Interface praticien - nouvelles technologies en orthopédie dento faciale : apport des sciences cognitives

Masrour Makaremi

▶ **To cite this version:**

Masrour Makaremi. Interface praticien - nouvelles technologies en orthopédie dento faciale : apport des sciences cognitives. Ingénierie biomédicale. Université de Bordeaux, 2022. Français. NNT : 2022BORD0195 . tel-03771549

HAL Id: tel-03771549
https://theses.hal.science/tel-03771549

Submitted on 7 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

université de BORDEAUX

THÈSE PRÉSENTÉE POUR OBTENIR LE GRADE DE

DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE SP2

SPÉCIALITÉ : SCIENCES COGNITIVES

Par Masrour MAKAREMI

Interface praticien – nouvelles technologies en orthopédie dento-faciale : apport des sciences cognitives

Sous la direction de Bernard N'KAOUA

Soutenue le 23/06/2022

Membres du jury :

M. ALLAIN Philippe, Professeur à l'Université d'Angers, Rapporteur

M. BORAUD Thomas, Directeur de recherche au C.N.R.S, Président du Jury

M. DJAFARI Ali, Directeur de recherche émérite au C.N.R.S, Invité

Mme MACKAY Wendy, Professeur au Collège de France, Membre du jury

M. N'KAOUA Bernard, Professeur à l'Université de Bordeaux, Directeur de thèse

M. SOREL Olivier, Professeur à l'Université de Rennes, Rapporteur

Résumé :

Cette thèse aborde, sous l'angle des sciences cognitives, la question de l'optimisation de l'interface entre les nouvelles technologies et le praticien en orthopédie dento-faciale. Elle interroge notamment l'impact de la transition numérique et digitale sur la perception des données paracliniques par l'orthodontiste et, par conséquent, sur le processus de prise de décision. Notre travail s'articule autour de deux axes :

Le premier étudie l'impact de la transition numérique sur la perception et l'utilisation par l'orthodontiste de modèles virtuels en 3D des arcades dentaires.

L'analyse de l'estimation des distances par les orthodontistes sur des modèles numériques en 3D des arcades dentaires a mené à au moins trois constats :

- que les distances mesurées sur les modèles virtuels ont été surestimées ;
- que la performance dans l'estimation des distances est moins bonne sur les modèles virtuels que sur les modèles réels ;
- que la charge mentale induite par l'estimation des distances sur des modèles orthodontiques est significativement plus élevée lorsque l'opération est réalisée sur des modèles virtuels que sur des modèles réels.

Le second axe explore la manière dont l'interprétabilité des réseaux de neurones convolutifs peut contribuer à une meilleure compréhension des interactions entre les différentes régions anatomiques crânio-faciales en fonction de la sévérité de la rétrognathie mandibulaire. Notre étude confirme que le morphotype pathologique de la rétrognathie mandibulaire est étendu au-delà de la mandibule mais également que, au travers de l'amplitude de la rétrognathie mandibulaire, une cinétique croissante du score de classification et de la surface discriminante apparaît sur la surface du crâne.

Mots clés :

Environnement numérique, réalité virtuelle, expertise perceptive, réseau de neurones convolutifs, interprétabilité, orthopédie dento-faciale, orthodontie.

Abstract:

This thesis addresses from the perspective of cognitive science, the issue of optimizing the interface between new technologies and the dentofacial orthopedic practitioner.

In particular, it questions the impact of the digital transition on the perception of paraclinical data by the orthodontist and, consequently, on the decision-making process. Our work is articulated around two axes :

The first one studies the impact of the digital transition on the orthodontist's perception and use of 3D virtual models of dental arches.

Analysis of orthodontists' estimation of distances on 3D digital models of dental arches led to at least three findings:

- distances measured on the virtual models were overestimated.
- performance of distance estimation is worse on virtual models than on real models.
- mental load of distance estimation on orthodontic models is significantly higher when the operation is performed on virtual models than on real models.

The second axis explores how the interpretability of convolutional neural networks can contribute to a better understanding of interactions between different craniofacial anatomical regions depending on the severity of mandibular retrognathia. Our study confirms that the pathological morphotype of mandibular retrognathia is extended beyond the mandible but also that, through the amplitude of mandibular retrognathia, increasing kinetics of the classification score and discriminative surface appear on the skull surface.

Key words:

Digital environment, virtual reality, perceptual expertise, convolutional neural network, interpretability, dentofacial orthopedics, orthodontics.

A la mémoire de ma mère Fetemeh Zarei.

La fatalité du destin ne nous a accordé qu'une rencontre de cinq minutes lorsque j'étais un petit garçon et toi, une prisonnière politique, mais son intensité, la force de tes paroles et l'exemple que tu nous as laissé par ton action m'ont donné le courage de déplacer des montagnes. Que cette thèse en soit un témoignage.

Table des matières :

1 : Introduction

2 : Interface orthodontiste-nouvelles technologies : pourquoi une thèse en sciences cognitives ?

2-1 Mon cheminement vers les sciences cognitives

2-2 Apport des sciences cognitives à l'optimisation de l'interface entre le praticien et les nouvelles technologies en orthodontie.

2-2-1 Introduction

2-2-2 L'orthodontie numérique et digitale

2-2-3 Les sciences cognitives

2-2-4 Conflit de représentation

2-2-5 Conclusion

2-2-6 Rôle de l'article dans ma démarche : ma rencontre avec les sciences cognitives

2-3 Les enjeux d'une nouvelle interface

2-3-1 Introduction

2-3-2 La complexité

2-3-3 Voies de connexion de l'expertise algorithmique et humaine

2-3-4 Conclusion

2-3-5 Rôle de l'article dans ma démarche : une conviction quant à l'apport des sciences cognitives

2-4 Le workflow numérique au service de l'apprentissage clinique du praticien

2-4-1 Introduction

2-4-2 Décryptage du workflow

2-4-2.1 Le diagnostic

2-4-2.2 Planification thérapeutique

2-4-2.3 Fabrication sur mesure de l'outil thérapeutique

2-4-2.4 Suivi thérapeutique

2-4-3 Les théories de neurosciences cognitives dans l'apprentissage

2-4-3.1 L'inférence bayésienne

2-4-3.2 Les 4 piliers de l'apprentissage

2-4-3.2.1 L'attention

2-4-3.2.2 L'engagement actif

2-4-3.2.3 Le retour d'information

2-4-3.2.4 La consolidation

2-4-4 Les 4 piliers de l'apprentissage au sein du flux de travail numérique

2-4-4.1 Diagnostic/Attention

2-4-4.2 Planification/engagement actif

2-4-4.3 Suivi thérapeutique/retour sur erreur

2-4-4.4 La consolidation

2-4-5 Conclusion

2-4-6 Rôle de l'article dans ma démarche : nécessité d'aller plus loin avec les sciences cognitives

2-5 La perception de l'expert :

2-5-1 Introduction

2-5-2 « Expert » et « expertise »

2-5-2.1 Définition du terme « expert »

2-5-2.2 Domaines d'expertise

2-5-2.3 Mécanisme cognitifs et neurobiologique et l'expertise

2-5-2.4 Le concept de chunking

2-5-3 Expertise perceptive visuelle

2-5-3.1 La notion d'expertise perceptive visuelle

2-5-3.2 Plasticité cérébrale et expertise visuelle

2-5-3.3 Perception et fiabilité de l'expert dans un environnement complexe

2-6 L'essentiel à retenir

2-7 Conséquence logique : une thèse en sciences cognitives

3 : La réalité virtuelle

3-1 Définition de la réalité virtuelle

3-2 Perception de la distance et réalité virtuelle

3-2-1 Perception des distances dans le monde réel

3-2-2 Perception des distances en réalité virtuelle
3-2-3 Evaluation des distances au sein de dispositifs immersifs
3-2-4 Evaluation des distances au sein de dispositifs non-immersifs
3-2-4.1 Les techniques de visualisation basées sur la physique
3-2-4.2 Les techniques de visualisation illustratives
3.3 Réalité virtuelle, charge mentale et complexité de la tâche
3-4 Application de la réalité virtuelle dans les spécialités médicales
3-4-1 Utilisation de la réalité virtuelle en médecine
3-4-2 Réalité virtuelle et orthodontie
3-5 Conclusion
3-6 Essentiel à retenir

4 : Estimation des distances en 3D par l'orthodontiste utilisant des modèles numériques

4-1 Objectif de l'étude
4-2 Matériel et méthode
4-2-1 Matériel
4-2-2 Analyses statistiques
4-3 Résultats
4-4 Discussion
4-5 L'essentiel à retenir

5 : Performance et charge mentale de l'orthodontiste dans l'estimation des distances sur modèles réels et virtuels

5-1 Contexte et intérêt de l'étude
5-2 Matériel et méthode
5-2-1 Sujets
5-2-2 Matériel
5-2-2.1 Modèles d'arcade dentaire
5-2-2.2 Questionnaires

5-2-3 Procédure
5-2-4 Analyse statistique
5-3 Résultats
5-3-1 Charge mentale
5-3-2 Estimations des distances
5-4-2.1 Encombrement mandibulaire
5-3-2.2 Distance inter-canine
5-3-2.3 Distance inter-molaire
5-4 Discussion
5-5 Conséquences cliniques des évaluations
5-6 Conclusion
5-7 Essentiel à retenir

6 : Apport de l'interprétabilité des réseaux de neurones convolutionnels (CNN) à une meilleure perception des dysmorphoses par l'orthodontiste

6-1 Vision par ordinateur et réseaux neuronaux convolutionnels (CNN)
6-1-1 Définition
6-1-2 Fonctionnement
6-1-2.1 Les neurones artificiels
6-1-2.2 Apprentissage et classification
6-1-2.3 CNN et représentation interne
6-1-2.4 « Black Box Problem » et interprétabilité
6-2 Application en imagerie médicale
6-2-1 Principales applications de la vision par ordinateur en médecine
6-2-2 Apport de l'interprétabilité des réseaux neuronaux convolutionnels dans le domaine médical
6-3 Application des CNN en orthopédie dento-faciale (ODF)
6-4 Détermination du stade de maturation vertébrale sur téléradiographie de profil par l'apprentissage d'un réseau neuronal multicouche
6-4-1 Introduction
6-4-2 Matériel et méthode

6-4-3 Résultats

6-4-4 Discussion

6-5 Utilisation de l'interprétabilité du deep learning pour l'extraction d'informations morphologiques en imagerie médicale : Cas d'applications pour la rétrognathie mandibulaire

6-5-1 Introduction

6-5-2 Méthodes

6-5-2.1 Aperçu des méthodes

6-5-2.2 Ensemble de données

6-5-2.3 Implémentation du réseau de classification

6-5-2.4 Prétraitement des images

6-5-2.5 Précision du modèle

6-5-2.6 Carte d'activation : Technique de superposition

6-5-2.7 Comparaison par zones

6-5-3- Résultats

6-5-3.1 Analyse descriptive

6-5-3.2 Analyse statistique

6-5-4 Discussion

6-5-4.1 Extraction d'informations morphologiques

6-5-4.2 Impact de la rétrognathie mandibulaire sur l'architecture cranio-faciale

6-5-4.3 Implication anthropologiques

6-5-5 Conclusion

7 : Discussion générale

7-1-1 Flux de travail numérique et sens clinique

7-1-2 Base cognitive de l'expertise

7-1-3 Étude de l'estimation des distances sur des modèles d'arcades dentaires en réalité virtuelle

7-2-1 Vision par ordinateur et expertise médicale

7-2-2 Elaboration d'un CNN de classification de maturation osseuse

7-2-3 Réseau de neurones convolutionnels et expertise médicale

8 : Conclusion

Liste des Figures et Tableaux

Bibliographie

Annexes

Annexes-1 :

- Makaremi, M., & N'Kaoua, B. (2021). Estimation of Distances in 3D by Orthodontists Using Digital Models. *Applied Sciences*, *11*(18), 8285.

Annexes-2 :

- Estimation of distances within real and virtual dental models as a function of task complexity.(en cours de soumission).

Annexes-3 :

- Makaremi, M., Lacaule, C., & Mohammad-Djafari, A. (2019). Deep learning and artificial intelligence for the determination of the cervical vertebra maturation degree from lateral radiography. *Entropy*, *21*(12), 1222.

« Plus la technologie progresse, plus les gens s'intéressent aux possibilités du seul esprit humain. » Isaac Bashevis Singer

1. Introduction

Les progrès des sciences numériques ont généré un changement de paradigme dans de très nombreuses disciplines appliquées : la transformation numérique et son corollaire, la transformation digitale, ayant amené l'homme à coexister étroitement avec la machine dans des domaines où, jadis, il régnait en maître. En effet, le mot « technologie » tire son origine du grec ancien *τέχνη* (la « tekhnè »), qui, selon la conception aristotélicienne, signifie « une capacité générale à fabriquer des choses avec intelligence » (Bruzina, 1982). Pour ce qui est de l'artisan, elle désigne son aptitude à envisager des formes particulières et à déployer ses compétences manuelles et son acuité visuelle pour les réaliser.

Pour que le rôle que les nouvelles technologies sont appelées à tenir soit un véritable progrès, tous les acteurs impliqués doivent avoir conscience de la nécessité de s'inscrire dans une stratégie d'augmentation résolument distincte d'une stratégie d'automatisation. L'enjeu est de prendre le contre-pied de l'automatisation, qui vise à assigner à des ordinateurs, une à une, des tâches traditionnellement effectuées par des humains dès que celles-ci peuvent être codifiées. Adopter une stratégie d'augmentation est un gage d'un progrès digne de ce nom : car en prenant en considération le fait que les capacités techniques peuvent atteindre des niveaux de performance nettement supérieurs chez les experts, et en valorisant cette performance, elle rend possible une synergie aussi fructueuse que possible entre les humains et les outils numériques.

L'orthopédie dento-faciale est une spécialité médicale fortement impactée par les progrès des sciences numériques et digitales. En particulier, l'imagerie tridimensionnelle (ci-après : imagerie en 3D) et l'intelligence artificielle génèrent de nouvelles données paracliniques à partir d'un patient virtuel : le set up virtuel dynamique, véritable modélisation tridimensionnelle (ci-après : modélisation en 3D) des objectifs thérapeutiques, et l'impression tridimensionnelle (ci-après : impression en 3D) permettent, à travers les procédés CAD/CAM, la fabrication des outils thérapeutiques individualisés suppléant la compétence manuelle du praticien. L'optimisation de l'interface entre l'orthodontiste et les nouvelles technologies est donc le défi majeur posé dans ce domaine : permettre au praticien d'engager utilement l'ensemble de son expertise dans le cadre d'une stratégie d'augmentation.

Cette thèse aborde, sous l'angle des sciences cognitives, la question de l'optimisation de l'interface entre les nouvelles technologies et le praticien en orthopédie dento-faciale. Elle

interroge notamment l'impact de la transition numérique et digitale sur la perception des données paracliniques par l'orthodontiste et, par conséquent, sur le processus de prise de décision. Notre travail s'articule autour de deux axes :

- Le premier étudie l'impact de la transition numérique sur la perception et l'utilisation par l'orthodontiste de modèles virtuels en 3D des arcades dentaires (Chapitres 2 à 4).
- Le second explore la manière dont l'interprétabilité des réseaux de neurones peut contribuer à étendre le champ de perception de l'orthodontiste (Chapitre 5).

Le premier chapitre expose la réflexion qui m'a amené à une double prise de conscience : premièrement, conscience du défi que nous devrons relever en tant que cliniciens – permettre à nos cerveaux de travailler avec l'outil numérique tout en conservant notre intuition ; deuxièmement, conscience du fait que les sciences cognitives sont notre meilleur allié dans ce processus. Ce chapitre retrace ce chemin initiatique tel qu'il se déploie à travers trois articles académiques. Il est complété par une revue de la bibliographie sur l'expertise démontrant que cette dernière repose davantage sur des processus holistiques que sur des processus analytiques.

Le second chapitre est une analyse de la littérature spécialisée dans ce domaine : les modèles numériques en 3D des arcades dentaires sont une réalité virtuelle non immersive. La littérature de la réalité virtuelle y est analysée sur trois plans :

- l'estimation des distances : la perception de la distance dans un milieu virtuel diffère de celle qui a lieu dans un milieu réel ;
- la charge mentale lors d'interaction avec la réalité virtuelle ;
- enfin, un état de l'art sur la place de la réalité virtuelle au sein des différentes spécialités médicales, et tout particulièrement, de l'orthopédie dento-faciale.

Le troisième chapitre décrit la première expérience réalisée dans le cadre de cette thèse et qui porte sur l'estimation des distances en 3D par l'orthodontiste à l'aide de modèles numériques. En effet, l'estimation des distances est au cœur de la pratique diagnostique de l'orthodontiste et le passage d'estimations réalisées sur des modèles réels à des estimations réalisées sur des modèles en 3D imposent de bien comprendre les différences entre les estimations réalisées dans ces deux environnements. Les résultats obtenus ont montré que 4 des 7 variables testées

étaient associées à un manque de précision dans l'estimation des distances, la taille de l'arc mandibulaire étant la variable la moins bien estimée.

Consacré à la seconde expérience, le quatrième chapitre analyse la performance et la charge mentale de l'orthodontiste dans l'estimation des distances tant sur des modèles réels que sur des modèles virtuels. La charge mentale de l'estimation des distances sur des modèles orthodontiques est significativement plus élevée lorsque l'opération est réalisée sur des modèles virtuels que sur des modèles réels. En revanche, la performance est significativement moins bonne dans ce même cas.

Le cinquième chapitre aborde le second axe de cette thèse : la manière dont l'interprétabilité des réseaux de neurones peut contribuer à étendre le champ de perception de l'orthodontiste. Il comporte deux parties : a) la première propose une revue de la bibliographie sur la vision assistée par ordinateur, l'interprétabilité des réseaux de neurones et leur application en médecine et particulièrement en orthopédie dento-faciale ; b) la deuxième présente nos travaux sur l'apport de l'interprétabilité des réseaux de neurones convolutionnels (CNN, pour Convolutional Neural Netwoks) à la recherche de corrélations entre la rétrognathie mandibulaire et l'architecture cranio-faciale. Cette étude confirme que le morphotype pathologique de la rétrognathie mandibulaire est étendu au-delà de la mandibule mais également que, au travers de l'amplitude de la rétrognathie mandibulaire, une cinétique croissante du score de classification et de la surface discriminante apparaît sur la surface du crâne.

2 : Interface Orthodontiste-nouvelles technologies : pourquoi une thèse en sciences cognitives ?

2-1 Mon cheminement vers les sciences cognitives

« C'est toujours par hasard qu'on accomplit son destin. » (Marcel Hachard)

Dès le début de ma pratique médicale, j'étais l'archétype d'un « early adopter », à savoir un individu qui a pour habitude d'acheter quasi systématiquement les nouveaux produits dans une catégorie de produit donnée (pour ce qui me concerne, les dispositifs médicaux orthodontiques à la fine pointe de la technique).

De même que, en tant que simple consommateur, en 2007 j'avais fait la queue à 6 h du matin devant une boutique France Télécom pour être parmi les premiers à acquérir l'iPhone1 que j'attendais depuis des mois, en 2011 j'étais l'un des premiers praticiens à acquérir une caméra intra-orale afin d'obtenir des empreintes tridimensionnelles des arcades dentaires de mes patients et des modèles d'étude dentaires indispensables dans le diagnostic des malocclusions. Cette transition des modèles d'étude réels vers des modèles virtuels me semblait ne présenter que des avantages : outre d'assouvir mon appétence pour les nouvelles technologies, l'effet bénéfique en termes d'image de marque pour mon cabinet, d'expérience pour le patient et de plus-value logistique dans l'organisation et le stockage des empreintes me paraissait évident. Pourtant, malgré mon enthousiasme, j'éprouvais un sentiment de plus en plus caractérisé à l'égard du recours aux empreintes numériques : mon diagnostic orthodontique à partir de modèles numériques était moins performant que lorsqu'il était effectué sur la base de modèles réels. Et je sentais bien que ce n'était pas qu'une question de familiarisation avec ce nouvel outil. J'avais le sentiment que la virtualisation des données cliniques avait en quelque sorte déstabilisé mon sens clinique, et, pour la première fois, mon attrait pour la modernité et les nouvelles technologies se retrouvait en contradiction avec mon éthique : offrir à mes patients les meilleurs soins possibles. Je faisais face à un dilemme, qui plus est désarmé, sans savoir vers qui me retourner. En pleine réflexion sur cette problématique et, plus globalement, sur les impacts de la transition numérique sur nos sociétés, un soir, j'ai croisé au hasard de mes recherches le podcast d'un entretien du philosophe Michel Serres, qui, en 2007, lors d'une conférence prononcée à l'occasion des 40 ans de l'INRIA, insistait sur l'ingéniosité et sur la créativité que les nouvelles générations devraient déployer pour que leur cerveau puisse

travailler efficacement avec l'outil numérique : « les nouvelles technologies vont obliger les humains à devenir intelligents ».

Cette conférence de Michel Serres a suscité en moi un déclic, me faisant comprendre que le défi que nous devrions relever en tant que cliniciens était de permettre à nos cerveaux de travailler avec l'outil numérique tout en conservant notre intuition. Pour cela il fallait repenser notre rapport aux nouvelles technologies, en plaçant au centre de notre réflexion l'expertise humaine et non pas l'innovation technologique.

Mais il ne suffit pas de vouloir ce changement de paradigme pour qu'il se réalise, il faut avoir les outils pour mieux comprendre les interactions du praticien avec les nouvelles technologies dans le but de les améliorer. C'est alors que j'ai découvert les sciences cognitives et compris que cette discipline serait l'allié idéal pour répondre à ce défi. Ce chapitre retrace ce chemin initiatique à travers les articles que j'ai eu la chance de publier ces cinq dernières années.

Afin que ma démarche soit pleinement appréhendée par le lecteur, j'ai aussi fait le choix de laisser ces articles dans leur schéma originel, sans y apporter aucune modification, ce qui a abouti à quelques répétitions de texte et de figures çà et là.

Le premier de ces articles, qui illustre ma prise de conscience des problématiques exposées précédemment, a été rédigé juste avant le début de mon Master 2 en sciences cognitives. Le deuxième, rédigé après la finalisation de ce Master 2, montre comment une réflexion au départ très personnelle et hésitante s'est renforcée au fil du cursus que je venais d'achever. Le troisième article a été rédigé dans le cadre de ma thèse de doctorat, et témoigne de l'avancement de ma réflexion au contact des travaux scientifiques constituant une bibliographie considérablement enrichie.

Au cours de ma réflexion et de ma recherche bibliographique, ma vision s'est progressivement élargie d'interrogations inhérentes à ma spécialité, ma réflexion s'est portée sur l'interaction de l'expert avec la transition numérique et la virtualisation de ses outils de travail. Le quatrième chapitre relate le travail bibliographique qui a été mené sur le fonctionnement cognitif de l'expert. La conclusion de ces cinq années de réflexion synthétisées dans ce chapitre est que, si la maitrise de l'outil numérique par l'homme dans l'optique d'une stratégie d'augmentation est une évidence, sa mise en œuvre est un défi qui peut être relevé de manière optimale à l'aide des sciences cognitives. En effet, réussir une stratégie d'augmentation suppose et nécessite la connexion de l'intuition de l'expert au flux de travail numérique.

2-2 Apport des sciences cognitives à l'optimisation de l'interface entre le praticien et les nouvelles technologies en orthodontie.

Référence bibliographique de l'article :

Makaremi, M. (2016). Apport des sciences cognitives à l'optimisation de l'interface entre le praticien et les nouvelles technologies en orthodontie. *Revue d'Orthopédie Dento-Faciale*, *50*(3), 335-343.

2-2-1 Introduction

Au cours des dernières décennies, nos habitudes ont été bouleversées par la naissance de l'orthodontie numérique et digitale (Vassura et al., 2010). Nul doute que nous sommes aujourd'hui à l'aube d'une révolution. L'avènement de l'imagerie, des empreintes et de l'impression tridimensionnelles, l'expression de l'intelligence artificielle à travers les algorithmes de proposition thérapeutiques (ex. : set-up virtuel tel que clean check R) et le développement des procédés industriels de fabrication individualisée des appareils orthodontiques, génèrent de profondes mutations. Celles-ci interviennent dans un contexte de transformation de la pratique médicale (Alexandre, 2011) ; au-delà, c'est toute la société que nous voyons se métamorphoser au contact des nouvelles technologies numériques (Dervin et Abbas, 2009). Devant l'intensité de ces changements, se pose légitimement la question de l'identité du praticien et de sa place dans le nouveau cosmos orthodontique. Si l'on ne tient compte que de son rôle technique, ces bouleversements peuvent être vus comme une remise en cause du spécialiste, les progrès techniques se faisant au détriment de la valorisation de l'humain. On peut alors se demander, non sans anxiété, à quoi servira l'orthodontiste dans les schémas thérapeutiques du futur.

Mais un tout autre regard peut être porté sur ces mutations. Les propos de l'écrivain Isaac Bashevis Singer, évoquant la littérature, l'éclairent (Lewi, 2001) : ce regard consisterait à ne pas croire que l'orthodontiste ait quoi que ce soit à craindre de la technologie. Au contraire, plus la technologie progressera, plus les gens s'intéresseront aux possibilités du seul esprit humain. Sur ce chemin, qui reconnait une place centrale au praticien dans l'orthodontie du futur, les sciences cognitives peuvent être un précieux allié, grâce aux outils indispensables que cette discipline fournit pour décrypter et optimiser l'apport de l'intellect humain dans les schémas thérapeutiques du troisième millénaire.

2-2-2 L'orthodontie numérique et digitale

Le mérite de la conception d'un appareillage orthodontique sur mesure à l'aide de la technologie CAD/CAM (Computer-Aided Design/Computer-Aided Manufacturing) revient à Graig Andreiko (1994) qui, dans les années 1990, dans le cadre du projet Elan, développe pour Ormco les premiers appareillages.

Dès le début, son concepteur a conscience de la difficulté à appliquer des normes, compte tenu de la grande variabilité observable chez l'homme anatomiquement moderne (Andreiko, 1994).

Cette démarche thérapeutique s'inscrit dans une philosophie de médecine personnalisée : une médecine établie sur les caractérisations fines du patient, utilisant des moyens de modélisation individuels (Mathieu et al., 2013).

Dans les années 2000, plusieurs starts up, dont l'entreprise Align Technologie devenue depuis une multinationale, démocratisent les appareillages sur mesure conçus par des procédés industriels (Phulari, 2013).

La modélisation du projet thérapeutique et de l'appareillage orthodontique est basée sur un algorithme fermé qui détermine l'objectif thérapeutique et le chemin des déplacements dentaires pour y parvenir (figure 1). Elle s'appuie donc sur un procédé d'intelligence artificielle, dans lequel un logiciel informatique propose la solution thérapeutique (ex. : le clean check) après qu'un technicien ait traduit en langage informatique les données paracliniques (en particulier, les modèles d'études avec segmentation des dents présentes sur l'arcade) et les instructions. Le set-up virtuel est ensuite validé par le praticien, et l'appareillage orthodontique fabriqué par procédé industriel puis livré au praticien qui assure le suivi thérapeutique.

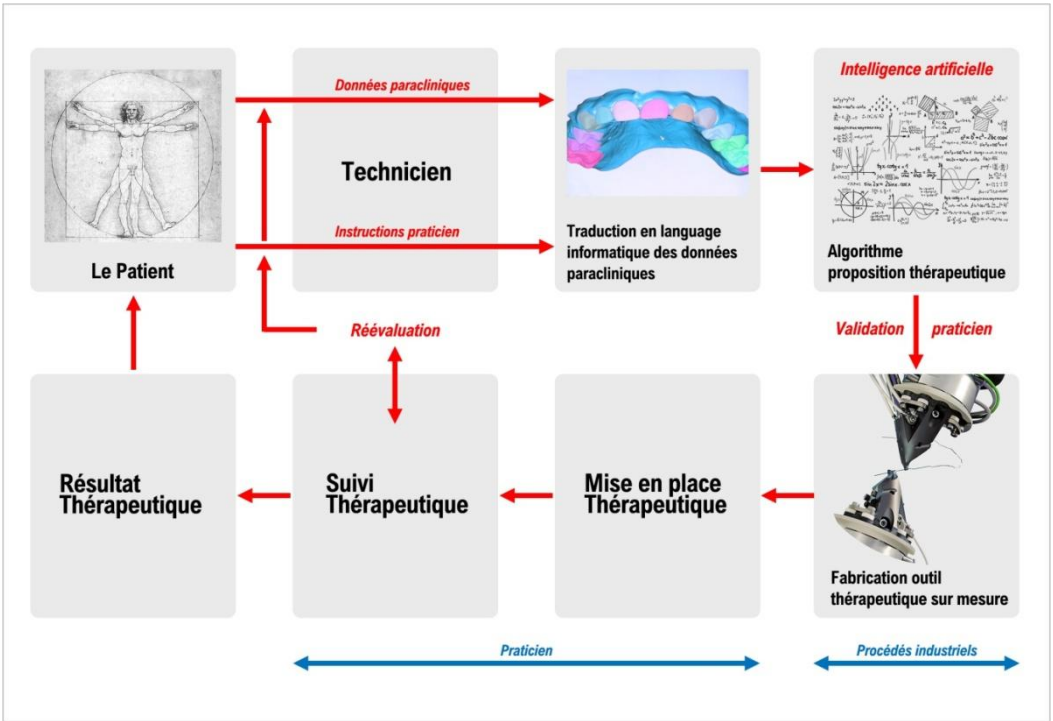


Figure 1 : Conception actuelle de l'outil thérapeutique orthodontique selon le procédé CAD/CAM.

La genèse du projet thérapeutique s'effectue donc à partir d'un algorithme basé sur des normes préétablies. Publié bien avant l'avènement des nouvelles technologies, l'analyse de Tweed (Tweed, 1966) ou les chevrons de Steiner (Steiner, 1969) peuvent également être perçus comme des algorithmes, soit une suite finie et non-ambiguë d'opérations ou d'instructions permettant de résoudre un problème.

Aujourd'hui comme hier, la question essentielle est donc la confiance que nous avons dans les normes pour établir notre plan de traitement (Philippe, 2013) : si nous considérons qu'un algorithme fermé peut trouver l'optimum individuel (Vassura et al., 2001), alors nous pouvons confier à l'intelligence artificielle la conception du projet thérapeutique. Mais si nous avons la conviction qu'il n'y a pas de formule mathématique magique permettant d'aboutir à l'optimum individuel, si nous croyons que ce dernier ne s'obtient qu'au prix d'une symbiose entre l'application des données et normes scientifiques et l'expression du ressenti clinique du praticien, alors immanquablement, naît le désir de faire progresser la symbiose entre la puissance des outils numériques, digitaux, de l'intelligence artificielle et la finesse de la perception clinique du praticien (Figure 2), afin que le projet thérapeutique soit le véritable fruit

de cette union : profiter des progrès des sciences numériques tout en gardant l'humain au centre.

Figure 2 : Le praticien doit connecter le patient à l'outil thérapeutique issu des nouvelles technologies.

2-2-3 Les sciences cognitives

Les sciences cognitives regroupent un ensemble de disciplines scientifiques dédiées à la description, à l'explication et, le cas échéant, à la simulation des mécanismes de la pensée humaine ou artificielle, et plus généralement de tout système complexe de traitement de l'information capable d'acquérir, conserver, utiliser et transmettre des connaissances (Grin, 2003). Cette discipline repose donc sur l'étude et la modélisation de phénomènes aussi divers que la perception, l'intelligence, le langage, la mémoire, l'attention, le raisonnement, les émotions ou même la conscience (Miller, Galanter et Pribram, 1960). Nées dans les années 1950, elles forment aujourd'hui un champ interdisciplinaire très vaste, dont les limites et le degré d'articulation des disciplines constitutives font toujours débat (Dupuy, 2013).

Les sciences cognitives utilisent conjointement des données issues d'une multitude de branches de la science et de l'ingénierie, comme la linguistique, l'anthropologie, la psychologie, les

neurosciences, la philosophie, l'intelligence artificielle. Ces différentes disciplines, comme le montrent les exemples cités, peuvent constituer un socle pour décrypter la place du praticien dans la thérapeutique orthodontique et son interface avec les nouvelles technologies numériques (figure 3).

La maîtrise linguistique est essentielle dès lors qu'il y a externalisation du procédé de fabrication de l'outil thérapeutique sur mesure.

L'intelligence artificielle est un nouveau paramètre de notre pratique orthodontique ; il nous faut en prendre pleinement conscience et en définir les contours.

L'étude de la dimension psychologique est fondamentale dans la relation avec le patient, mais également pour appréhender notre abord des nouvelles technologies, comprendre la fascination ou le rejet qu'elles peuvent générer. Les neurosciences ont étudié le recrutement neuronal différentiel entre la perception virtuelle d'un objet (un modèle d'étude virtuel) et sa manipulation réelle (modèle en plâtre).

Enfin, un regard philosophique permet d'entreprendre une réflexion de fond sur des questions telles que l'interface homme-machine.

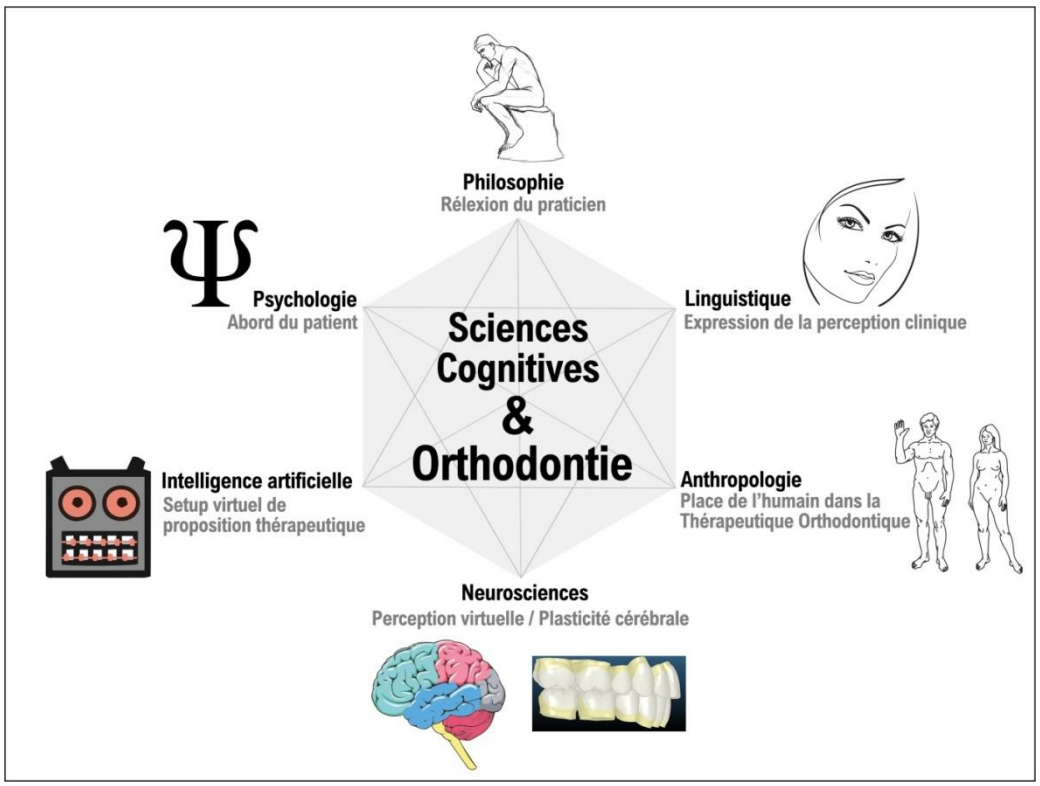


Figure 3 : Interactions entre les différentes disciplines constituant les sciences cognitives et l'orthodontie.

2-2-4 Conflit de représentation

La représentation mentale ou représentation cognitive est l'image qu'un individu se fait d'une situation. Elle est confluente des sensations et de la mémoire (Barsalou, 1999).
Comme toute activité humaine est organisée en vue d'une fin (Berthoz et Petit, 2006), la notion de représentation est proche du concept d'intentionnalité. La psychologie cognitive considère que toutes les représentations ne sont pas conscientes (Gallina, 2006).

Le concept de représentation cognitive a également été abordé dans le domaine de l'industrie, notamment pour les entreprises confrontées à des risques particulièrement importants (aviation, nucléaire, etc.). Des recherches ont été menées pour permettre aux opérateurs d'avoir une représentation mentale aussi proche que possible de la réalité. Ces recherches, fondées notamment sur la théorie de la communication (Falk et al., 2010), ont conduit à l'amélioration de l'ergonomie des interfaces homme-machine.

Lorsque le praticien détermine, à partir des données cliniques et paracliniques, un plan de traitement et une stratégie thérapeutique, il construit également une représentation mentale du schéma thérapeutique qui s'appuie sur son sens clinique. Ce n'est qu'à un stade d'orthodontiste confirmé que le praticien parvient à engrammer des représentations tridimensionnelles des mouvements intra- et inter arcades efficientes. Il les édifie à partir de sa sensibilité thérapeutique et de ses expériences cliniques passées.

Mais il entre également en jeu une composante de la personnalité dans la perception individuelle de la beauté et dans le ressenti de la demande du patient. Cette représentation intuitive est un précieux guide dans les comportements thérapeutiques et contribue à l'efficience des traitements.

Dans le schéma actuel de conception d'appareillages sur mesure (figure 1), l'industriel propose, en amont de la conception, une représentation virtuelle du projet thérapeutique (ex : clean check), construite à partir des données paracliniques et de quelques instructions, souvent succinctes, du praticien.

Ainsi, il existe un « conflit de représentation » entre la construction mentale que le praticien se fait du projet thérapeutique et le set-up virtuel, fruit de l'intelligence artificielle. Si les deux dessinent le chemin et les objectifs thérapeutiques d'une même situation clinique, l'une est l'œuvre de l'intellect du praticien et le second, le résultat d'une équation.

Si le praticien n'a pas fait l'effort de construire une représentation mentale du schéma thérapeutique avant de voir le set-up virtuel proposé par l'industriel, il lui sera très compliqué de l'élaborer a posteriori. Il devient alors le vassal du plan de traitement issu de l'intelligence artificielle. S'il a fait l'effort de le construire en amont, son adéquation avec le set-up virtuel est complexe. L'idéal est d'aboutir à une fusion entre ces deux représentations (figure 4), une symbiose qui permettrait d'utiliser les progrès du numérique tout en gardant la prédominance de la dimension humaine, indispensable à la pratique médicale et la totale maîtrise par le praticien du projet thérapeutique. Cet objectif nécessite de donner un rôle beaucoup plus actif à la cognition du praticien dans la construction du set-up virtuel. Il s'agit d'un projet ambitieux qui demande beaucoup plus d'efforts intellectuels et d'implication de la part du praticien mais également une plus grande volonté de coopération de la part des industriels. On entrerait alors pleinement dans une médecine individualisée et la qualité des traitements, ainsi que la valorisation du praticien, y gagneraient beaucoup.

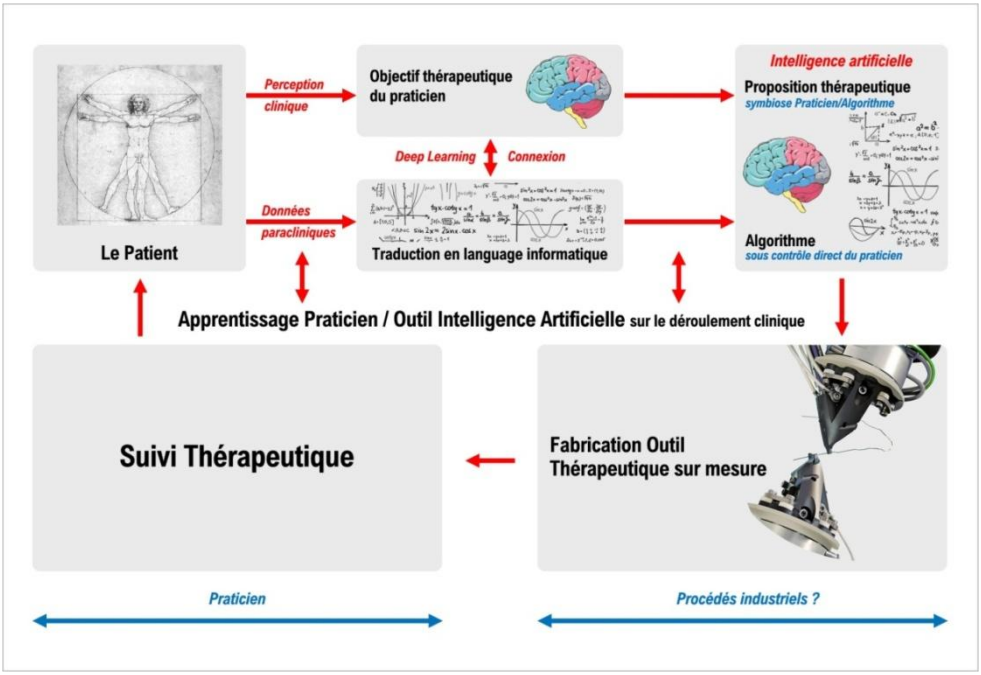


Figure 4 : Symbiose entre le praticien et l'intelligence artificielle (ci-après IA).

Aujourd'hui différents obstacles doivent être franchis pour fusionner les deux représentations et optimiser l'interface homme-machine :

Pour rendre les données cliniques et paracliniques lisibles aux logiciels informatiques, des manipulations techniques sont nécessaires, particulièrement la segmentation de l'arcade dentaire (Guang et al., 2014), afin que le logiciel puisse reconnaître et individualiser les dents. C'est un travail fastidieux qui relève de la compétence de techniciens, non du praticien. La relation entre les outils d'intelligence artificielle et l'orthodontiste est donc, dès le départ, scindée. Les progrès fulgurants dans le domaine du Deep Learning « apprentissage profond » (Ranzato et al., 2007), qui permettent la reconnaissance faciale ou la conduite de voiture sans chauffeur efficientes, peuvent nous faire espérer que dans un futur proche, le logiciel sera capable de traduire lui-même en langage informatique les données cliniques.

Un autre aspect qui appelle une attention particulière est la notion de perception. Différents travaux en neurosciences (Allami et al., 2008) ; (Decety, Jeannerod et Prablanc, 1989) ont montré que le recrutement neuronal et la plasticité cérébrale (figure 5) étaient différents, entre la manipulation virtuelle d'un objet à travers un écran d'ordinateur, et sa perception et sa manipulation concrètes. Nous devons chercher à progresser dans notre perception virtuelle des objets par un schéma d'apprentissage adéquat. Les paramètres intégrés dans la représentation du schéma thérapeutique constituent un autre biais de perception ; par exemple, lorsque nous observons un saut d'articulé de passage d'une Cl II à une Cl I s'effectuer sur le set-up virtuel de proposition thérapeutique, nous y voyons l'expression d'une stimulation de croissance alors qu'aucune donnée sur la croissance du patient n'a été soumise à la machine. C'est tout le paradoxe de la coexistence d'une double représentation.

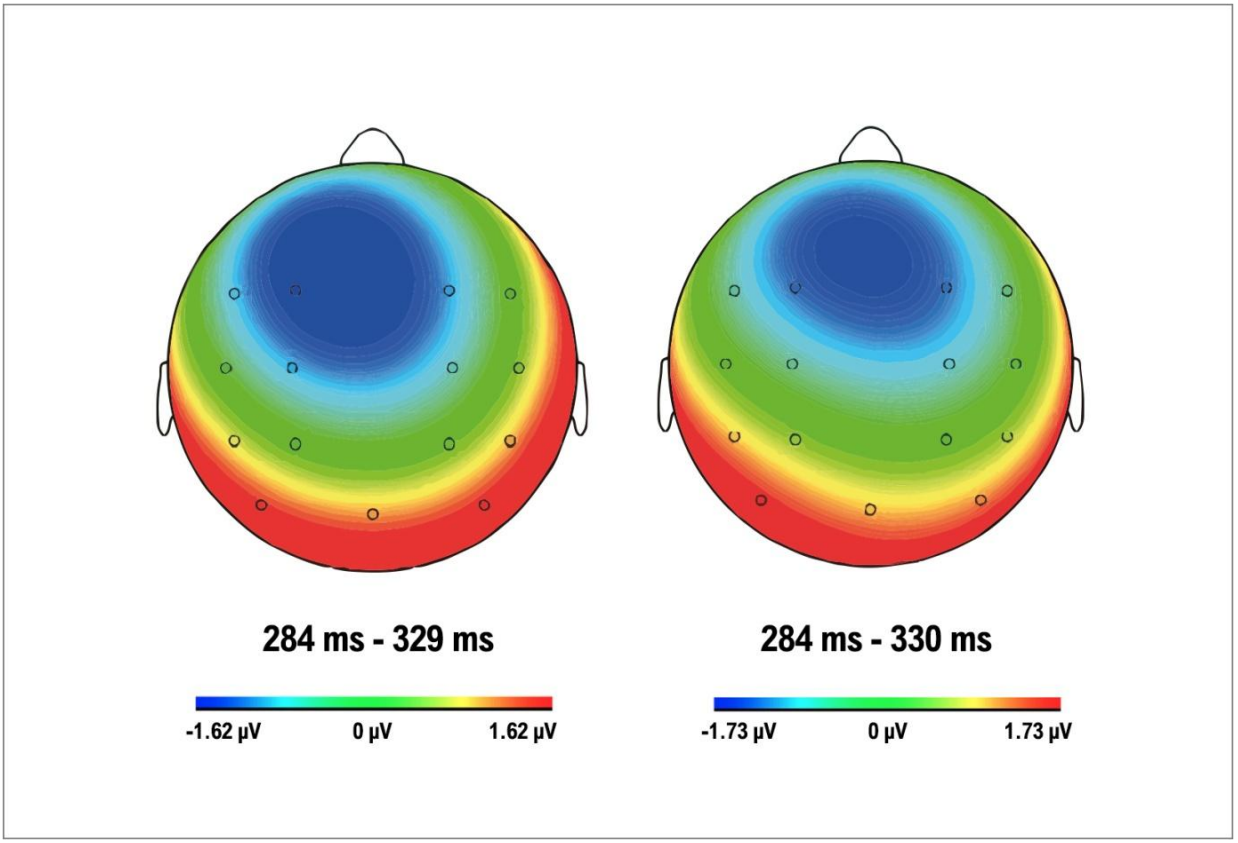


Figure 5 : Activité cérébrale différentielle lors d'une perception virtuelle (droite) et réelle (gauche) d'après Allami et al., 2008.

Le schéma actuel de conception des appareillages orthodontiques à partir de la technologie CAD/CAM a été conçu par les industriels. Il constitue un formidable progrès dans notre réponse thérapeutique. Légitimement, les industriels l'ont façonné pour qu'il s'adapte au mieux à leur procédé de fabrication. C'est notre responsabilité de praticien de conserver la maîtrise de la stratégie thérapeutique et d'apporter la réponse optimum tenant compte de la singularité de chaque individu. Nous pouvons le faire en coopération avec les industriels, en nous appuyant sur les sciences cognitives pour améliorer l'interface homme-machine, la connexion entre les outils d'intelligence artificielle et l'intellect du praticien.

Il est également opportun d'élargir notre horizon. Les fulgurants progrès de l'impression en 3D (Berchon et Luyt, 2014) laissent entrevoir l'espoir de redonner au praticien la maîtrise de la conception de l'appareillage orthodontique.

2-2-5 Conclusion

Faire retour sur l'un des actes fondateurs de l'intelligence artificielle permet d'illustrer l'importance de l'interface homme-machine. Durant la Seconde Guerre mondiale, Alan Turing, mathématicien anglais de génie (Turing et al., 1995), invente une machine dont l'objectif est de décrypter les codes de l'armée allemande (programme Enigma) (Turing et al., 1995). Dans un premier temps, il échoue car les calculs de la machine seuls n'ont pas la capacité de casser les codes. Il parvient à remplir sa mission que lorsqu'il combine la puissance de calcul de la machine avec un travail linguistique sur les messages, fait par des spécialistes en langue germanique, leur travail préliminaire guidant les calculs et augmentant la probabilité de la machine à décrypter les codes. Ainsi, pour Alan Turing, la réussite naît d'une bonne connexion entre l'intellect humain et l'intelligence artificielle. Cette connexion est fondamentale en médecine où l'on soigne l'homme, et tout particulièrement en orthodontie, où la perception de la beauté est éminemment subjective.

En conclusion, nous retiendrons que la symbiose cognitive du praticien et de l'intelligence artificielle permettant l'utilisation optimale des nouvelles technologies doit être notre objectif. Nous pouvons, dans cette quête, nous appuyer sur les sciences cognitives. Aujourd'hui, de nombreux obstacles demeurent et rendent le chemin difficile et incertain. Mais nous devons garder espoir et persévérer. Devant la révolution technologique que nous vivons, citons les propos de l'historien Robert Rosentone (Rosenstone, 1981) : « La révolution est une tentative pour faire aboutir les rêves. »

2-2-6 Rôle de l'article dans ma démarche : ma rencontre avec les sciences cognitives

L'écriture en 2016 de l'article ci-dessus qui a pour la première fois établi un lien entre l'orthopédie dento-faciale, les nouvelles technologies numériques et les sciences cognitives, alors que j'étais encore néophyte en matière de sciences cognitives, était un projet pour le moins ambitieux. Ce travail m'a amené à expliciter par écrit devant mes pairs que ce qui était de l'ordre de l'intuition m'a permis de m'immerger dans l'univers des sciences cognitives. Cet exercice m'a confirmé dans ma démarche. Mais j'ai pu également prendre la mesure de toute la diversité et la complexité des sciences cognitives, ainsi que de mes lacunes dans cette discipline : de simples connaissances de culture générale étaient loin de suffire à la réussite de

mon projet scientifique. C'est dans le but d'acquérir des bases solides en sciences humaine que j'ai entrepris un master 2 en sciences cognitives, alors que, titulaire d'un Master 2 en anthropologie biologique, j'aurais pu entamer directement une thèse en sciences cognitives, comme cela m'avait été proposé. Le second article relate l'expérience du Master 2 en sciences cognitives et la manière cette expérience m'a permis de renforcer ma profonde conviction quant au rôle fondamental de l'intuition du praticien dans la pratique de la médecine de l'avenir.

2-3 Les enjeux d'une nouvelle interface

Référence bibliographique de l'article :
Makaremi, M. (2018). Les enjeux d'une nouvelle interface. *Revue d'Orthopédie Dento-Faciale*, *52*(4), 319-324.

2-3-1 Introduction

Les progrès effectués dans les domaines de l'intelligence artificielle et des sciences informatiques auront un impact majeur sur la pratique de l'orthopédie dento-faciale au cours des décennies à venir. Ce constat fait l'objet d'un ample consensus, comme en témoigne les investissements colossaux et les fusions acquisitions en cours chez les principales industries spécialisées dans la thérapeutique orthodontique.
Devant l'intensité de ces changements, se pose légitimement la question de l'identité du praticien et de sa place dans ce nouvel environnement.

Les écrits du docteur Laurent Alexandre reflètent un courant de pensée qui s'apparente au néo-luddisme (Jones, 2013) prévalant de nos jours parmi le corps médical. Selon ce spécialiste des révolutions technologiques et de leurs enjeux, l'enrichissement incessant des compétences des robots comporte le risque majeur « qu'en 2030 le médecin soit subordonné à l'algorithme, comme l'infirmière l'est aujourd'hui au médecin » (Alexandre, 2017).
Tout en ayant le mérite d'amener la réflexion sur les enjeux de l'interface homme-machine (IHM) dans les pratiques médicales, les prédictions afférentes aux stratégies d'automatisation que les multinationales du numérique veulent appliquer au domaine de la santé et à

l'orthopédie dento-faciale en particulier sont génératrices de défiance et d'angoisse chez de nombreux praticiens.

La spécialité que j'exerce, est dans l'œil du cyclone : si l'on en croit le docteur Alexandre, elle est vouée à mourir à petit feu, aucun praticien n'arrivant « à la cheville » d'un système basé sur l'intelligence artificielle (IA).

J'ai donc voulu relever mon regard vers mon futur bourreau et, après avoir plongé au cœur du monde de l'IA par le biais d'un Master 2 de Neurosciences computationnelles, je n'ai jamais été aussi optimiste quant à la plus-value du clinicien. Aux antipodes du mythe de la singularité technologique (Ganascia, 2017), je suis persuadé que la bonne symbiose entre la médecine et l'IA est en quête d'une complémentarité (Autor, Katz et Krueger, 1998) mobilisant une stratégie d'augmentation résolument distincte d'une stratégie d'automatisation, pour reprendre les termes de Davenport et Kirby. Cette dernière « part des tâches de base qui constituent un poste donné, et vient les soustraire à travers le déploiement d'ordinateurs qui reprennent une à une ces tâches effectuées par des humains dès que celles-ci peuvent être codifiées » (Davenport et Kirby, 2016). Cependant, tout en assurant « des économies en termes de coûts », ce mouvement vers l'automatisation « circonscrit notre pensée dans un cadre délimité par les paramètres du travail tel qu'il est accompli aujourd'hui » (Davenport et Kirby, 2016). En revanche, l'augmentation « signifie que l'on part de ce que les humains font aujourd'hui pour arriver à comprendre la manière dont ce travail pourrait être approfondi, et non diminué, par une utilisation plus poussée des machines » (Davenport et Kirby, 2016).

Pour illustrer les implications possibles de chacune de ces stratégies dans le domaine de l'orthopédie dento-faciale, nous proposons l'exemple suivant. Une empreinte 3D des arcades dentaires est effectuée au moyen d'un scanner intra-oral est traitée par un algorithme qui va modéliser les arcades dentaires dans une position idéale en fonction de l'anatomie dentaire du patient. À partir de cette modélisation est ensuite conçu, à l'aide d'un procédé CAD/CAM (« computer-aided design », « computer-aided manufacturing »), un dispositif orthodontique sur mesure (bague/gouttière) dont la mise en place technique ne demande qu'une compétence clinique limitée. Dans ce schéma thérapeutique, la valeur ajoutée de l'orthodontiste, aussi bien technique qu'intellectuelle, se réduit, de prime abord, à un rôle minime.

2-3-2 La complexité

Pour que tout se déroule parfaitement suivant un procédé calibré au millimètre près, il faut faire abstraction du fait que les déplacements dentoalvéolaires ont lieu au sein de la face, carrefour de nombreuses fonctions oro-faciales (respiration mastication, déglutition, phonation, posture linguale…) ayant une incidence esthétique majeure dans un environnement biologique de tissu mou et dur (parodonte) réagissant aux contraintes biomécaniques de façon variable d'un individu à l'autre. À ces paradigmes viennent souvent se rajouter les paramètres de croissance faciale complexes et variables qu'il faut prendre en compte.

Tout bien considéré, au contraire de ce que l'orthodontiste peut parfois être amené à croire de lui-même, sa valeur et celle de ses dix années d'études ne résident pas seulement dans sa capacité à mettre en œuvre une procédure technique parfaitement codifiée et maitrisée, mais aussi dans son aptitude à effectuer un codage des valeurs des différents paramètres faciaux sur une échelle commune (établir une « monnaie commune » des différentes paramètres) (figure 6).

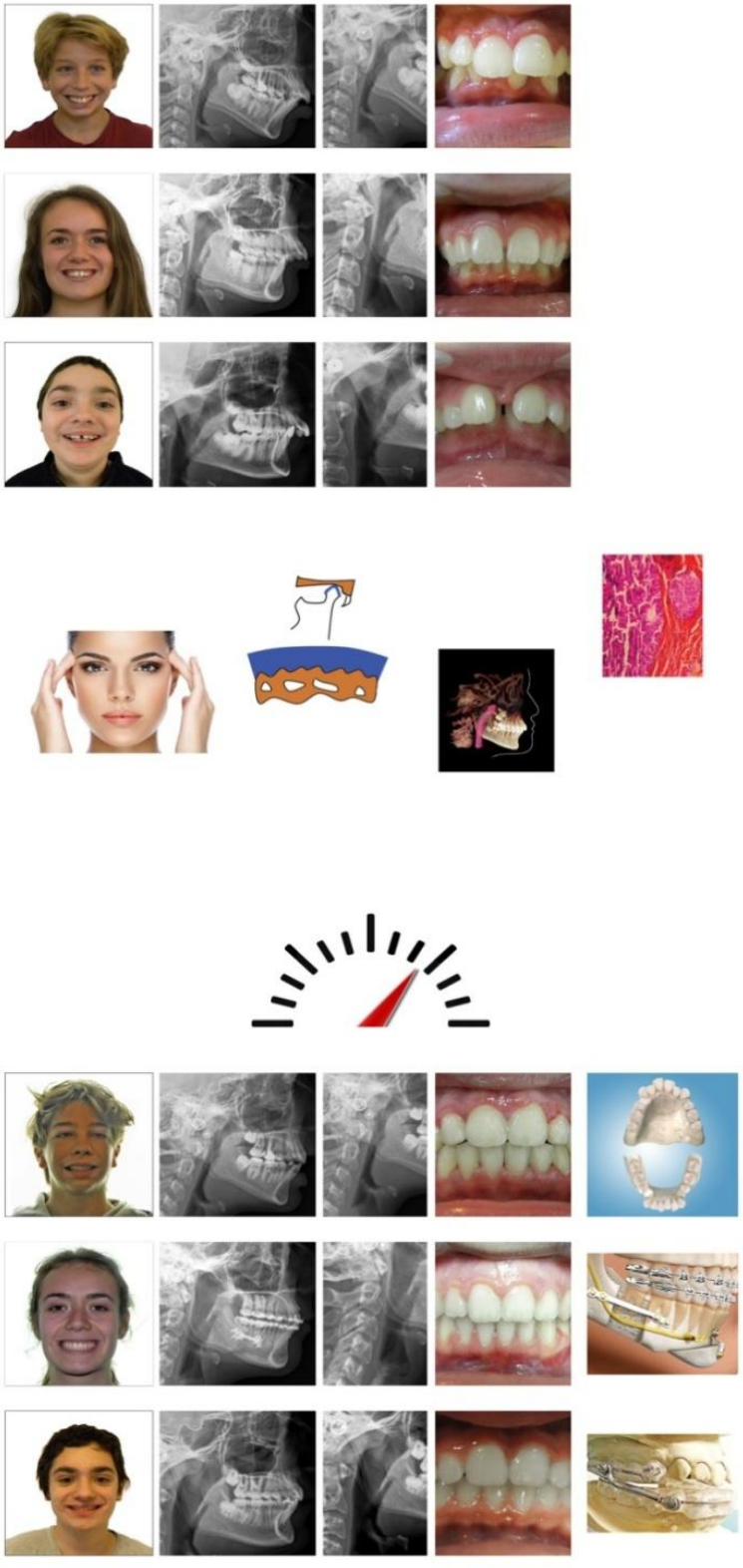


Figure 6 : C'est en créant une monnaie commune entre différents paramètres (esthétique, fonctions oro-faciales**,** croissance faciale) que le praticien parvient à l'optimum individuel pour le patient.

Par exemple, dans deux situations cliniques où le positionnement tridimensionnel des arcades est très similaire, les empreintes 3D seront identiques ; qui plus est, si nous ne faisons que suivre l'algorithme, les outils thérapeutiques individualisés seront également très proches. Or, si l'on se transpose dans le contexte biologique au sein de la face, les situations cliniques s'éloigneront l'une de l'autre. À travers son diagnostic, le praticien va placer sur la même échelle de valeur le potentiel de croissance osseuse de la mandibule, une perception esthétique du profil et son devenir ainsi que l'efficience d'une fonction linguale et respiratoire de sorte qu'il décidera de façon quasiment intuitive (sens clinique), en intégrant ces paramètres cliniques dans une même matrice d'analyse. Selon qu'il prône une forte stimulation de la croissance mandibulaire ou des extractions de prémolaires, la planification et l'outil thérapeutiques seront diamétralement différents.

En l'espèce, l'inefficacité de la machine n'est pas due au fait qu'il lui manque encore les données des différents paramètres faciaux qui sont nécessaires à son algorithme et qui, s'ils lui étaient un jour transmis, lui permettraient de « prendre la bonne décision ». Il s'agit plutôt de son incapacité structurelle à forger une « monnaie commune » entre paramètres étrangers les uns aux autres (par exemple, entre une fonction linguale et une esthétique faciale). Il n'est que de se pencher sur le fonctionnement du « Deep Learning » (Apprentissage profond machine) et tout particulièrement sur les mécanismes d'apprentissage supervisé machines pour comprendre pourquoi, en l'état actuel des connaissances sur l'IA, il est impossible de lui apprendre à mettre en place cette « monnaie commune ».

En creusant dans différentes spécialités médicales, on arrive à des conclusions similaires. Une incarnation semble nécessaire à ce type de cognition. Les algorithmes vont pouvoir décrypter et analyser des données : voilà en quoi ils sont supérieurs à notre cerveau. Mais lorsqu'il s'agira de les placer dans le contexte global de l'individu de sorte à créer une « monnaie commune » entre différents paramètres, ils seront dans une impasse : il ne leur est pas possible de créer des codes informatiques pour modéliser le sens clinique. C'est là que réside la plus-value de nos cerveaux – et pour longtemps encore.

2-3-3 Voies de connexion de l'expertise algorithmique et humaine

Différentes pistes peuvent être envisagées dans le cadre de la stratégie d'augmentation dont il a été question, le défi majeur pour le praticien étant d'accepter de sortir de sa zone de confort pour former un tandem avec l'intelligence artificielle et pour le programmateur de concevoir un

modèle au service de la cognition humaine et non son supplétif. S'agissant de l'orthopédie dento-faciale, je proposerais les hypothèses suivantes :

- Mieux connecter le praticien à l'environnement numérique

Rechercher des interfaces plus immersives entre la modélisation 3D issue des algorithmes et le praticien, en prenant en compte le fait qu'au-delà d'un degré de complexité de représentation 3D, le cerveau a du mal à maintenir son attention (Frey, Hachet et Lotte, 2017) (d'où l'intérêt pour la réalité virtuelle ou augmentée) ; garder aussi à l'esprit que le cerveau ne procède pas de manière verbale, nous aurons du mal à expliciter nos intuitions cliniques afin de les transmettre aux dispositifs CAD/CAM. Une autre problématique est que nous manipulons un dispositif avec 2 degrés de liberté (DOF), comme une souris, pour interagir avec un environnement 6 DOF.
En recherchant un environnement plus immersif nous réduirons le gouffre entre le monde réel et virtuel qui limite notre capacité à exploiter les nombreuses données que nous possédons pour prendre nos décisions.

- Faire du praticien une interface efficace

Chercher à former des praticiens capables d'offrir une interface efficace entre le patient et les dispositifs CAD/CAM plus que des techniciens performants qui se placeront en concurrence avec lesdits dispositifs. Cet objectif suppose de considérer la formation en internat (DES ODF) non pas comme une simple suite de procédés techniques, mais aussi comme un milieu censé nourrir la connaissance et la sensibilité du praticien à parts égales, de sorte que celui-ci soit à même d'évaluer de façon transversale les paramètres dont dépend la réussite de sa thérapeutique. Le rapport du député Cédric Villani (Villani, 2018) va dans cette même philosophie « S'il n'est pas question de remplacer les médecins par la machine, l'enjeu est bien d'organiser des interactions vertueuses entre l'expertise humaine et les rapports de l'IA dans les exercices quotidien de la médecine ». Les auteurs recommandent de former les professionnels de santé aux usages de l'intelligence artificielle, de l'internet des objets (IOT) et du big data en santé.

- Renforcer les réseaux entre praticiens au moyen du « deep learning »

Utiliser les capacités du « deep learning », notamment dans la reconnaissance des formes, pour identifier des cas cliniques susceptibles de connecter, grâce à cette reconnaissance, deux praticiens confrontés à des cas cliniques similaires en deux endroits distincts du globe. Offrir

aussi à chaque praticien la possibilité de mener des recherches thématiques ou des analyses critiques dans sa propre base de données.

L'objectif étant de combiner l'intelligence collective et les archives numériques. L'ensemble des données numériques issues de nos patients exploités par le levier de l'intelligence artificielle peut être au service du progrès de notre discipline générant un écosystème de connaissance. Il faudra sans doute faire preuve d'innovation pour stabiliser le cadre juridique et technique de cette démarche.

2-3-4 Conclusion

Valoriser une stratégie d'augmentation dans les rapports entre l'humain et l'intelligence artificielle, pour l'orthodontie de l'avenir, c'est se souvenir des paroles du prix Nobel de littérature Isaac Bashevis Singer : « plus la technologie progresse, plus les gens s'intéressent aux possibilités du seul esprit humain ».

La réussite de ce modèle passe par une symbiose cognitive entre le praticien et l'intelligence artificielle (figure 7) assurant l'utilisation optimale des nouvelles technologies, dans l'intérêt du patient.

Aujourd'hui, de nombreux obstacles rendent le chemin difficile et incertain, mais nous devons garder espoir et persévérer dans la quête d'une coopération transversale entre sciences informatiques et sciences médicales. Au-delà des enjeux propres à l'orthopédie dento-faciale, c'est un véritable enjeu de société qui doit être pris en compte : celui de valoriser l'humain dans un rapport pacifié avec les machines intelligentes. Face à la révolution technologique à l'œuvre, il importe de faire résonner les mots de l'historien Robert Rosenstone : « La révolution, c'est une tentative pour faire aboutir des rêves ».

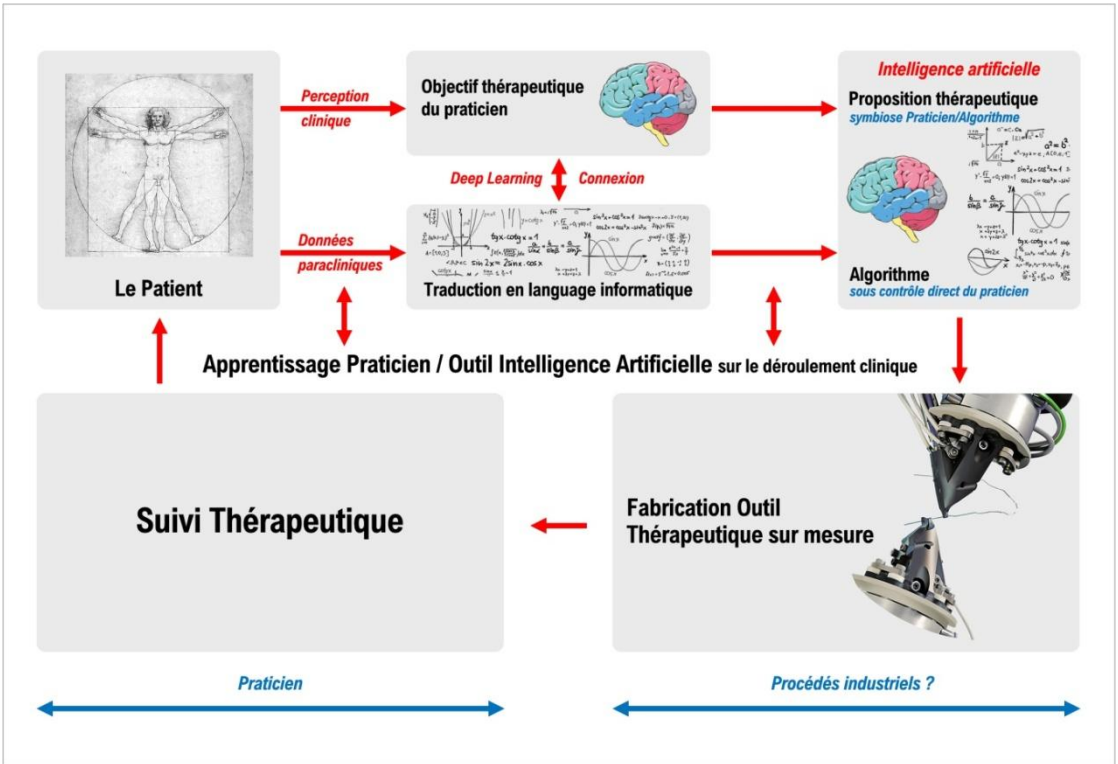


Figure 7 : Un workflow (flux de travail) optimisé permettant une symbiose entre le praticien et l'IA dans le procédé CAD/CAM.

2-3-5 Rôle de l'article dans ma démarche : une conviction quant à l'apport des sciences cognitives

Au terme de la réflexion menée dans le cade de mon Master 2, en partie illustrée par l'article ci-dessus, ma conviction quant à l'apport des sciences cognitives à l'optimisation de l'interface entre les nouvelles technologies et le praticien s'est trouvée renforcée. Deux points m'ont paru particulièrement importants à approfondir: mieux définir et comprendre l'impact de la virtualisation des données cliniques du patient sur la perception de ces données par le praticien (surtout s'agissant de données tridimensionnelles) et interroger le rôle fondamental de l'adéquation entre formation des praticiens et stratégie d'augmentation. Le travail de thèse devait donc être axé sur les problématiques de la virtualisation des données cliniques, en particulier des modèles en 3D, et son impact sur la perception qu'en a le praticien. En parallèle, j'ai poursuivi la rédaction d'un article de réflexion animé par la volonté de montrer que la transition numérique pouvait bel et bien s'inscrire dans une stratégie d'augmentation. L'illustration de cet état d'esprit est l'article présenté ci-dessous, mettant en relief l'intérêt que

revêt la transition numérique pour l'apprentissage et les progrès cliniques du praticien dans sa discipline.

2-4 Le workflow numérique au service de l'apprentissage clinique du praticien

Référence bibliographique de l'article :

Makaremi, M., de Brondeau, F., Lacaule, C., & N'Kaoua, B. (2020). Le workflow numérique au service de l'apprentissage clinique du praticien. *Orthod Fr*, *91*, 93-99.

2-4-1 Introduction

Impact du numérique sur le praticien :

Au cours des dernières décennies, nos habitudes ont été bouleversées par la naissance de l'orthodontie numérique et digitale. L'avènement de l'imagerie, des empreintes tridimensionnelles, l'expression de l'intelligence artificielle à travers les algorithmes de décision thérapeutique (p. ex. : set-up virtuel dynamique) et le développement des procédés de fabrication individualisée des appareils orthodontiques (CFAO : « Conception et fabrication assistées par ordinateur »), génèrent de profondes mutations. Celles-ci interviennent dans un contexte de transformation de la pratique médicale, au-delà duquel c'est toute la société que nous voyons se métamorphoser au contact des nouvelles technologies numériques (Makaremi et al., 2020).

L'impact de cette révolution numérique et digitale sur notre prise en charge des thérapeutiques orthodontiques est l'objet d'un riche débat (Makaremi, 2018).
Un point clé dans cette mutation est la place et l'identité du praticien : aura-t-il la maitrise du workflow numérique ou sera-t-il plutôt un fournisseur ?

Si l'objectif est une stratégie d'augmentation valorisant la place du praticien, le défi décisif pour sa réussite est de parvenir à donner la maîtrise de l'outil numérique aux praticiens tout en les faisant progresser cliniquement au contact de l'outil numérique.

L'objectif de cet article est de démontrer que si le workflow numérique a été suffisamment bien organisé et structuré en s'appuyant sur les théories issues des neurosciences cognitives, il peut devenir un formidable outil optimisant l'apprentissage et l'expérience clinique du praticien, lui assurant une progression tout au long de sa carrière, au fur et à mesure des cas auxquels il sera confronté.

2-4-2 Décryptage du workflow

L'interaction entre l'outil numérique et le processus thérapeutique peut être présentée en 4 étapes successives : le diagnostic, la planification thérapeutique, la fabrication de l'outil thérapeutique sur mesure et le suivi thérapeutique (figure 8).

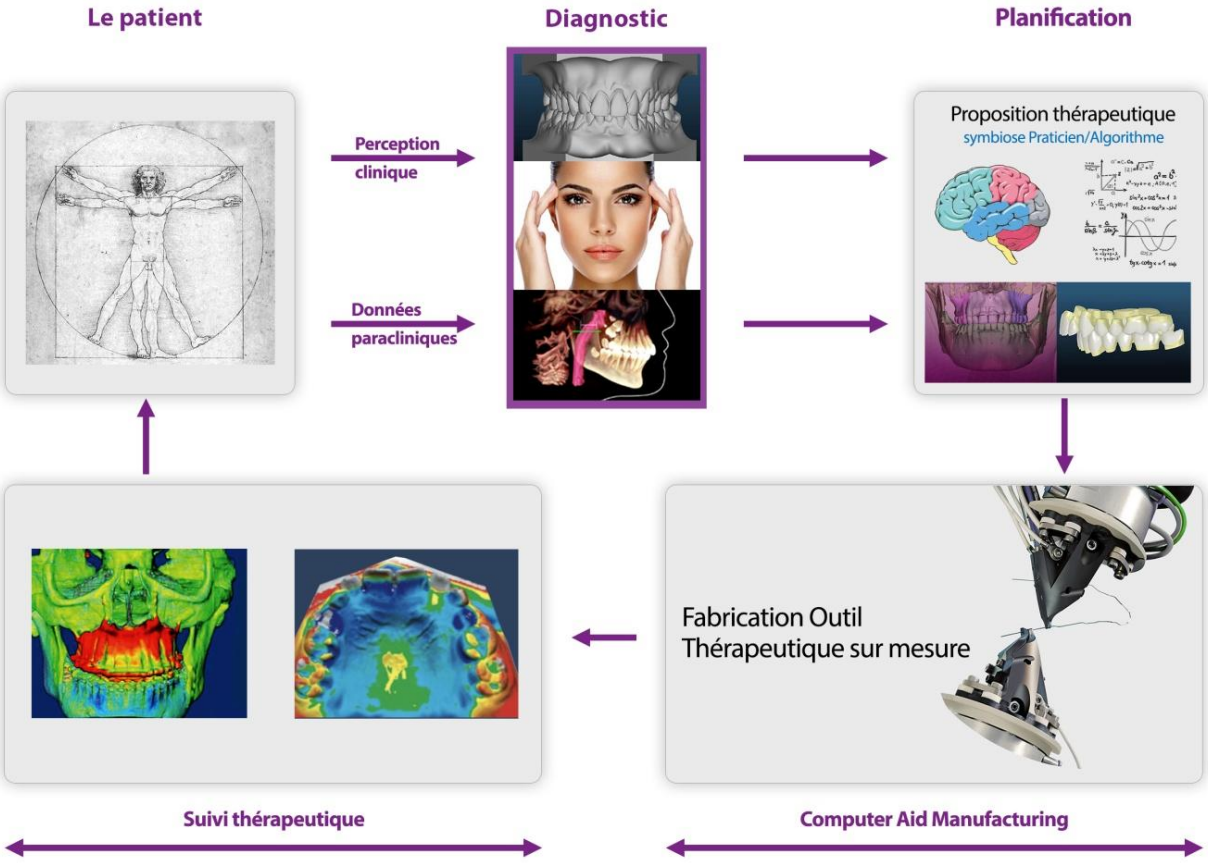


Figure 8 : Le workflow numérique.

2-4-2.1 Le diagnostic

Les progrès de l'imagerie maxillo-faciale grâce au « Cone-Beam grand champ », au scanner intra-oral et à la photographie 3D permettent de voir à travers les tissus, de mesurer des distances et d'avoir une vue d'ensemble du patient. De plus en plus précise et interactive, cette imagerie rend possible une vision plus transparente du patient. Nous voyons également se développer les premières applications basées sur l'intelligence artificielle, qui permettent par exemple d'effectuer des tracés céphalométriques directement sur la téléradiographie de profil sans intervention du praticien (Nishimoto et al., 2019 ; Zamora et al., 2012), allégeant la charge cognitive induite par ce travail fastidieux.

2-4-2-2 Planification thérapeutique

L'augmentation de la puissance de calcul des processeurs permet aux outils numériques de stimuler les objectifs thérapeutiques. Le set-up virtuel dynamique en est l'illustration parfaite : il permet d'obtenir une modélisation des objectifs thérapeutiques en 3D (Hennessy et Al-Awadhi, 2016). Il existe deux procédés de construction du set-up : soit on envoie les empreintes avec des instructions à un laboratoire chargé de présenter un set-up virtuel dynamique sur lequel le praticien peut également agir (clean check, insignia), soit on procède à la construction du set-up directement à partir de logiciels (3 shape, sky plan).

Le set-up virtuel dynamique est un précieux outil de visualisation des mouvements orthodontiques, de leur faisabilité, ainsi que de la biomécanique à mettre en place pour atteindre les objectifs thérapeutiques. Il constitue également un outil efficace d'explication et de dialogue avec le patient afin d'obtenir son adhésion et son implication dans le protocole orthodontique.

2-4-2.3 Fabrication sur mesure de l'outil thérapeutique

Dans la continuité de la planification, la seconde phase de la technologie CAD/CAM (Computer Aid Design, Computer Aid Manufacturing) permet de fabriquer des dispositifs médicaux sur mesure. L'impression 3D est une des avancées de la révolution numérique vouées à avoir un impact considérable sur notre métier d'orthodontiste (Lagravere et Flores-Mir, 2005) : elle supplée le travail manuel de conception, en donnant une réalité physique au projet

thérapeutique virtuel et en diminuant le biais humain. Elle est la traduction du passage d'un vecteur manuel vers un vecteur cérébral désormais au cœur de la mutation de notre profession.
Ici aussi, deux procédés s'offrent aux praticiens : après la validation de la proposition thérapeutique, le dispositif orthodontique sur mesure est fabriqué soit à distance par un laboratoire soit par le praticien lui-même si tant est qu'il possède les imprimantes 3D nécessaires à l'impression des dispositifs en question à partir des fichiers STL. Ce secteur est en plein développement, un des principaux défis étant l'impression directe des aligneurs.

2-4-2.4 Suivi thérapeutique

La possibilité de visualiser et de quantifier l'impact de la thérapeutique sur les différentes structures anatomiques présente de nombreux avantages dans le cadre d'une thérapeutique orthodontique.
D'abord, elle permet au praticien de valider les objectifs ou de rechercher ce qui n'a pas fonctionné et au patient de visualiser les effets du traitement. Son intérêt médico-légal semble également évident. Enfin, elle permet d'optimiser le travail d'équipe dans des cas pluridisciplinaires en permettant d'objectiver l'impact des différentes phases et, à travers l'analyse des résultats obtenus, de mieux appréhender les défis cliniques futurs.
L'option idéale pour le suivi thérapeutique est la possibilité d'effectuer des superpositions à l'aide d'outils numériques basés sur l'intelligence artificielle. Ces superpositions numériques sont possibles à différentes échelles. Des logiciels construits à partir de différents modèles d'apprentissage automatique (« machine learning ») disponibles dans certains scanners intra-oraux permettent de superposer deux modèles numériques d'un même patient (Vinas, De Hierro et Ustrell-Torrent, 2018 ; Ganzer et al., 2018), ce qui rend possible la visualisation des mouvements des arcades dentaires. Il existe également d'autres solutions basées sur l'intelligence artificielle, combinant la photographie numérique et les empreintes numériques : c'est le 3D Matching (Roisin, Brézulier et Sorel, 2016). Cette technologie permet de superposer sur les photographies des modèles 3D.
Naturellement, les superpositions en imagerie radiologique 3D sont aussi un moyen très précieux pour objectiver l'impact de nos thérapeutiques (Cevidane et al., 2005).

2-4-3 Les théories de neurosciences cognitives dans l'apprentissage

Les neurosciences cognitives ont effectué d'importants progrès dans la compréhension des processus d'apprentissage. Les travaux réalisés dans ce domaine permettent de voir sous un jour nouveau l'utilisation quotidienne du workflow numérique. Deux théories méritent d'être expliquées plus en détail : la théorie de l'inférence bayésienne et la théorie des 4 piliers de l'apprentissage selon le professeur Stanislas Dehaene.

2-4-3.1 L'inférence bayésienne

Comme le soulignent Tenebaum et al. (Tenebaum et al., 2011), pour les scientifiques qui étudient les mécanismes de l'apprentissage humain, le principal défi est de comprendre comment les humains apprennent à partir de si peu. C'est ce que le philosophe Bertrand Russell nomme « le scandale de l'induction ». En somme, il s'agit de notre capacité à construire un modèle théorique à partir d'un exemple vécu.

L'application de la théorie bayésienne aux sciences cognitives permet d'aborder en partie cette problématique. L'inférence bayésienne est une théorie mathématique simple qui caractérise le raisonnement plausible en présence d'incertitude (Jaynes, 2003). Elle rend bien compte des processus de perception étant donné des entrés ambiguës, notre cerveau en reconstruit l'interprétation la plus probable.

En résumant, on peut dire que notre cerveau construit son modèle d'apprentissage à partir d'exemples qui lui permettent d'ajuster a priori la vraisemblance et la probabilité face à de nouvelles situations avant de prendre ses décisions (Gweon, Tenenbaum et Schulz, 2010). L'application de ce cadre théorique d'inférence bayésienne à l'orthodontie permet de mettre en exergue le rôle majeur de l'étude des cas cliniques dans l'apprentissage au cours de la formation initiale et tout au long de notre carrière. En effet, c'est en grande partie à partir de ces exemples que l'on forme notre œil de praticien et que l'on modèle notre matrice de décision.

2-4-3.2 Les 4 piliers de l'apprentissage

Le professeur Stanislas Dehaene (Dehaene, 2018 ; Dehaene, 2014 ; Dehaene et al., 2011 ; Dehaene et al., 2010) a énoncé les 4 piliers essentiels qui permettent à notre cerveau, à la fois structuré et plastique, d'optimiser son apprentissage : l'attention, l'engagement actif, le retour d'information et la consolidation. Voici comment ces piliers ont été définis (figure 9) :

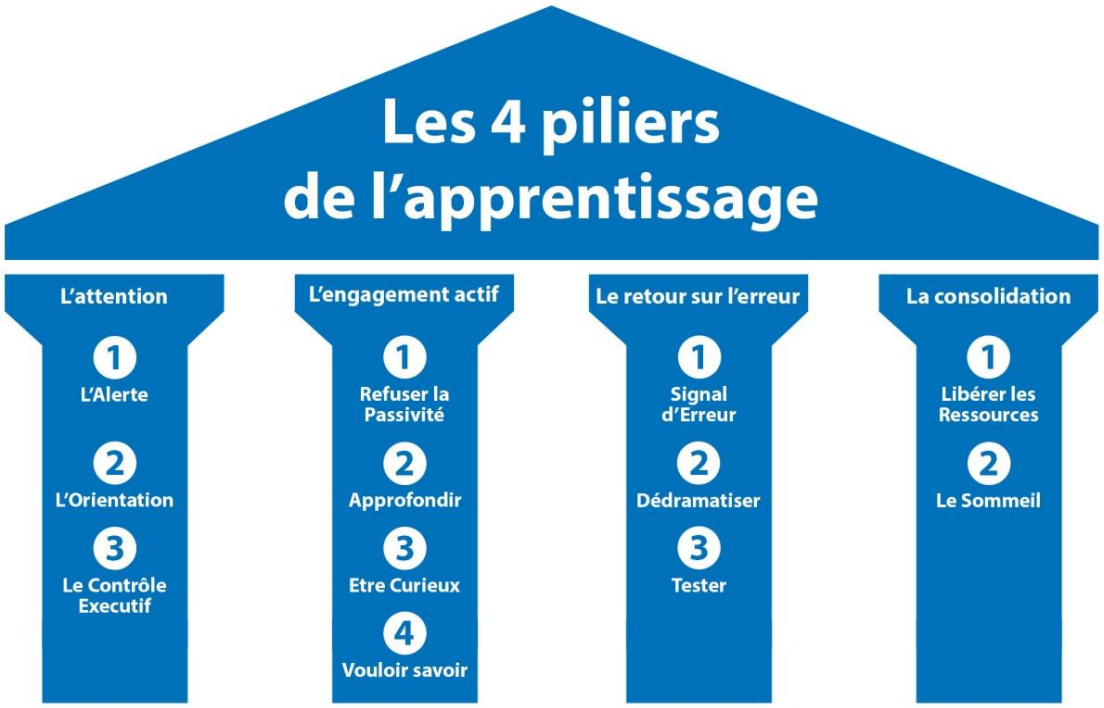


Figure 9 : Les quatre piliers de l'apprentissage d'après Dehaene (2018).

2-4-3.2.1 L'attention

Le premier des quatre piliers est la capacité de notre cerveau à sélectionner une information, l'amplifier, la canaliser et l'approfondir. En effet, abreuvé d'information et stimulé par les sens, le cerveau a besoin d'un filtre pour conserver ce qu'il estime essentiel. Ce qui signifie que « faire attention », c'est faire taire certaines informations jugées non pertinentes pour opérer une sélection parmi elles (Desimone et Duncan, 1995 ; Duncan, 2013).

2-4-3.2.2 L'engagement actif

Pour apprendre, il faut un engagement actif, ce qui incite à échafauder sans cesse de nouvelles hypothèses. Si vous êtes attentif, concentré, véritable acteur de votre apprentissage, et que vous traitez l'information en profondeur, vous faciliterez grandement votre travail de mémorisation (Freman, 2014). L'engagement actif est la curiosité (Cevidanes et al., 2005) qui pousse chacun d'entre nous à explorer son environnement, et la curiosité est un ingrédient essentiel de l'algorithme d'apprentissage de notre cerveau.

2-4-3.2.3 Le retour d'information

Le retour d'information est le troisième pilier de l'apprentissage : nous progressons en nous trompant à condition toutefois de recevoir un signal en retour, un feedback : la surprise qui émane du décalage entre la prédiction et la réalité. Le modèle créé par (Rescola et Wagner, 1972) nous apprend que le cerveau opère une prédiction puis calcule la différence entre sa prédiction et le stimulus effectivement rencontré, appelée « erreur de prédiction ». Ensuite, il effectue la correction nécessaire pour que sa prochaine prédiction soit plus proche de la réalité. Ainsi, l'option la plus favorable à l'apprentissage est d'avoir un retour sur erreur rapide et précis, sans qu'il s'agisse d'une punition.

2-4-3.2.4 La consolidation

Le quatrième pilier de l'apprentissage consiste à passer de l'effort, passage obligé de tout apprentissage, à la consolidation, ce qui permettra à notre cerveau de conserver nos nouvelles connaissances plus longtemps. Par exemple, dans le cas d'un enfant de 6 ans, le cerveau fournit un effort considérable pour lire, et déchiffre lentement. Pour que cette activité devienne une routine, un automatisme, il faudra encore quelques années de pratique intensive. Très énergivore au début, cette activité va de moins en moins solliciter le cortex pariétal et le cortex préfrontal. Après quelques années de pratique, l'enfant lira même sans effort. Cette routine est primordiale, car elle permet à l'individu de se concentrer sur autre chose.

2-4-4 Les 4 piliers de l'apprentissage au sein du flux de travail numérique

En combinant les séquences du workflow numérique et les théories sur l'apprentissage issues des travaux en neurosciences cognitives, on constate que chacune des 4 étapes du workflow peut s'assimiler à un des piliers de l'apprentissage. Cet état de fait semble donner raison au Prix Nobel de littérature Isaac Bashevis Singer : « Plus la technologie progresse, plus les gens s'intéressent aux possibilités du seul esprit humain. »

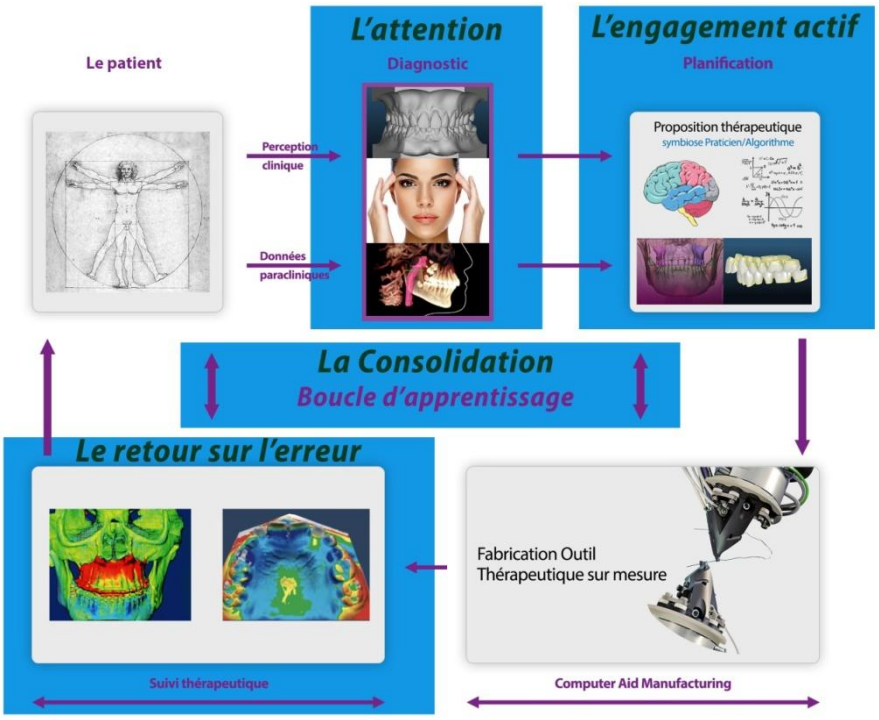


Figure 10 : Le flux du travail numérique renferme les quatre piliers de l'apprentissage.

2-4-4.1 Diagnostic/Attention

Les outils de diagnostic développés à partir du numérique tendent à amener une plus grande transparence et une vision tridimensionnelle de la face du patient. La quantité et la qualité de l'information disponible concernant le patient sont nettement amplifiées, ce qui permet à notre cerveau de mieux sélectionner l'information jugée importante et de l'approfondir, étant à même d'effectuer une observation plus précise. En outre, grâce à l'utilisation de l'IA, des mesures comme le tracé céphalométrique peuvent être automatisées, ce qui allège notre cerveau d'une charge cognitive, en le rendant plus disponible au traitement qualitatif de

l'imagerie clinique du patient. Mais le facteur le plus important pour réussir à améliorer l'attention du praticien grâce à l'outil numérique, c'est notre capacité à créer un filtre pour conserver ce qu'on estime essentiel. En effet, le piège est de se noyer dans une masse de données rendue possible grâce à la puissance de l'outil numérique. Au contraire, si le praticien désire augmenter son attention, il doit être capable de faire un tri pour conserver dans une matrice de prise de décision uniquement les données vraiment pertinentes pour chaque patient : « faire attention », c'est faire taire certaines informations et en sélectionner quelques-unes.

En somme, il s'agit de toujours maintenir une certaine cohérence entre ses observations cliniques et le choix des examens paracliniques en appliquant le principe dit du rasoir d'Ockham (qui peut se résumer dans la formule « les multiples ne doivent pas être utilisés sans nécessité »). Ainsi, les différents outils d'imagerie numérique peuvent potentialiser l'attention du praticien si tant est qu'on soit capable de mettre en place une véritable stratégie de gestion de cette imagerie, tout en évitant une systématisation intempestive.

2-4-4.2 Planification/engagement actif

La mise en place d'un schéma thérapeutique basé sur le CAD/CAM (Computer Aid Design, Computer Aid Manufacturing) nécessite la modélisation des objectifs thérapeutiques en 3D, dont l'expression la plus manifeste est le set-up virtuel dynamique. Bien sûr, tout dépend du degré d'implication du praticien dans l'édification du set-up. Cependant, il semble évident qu'un praticien investi dans la construction et la validation du set-up virtuel dynamique va être amené à être attentif aux objectifs thérapeutiques et traiter les informations en profondeur pour valider la faisabilité du projet thérapeutique. En outre, s'il possède une bonne maîtrise de l'outil numérique, couplée à une bonne vision des données du patient ainsi qu'à une connaissance de la biomécanique, il va pouvoir échafauder sans cesse de nouvelles hypothèses, verbaliser à travers la construction du set-up son intuition clinique. Ce travail de planification virtuelle constitue donc un véritable moteur de l'engagement actif tout à fait compatible avec une vision quasi ludique du travail du praticien, attisant sans cesse sa curiosité, un ingrédient essentiel de l'algorithme d'apprentissage de notre cerveau. Ainsi, la planification thérapeutique peut être l'épicentre de l'engagement actif, pilier d'un meilleur apprentissage chez le praticien, à condition que celui-ci soit totalement impliqué dans la construction du set-up virtuel dynamique. Plus il s'amusera et aura l'occasion de cultiver sa curiosité plus son apprentissage

sera performant ; et plus il jouera un rôle de « provider » (fournisseur), conformément aux diktats de certaines multinationales, plus l'intérêt clinique se verra amoindri.

2-4-4.3 Suivi thérapeutique/retour sur erreur

Nous progressons si nous nous trompons, à condition toutefois de recevoir un signal en retour. Les outils numériques permettent un bien meilleur suivi du patient, de plus en plus précis et rapide, ceci d'autant plus qu'ils peuvent être reproduits fréquemment tout au long du traitement (à l'exception de l'imagerie radiologique ionisante, qui doit être justifiée chaque fois). Couplées à la possibilité d'effectuer des mesures automatiques grâce aux techniques de « Deep learning », ces technologies permettent de voir et de mesurer instantanément l'impact de nos thérapeutiques, et de visualiser ce qui n'a pas fonctionné. Une véritable stratégie de monitoring du patient permet d'avoir un retour sur erreur rapide et précis, à condition de rester attaché à une forme d'honnêteté intellectuelle et de ne pas se blâmer soi-même pour ses erreurs, comme on a souvent tendance à le faire, d'une manière on ne peut plus contre-productive sur le plan des techniques d'apprentissage.

2-4-4.4 La consolidation

On le constate tous les jours, l'intensification d'environnement numérique génère une accélération de rythme. Une mise en place structurée du workflow numérique et des techniques (CAD/CAM) utilisées de façon quotidienne rend possible une consolidation. L'enjeu majeur est pour le praticien de rester connecté au workflow numérique au lieu de simplement le subir ; il s'agit, par exemple, de palier la perte du sens haptique ou de créer une interface homme-machine ergonomique privilégiant le confort du praticien au respect du process industriel.

En somme, si le workflow est construit autour du travail cognitif du praticien, chaque cas clinique traité viendra consolider son expérience clinique. En revanche, si le workflow est centré sur l'appareil orthodontique (aligneur, appareil lingual), le praticien se voyant relégué à un rôle plus proche de celui du « provider » (fournisseur) ou du contrôleur de qualité, l'impact positif sur le renforcement de l'apprentissage du praticien sera significativement moindre.

2-4-5 Conclusion

Notre rapport aux nouvelles technologies sera un élément déterminant dans la construction de notre identité d'orthodontistes du 21e siècle.

Si l'enjeu de cette nouvelle interface est de tendre vers une stratégie d'augmentation, le clinicien doit se fixer pour objectif de progresser et d'augmenter son expertise tout au long de sa carrière. Les neurosciences cognitives nous montrent que notre cerveau fonctionne sur un mode bayésien, en construisant ses modèles théoriques et sa matrice décisionnelle à partir d'exemples concrets. Ce qui s'illustre dans notre pratique de l'orthodontie par l'importance fondamentale que revêt l'observation de cas cliniques. Lorsqu'on met en parallèle le schéma du workflow numérique en orthodontie et les 4 piliers de l'apprentissage théorisés par le professeur Dehaene, on s'aperçoit que l'utilisation du flux numérique peut être un formidable outil d'apprentissage au service du praticien, désormais à même d'obtenir un précieux retour sur expérience lors de chaque cas clinique traité. Cet objectif suppose cependant une véritable stratégie numérique, centrée sur la cognition du praticien. Cette stratégie n'est possible que si le praticien garde sa liberté d'action et de réflexion, un défi très difficile à relever, mais qui en vaut la peine ; songeons au mot de Malraux : « La liberté appartient à ceux qui l'ont conquise ».

2-4-6 Rôle de l'article dans ma démarche : nécessité d'aller plus loin avec les sciences cognitives

L'écriture de l'article ci-dessus a été une étape importante dans mon parcours de thèse, car elle a permis de connecter ma réflexion basée sur les problématiques de l'orthopédie dento-faciale à une réflexion plus générale, à un niveau plus élevé d'abstraction, comme il se doit dans le cadre d'une thèse scientifique. J'ai pris conscience que la question de fond que j'entreprenais de traiter était celle de l'impact de la transition numérique sur la perception de l'expert au cours de sa prise de décision, et que mes travaux sur la perception par l'orthodontiste des modèles numériques était un exemple qui permettait d'illustrer cette problématique. Ainsi, le point suivant présente une synthèse de la bibliographique sur la perception de l'expert, permettant d'ancrer mon sujet dans l'impact de la transition numérique et digitale sur la perception et la prise de décision de l'expert.

2-5 La perception de l'expert

2-5-1 Introduction

Considérés sous l'angle de la finesse et de l'habileté qu'ils exigent de ceux qui les pratiquent, les différents métiers présentent chacun leur spécificité. Cependant, en se référant à la littérature du talent et de l'expertise (Gobet, 2011), on se rend compte à quel point les schémas de l'acquisition de l'expertise et de sa mise en œuvre peuvent avoir des invariants dans les différents corps de métiers. Pour l'orthodontiste, la capacité à effectuer un diagnostic sur des modèles d'études orthodontiques est une expertise perceptive visuelle similaire, quant à son objet, à bien d'autres domaines ou l'expertise perceptive visuelle est sollicitée.

L'expertise humaine est un sujet étudié en profondeur, entre autres, par les sciences cognitives, la psychologie et les neurosciences. Au cours de ce chapitre, après avoir posé les bases de la définition de l'expert et de l'expertise, nous aborderons les mécanismes cognitifs et neurobiologiques de l'expertise en présentant plus particulièrement la théorie des chunks. Ensuite, seront détaillées les notions d'expertise perceptive visuelle, et de plasticité cérébrale dans sa mise en place et sa fiabilité dans un environnement aussi complexe que celui constitué par l'anatomie tridimensionnelle des arcades dentaires.

2-5-2 « Expert » et « expertise »

2-5-2-1 Définition du terme « expert »

Le dictionnaire Webster's New World Dictionary (Schell, Schell et Martin, 2006) définit l'expert comme une personne qui possède compétence et habileté dans un domaine particulier, ou une personne largement reconnue comme une source fiable de connaissances, de techniques ou de compétences, ou dont le jugement se voit accorder une autorité et un statut élevés par le public ou par ses pairs.

Gobet (2011) définit l'expert comme une personne qui obtient des résultats supérieurs à ceux de la majorité de la population. Selon l'auteur, cette définition a le mérite de pouvoir être appliquée récursivement : « un super expert » est un expert qui obtient une performance supérieure à la majorité des experts. La notion d'expertise est alors corrélée à celle de performance.

Comme le souligne (Ericsson et al., 2006), la compétence de l'expert n'est pas quelque chose qui va et vient. Si vous réveilliez des joueurs d'échecs qualifiés au milieu de la nuit et leur montriez une situation de jeu difficile, ils trouveraient la solution sans trop de difficulté. Tout comme les radiologues expérimentés sont à même de détecter les lésions à partir des images radiologiques quel que soit le contexte.

Dans certains domaines, il existe des critères objectifs pour déterminer les experts qui sont constamment en mesure d'afficher des performances supérieures dans des tâches représentatives (Ericsson et al., 2006). Par exemple, les maîtres d'échecs gagneront presque toujours les parties contre des joueurs amateurs dans un tournoi, les médecins spécialistes sont beaucoup plus susceptibles de diagnostiquer correctement une maladie que les étudiants en médecine, et les musiciens professionnels peuvent interpréter des morceaux de musique d'une manière difficilement atteignable par des musiciens moins qualifiés. Selon Ericsson (Ericsson et al., 2006), il existe néanmoins des domaines où il est difficile pour les non-experts d'identifier des experts. Dans ces domaines, les chercheurs s'appuient sur des nominations par des pairs ou par des professionnels du domaine concerné. Cependant, les personnes reconnues comme experts par leurs pairs n'affichent pas systématiquement des performances supérieures dans les tâches liées au domaine considéré. Parfois, ils ne sont pas meilleurs que les novices, même pour ce qui est des tâches se situant au cœur de leur expertise, comme dans le domaine boursier (Bedard, 1991).

Dans son ouvrage « The Neuroscience of Expertise » (Bilalić, 2017), Merim Bilalić souligne que les domaines d'expertise classiques sont des environnements stables et que les changements qui se produisent (par exemple, de nouvelles maladies, de nouvelles marques de balles de tennis ou de nouveaux cordages de raquette) sont généralement suffisamment mineurs pour ne pas modifier radicalement l'environnement concerné en rendant non pertinentes les connaissances antérieures. Chaque domaine d'expertise fournit une mine d'informations cohérentes à ses praticiens. Selon le même auteur, les experts parviennent à trouver des moyens ingénieux pour contourner les limitations cognitives qui s’exercent sur eux. Néanmoins, comme peut en témoigner quiconque s'est essayé aux jeux de sport (ou à toute autre activité), les domaines d'expertise sont extrêmement complexes, et leur maîtrise exige de nombreuses années de pratique dévouée. Il y a beaucoup de choses à apprendre dans n'importe quel

domaine d'expertise, et c'est précisément cette connaissance des caractéristiques particulières qui se répète sans cesse dans un domaine spécifique qui permet à l'expert d'appréhender les problèmes avec des yeux différents de ceux d'un novice.

2-5-2-2 Domaines d'expertise

Bilalić (2017) identifie trois domaines d'expertise distincts :

L'expertise perceptuelle (perceptual expertise) : elle est relative à des domaines qui reposent principalement sur des informations provenant de nos sens. Par exemple, la recherche purement visuelle comme celle que doit effectuer le radiologue recherchant une lésion, même si, évidemment, les radiologues expérimentés doivent également solliciter leur mémoire, sans laquelle ils seraient incapables de repérer et de catégoriser les lésions.

L'expertise cognitive (cognitive expertise) : le jeu d'échecs est un exemple phare de ce type d'expertise. Dans cette discipline, l'information perçue par nos sens est déterminée par nos connaissances préalables. Les situations de jeu d'échecs précédemment stockées en mémoire influencent directement les mécanismes perceptifs et permettent d'aider à comprendre le problème à résoudre, puis à imaginer le déroulement du jeu (processus dits « top-down » ou descendants).

L'expertise sensorimotrice (motor expertise) : la performance de sportifs de haut niveau dans de nombreux sports, comme le tennis ou le football, est l'illustration de ce type d'expertise. L'acquisition d'habiletés sensorimotrices se traduit par une amélioration des performances reflétant une optimisation du contrôle sensorimoteur, en particulier en termes de processus de préparation et d'exécution ainsi que des capacités de réaction et d'anticipation.
Le concept d'habileté renvoie directement à la notion de différences inter-individuelles dans le contrôle du mouvement et, de ce fait, au principe d'expertise sensorimotrice. Cette expertise est ainsi associée à la capacité d'effectuer un mouvement rapide et précis ou, plus généralement, mieux adapté aux exigences de la tâche concernée.

2-5-2-3 Expertise et mécanismes neurobiologiques & cognitifs

L'expertise est un excellent exemple de la façon dont divers processus cognitifs, tels que la mémoire, l'attention ou la perception, se combinent pour permettre d'incroyables performances. Les mécanismes cognitifs sous-jacents à l'acquisition de cette expertise ont été au centre de nombreuses études.

Au fil de longues années d'expériences, les experts ont acquis des connaissances cohérentes dans leur domaine (Chase et Simon, 1973 ; Gobet et al., 2001).

Tous les domaines comportent des « règles » qui sont stables et des situations qui se reproduisent sans cesse sous une forme ou sous une autre. Les connaissances relatives à chacun d'eux sont stockées dans une « mémoire à long terme » (LTM), le processus de rétention auquel nous nous référons habituellement lorsque nous parlons de mémoire dans la vie de tous les jours. Le nom vient de l'idée que les informations stockées ici resteront disponibles à des fins de récupération pendant des semaines, des mois voire des décennies, au contraire de la « mémoire à court terme » (STM), où le contenu est retenu pendant une durée très limitée.

Lorsqu'un expert rencontre une situation apparemment nouvelle dans son domaine, il active automatiquement les connaissances spécifiques à son domaine d'expertise, stockées dans sa mémoire à long terme (Richman, Staszewski et Simon, 1995). La nouvelle situation est ensuite comparée à une situation précédemment rencontrée et stockée en mémoire à long terme. La conséquence de cette correspondance automatique des modèles du monde extérieur et des modèles stockés en mémoire est que les experts saisissent rapidement l'essence de la nouvelle situation. Les experts n'ont donc pas besoin de capacités extraordinaires pour comprendre la situation complexe à laquelle ils sont confrontés. Leurs connaissances leur permettent de chercher la solution en se référant à une situation similaire, sauvegardée dans la mémoire à long terme.

Un autre aspect des mécanismes cognitifs propres aux joueurs d'échecs a été exposé par les travaux de De Groot (1978), qui ont démontré que les stratégies des experts sont très différentes de celles de leurs pairs moins qualifiés (amateurs). Alors que les novices ont besoin d'examiner chaque recoin de la position pour identifier les pièces d'intérêt, les experts n'ont besoin que de quelques regards pour effectuer la même tâche. L'auteur a également démontré que les joueurs d'échecs experts ne regardent pas nécessairement plus loin mais examinent immédiatement des solutions de haute qualité : leurs connaissances leur permettent de

s'orienter instantanément et de se concentrer sur des solutions prometteuses quant au problème posé.

Ainsi, comme le soulignent Bilalić et al. (2010), les stratégies des experts peuvent être rapides et même donner l'impression de ne pas exiger d'effort, mais leur efficacité est basée sur une interaction complexe entre de nombreux processus cognitifs, ce qui a un effet profond sur la façon dont le cerveau met en œuvre les performances des experts, grâce à sa remarquable capacité d'adaptation à la demande environnementale.
Grâce aux progrès de la neuro-imagerie, il est possible de mieux comprendre les mécanismes cérébraux en jeu dans l'expertise. La technique de neuro-imagerie la plus couramment utilisée est sans doute l'imagerie par résonance magnétique fonctionnelle (IRMf). Les changements structurels du cerveau sont généralement étudiés à l'aide de la morphométrie à base de voxel (VBM) pour ce qui est de la matière grise et à l'aide de l'imagerie du tenseur de diffusion (DTI) pour ce qui est de la substance blanche (Ericsson et al., 2006).

2-5-2-4 Le concept de chunking

Une des théories expliquant les performances d'un expert est la théorie des chunks, qui détaille le mode de résolution de problèmes et les processus de mémoire utilisés par les experts. Dans ce modèle, l'importance de la perception est soulignée. Le modèle en question met en avant la vitesse avec laquelle les experts « perçoivent » les éléments clés d'une situation problématique et peuvent, instantanément, reconnaitre les configurations déjà rencontrées et stockées dans leur mémoire à long terme sous forme de « chunks ». Ce concept repose sur les travaux de Miller (1956) qui avait montré l'efficacité d'un regroupement des informations à l'aide du code binaire. Il postule que ce mécanisme d'association représente le fonctionnement de la mémoire dans son ensemble et, donc, que les nouvelles connaissances se créent en regroupant des unités de bas niveau en de nouvelles unités signifiantes. Comme nous l'avons mentionné précédemment, la mémoire à court terme, dont la capacité est limitée, aurait un empan mnésique d'environ sept éléments. Ainsi, en procédant par chunking, il est possible de retenir plus d'informations, puisque les unités ne sont plus mémorisées individuellement mais par regroupement au sein d'unités signifiantes plus vastes. La taille du chunk, qu'il soit composé de trois unités ou bien de six, ne représentera qu'une seule information à mémoriser. Cela a été démontré notamment par Ehrlich (1972), au moyen d'un matériau langagier.

Dans le principe, le chunking est donc proche de la théorie des géons formulée par Biederman (1987). Selon lui, la construction des objets qui nous entourent repose sur des formes de base qui s'associent entre elles pour former de nouveaux objets ayant une signification propre. Les individus seraient alors composés d'un répertoire, estimé à environ 30000 objets qui se construisent à partir de quelques éléments de base. Nous pouvons également citer le modèle du pandémonium de Selfridge et al. (1959), qui repose sur la reconnaissance de traits particuliers en provenance du système sensoriel pour expliquer le processus de reconnaissance des lettres.

Johnson (1970) décrit, quant à lui, quatre caractéristiques des chunks :

- Le chunk lui-même, c'est-à-dire la nouvelle unité signifiante construite à partir d'unités de plus bas niveau.
- Le "code mnémotechnique" qui correspond au symbole enregistré par le sujet lors de la phase de mémorisation et qui est associé à un chunk spécifique.
- Le processus de recodage qui consiste à mémoriser le lien entre le "code mnémotechnique" et le chunk dont on souhaite se rappeler.
- Le processus de décodage qui permet de retrouver les informations que l'on souhaite mémoriser à partir de la réactivation du "code mnémotechnique". La performance des experts semble, donc, être en partie imputable à la reconnaissance de certains patterns spécifiques que l'on nomme des chunks. Par exemple, dans le domaine du jeu d'échecs, ces chunks sont des configurations particulières de pièces qui vont être reconnues immédiatement et vont permettre l'activation de certains plans d'action spécifiques à la situation. Sheridan et Reingold (2017) expliquent que cette capacité à percevoir immédiatement certaines structures de pièces spécifiques est une composante clé de l'expertise.

En psychologie de l'expertise, le modèle CHREST (Chunk Hierarchy and Retrieval Structures ; Gobet, 1993) propose que les chunks soient organisés à la manière de réseaux hiérarchiques. Il s'agit d'une architecture cognitive construite afin de reproduire et d'expliquer le fonctionnement de la mémoire des joueurs d'échecs. L'architecture cognitive proposée par le programme CHREST (figure 11) reprend plusieurs éléments de ces modèles afin de reproduire la performance experte de manière globale. Ce programme est composé d'un espace visuel duquel vont être extraites les informations importantes d'une mémoire à court terme ayant une capacité maximum de sept éléments ; d'une mémoire à long terme qui permet de récupérer

l'information à travers un réseau de discrimination ; et, enfin, de structures de récupération qui servent à encoder rapidement l'information et à la récupérer efficacement lorsque l'on en a besoin.

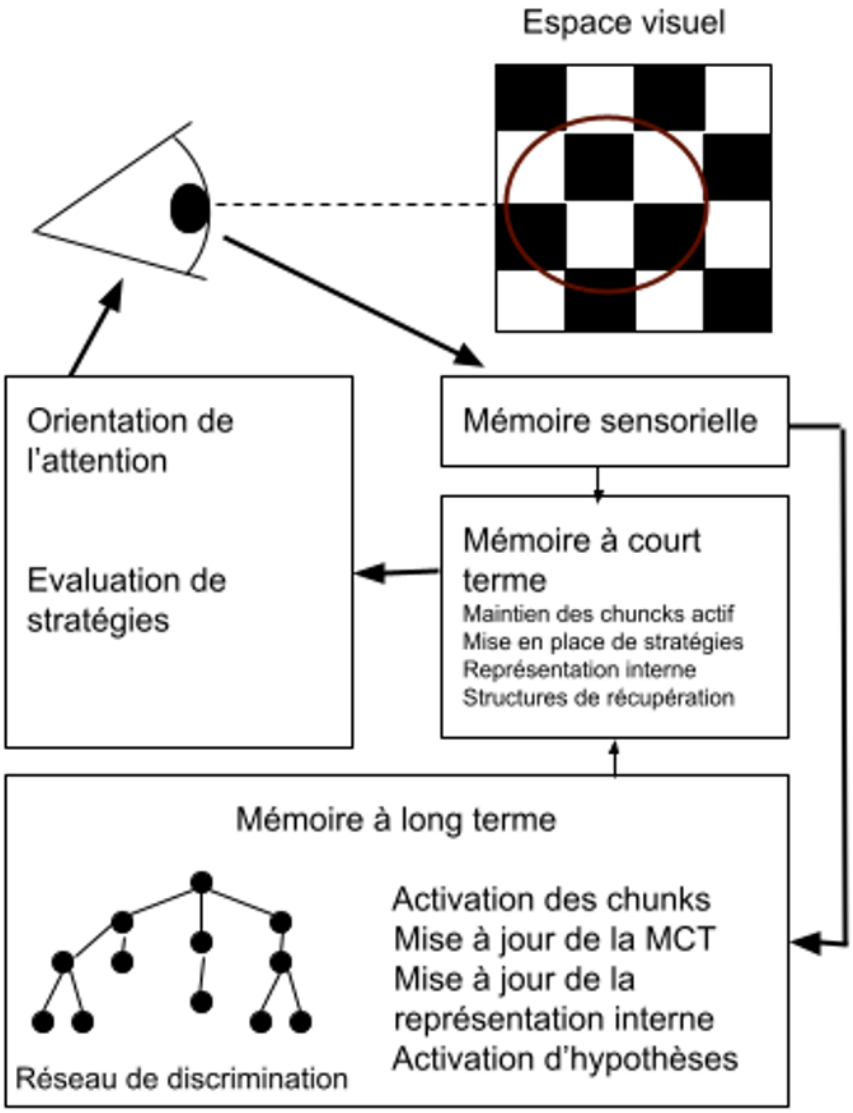


Figure 11 : Fonctionnement du modèle d'architecture cognitive CHREST d'après Gobet (1993).

2-5-3 Expertise perceptive visuelle

Les mécanismes perceptifs élémentaires ont un impact important sur ce que nous sommes capables de faire et jouent un rôle essentiel dans le développement de l'expertise. Ainsi, le développement de la perception est étroitement lié au développement de l'expertise.

2-5-3-1 La notion d'expertise perceptive visuelle

Les experts perceptifs sont ceux qui, ayant une expérience massive avec une catégorie particulière d'objets, peuvent faire des discriminations fines sur ces catégories visuelles.
Glemberg (Glemberg, 2010) précise que, dans le déploiement de l'expertise, la perception est essentiellement liée à l'action, car, à un certain niveau, les comportements des experts sont, par définition, dirigés vers une cible.

Comme le souligne Ericsson (2006), dans la problématique de la perception et de l'expertise, la question essentielle, qui a fait l'objet de nombreuses recherches est la suivante : comment l'expertise guide-t-elle le déploiement de la perception dans des tâches complexes ?
Une multitude de recherches suggèrent que l'apprentissage dans des domaines complexes implique fondamentalement d'apprendre à bien voir : la perception peut être ajustée pour mettre en évidence les caractéristiques pertinentes et les propriétés intrinsèques d'une tâche (Chase et Simon, 1973 ; Chi, Feltovich, et Glaser, 1981 ; Hegarty et al., 2009), pour contrôler la temporalité du raisonnement (Gleicher et al., 2013), pour recruter et restructurer les régions neuronales afin d'accomplir des tâches perceptives qualifiées (Gauthier et al., 2000) et même pour effectuer des tâches conceptuelles importantes dévolues à un réseau de traitement perceptivo-moteur (Clark, 1998). Ainsi, une partie de l'expertise peut résider dans l'apprentissage des compétences perceptives appropriées.

2-5-3-2 Plasticité cérébrale et expertise visuelle

La vision est le principal sens que nous employons lorsque nous explorons l'environnement qui nous entoure. L'importance de la vision est soulignée par le fait qu'un tiers des aires cérébrales sont consacrées au traitement des entrées visuelles. Il n'est pas étonnant que les humains soient bien équipés pour être des experts visuels. L'effet de l'expertise perceptive visuelle est traduit au niveau du cerveau par des modifications tant fonctionnelles que structurelles.

Deux mécanismes cérébraux distincts ont été décrits concernant la reconnaissance des objets : d'une part, le traitement holistique, qui correspond à un traitement perceptif des éléments en tant qu'unité découlant d'une logique globale, et d'autre part, le traitement analytique, où les diverses composantes sont perçues individuellement avant d'être transformées en un tout. La

performance des experts ressemble à bien des égards à un traitement holistique (Richler, Palmeri et Gauthier, 2012).

Sur le plan neurophysiologique, les travaux sur l'expertise perceptive visuelle se sont beaucoup intéressés aux études sur les mécanismes neurophysiologiques engagés dans la reconnaissance faciale. En effet, il existe un certain nombre de zones cérébrales impliquées dans les différentes composantes du traitement du visage, mais la plupart des chercheurs identifient la zone du visage fusiforme (FFA[1]) située dans la partie inférieure du lobe temporal comme le moteur principal du traitement holistique (Duchaine et Yovel, 2015). Non seulement la FFA est plus réactive aux visages que tout autre stimulus, mais des dommages du tissu cérébral dans ou autour de cette zone entraînent une prosopagnosie, une pathologie dans laquelle les personnes rencontrent des difficultés à percevoir les visages. Le constat important de différentes études est que la FFA est plus active chez les experts que chez les novices lorsqu'ils observent des objets dans leurs domaines de spécialisation (Ross et al., 2015 ; Bilalić et al., 2016), ce qui implique que la FFA est importante pour le traitement holistique, pas seulement pour la perception des visages.

L'expertise radiologique est typiquement identifiée comme holistique. L'étude de Harley et al. (2009) a examiné son lien avec la FFA. Les niveaux d'activation de la FFA ne différaient pas entre les radiologues lorsqu'ils devaient décider si la partie indiquée d'une radiographie du thorax contenait un nodule pulmonaire. En revanche, le complexe occipital latéral (LOC), la première structure cérébrale de la voie ventrale qui réagit aux objets et à leur forme, était négativement associé aux performances radiologiques, indiquant que seuls les radiologues les moins qualifiés avaient engagé le LOC. L'engagement de la FFA (zone importante pour les traitements holistiques) et le désengagement du LOC au fur et à mesure que l'expertise radiologique se développe correspondent bien à l'idée que l'expertise radiologique repose davantage sur des processus holistiques que sur des processus analytiques (Bilalić et al., 2016).

2-5-3-3 Perception et fiabilité de l'expert dans un environnement complexe

Au terme de ces éléments de réflexion, nous pouvons conclure que l'expertise perceptive visuelle, processus dont l'acquisition est longue et complexe, modifie le traitement de la tâche et permet d'aboutir à une surperformance. Cette expertise est le fruit d'une « engrammation »

[1] *Fusiform Face Area.*

au niveau de la plasticité cérébrale. L'expertise perceptive visuelle fonctionne préférentiellement de manière holistique plus qu'analytique, et, comme cela a été confirmé par la théorie des chunks, une bonne expertise visuelle dépend de processus descendants dans lesquels les connaissances et modèles stockés dans la mémoire guident la prise d'information de manière intuitive. La question sous-jacente au fonctionnement holistique de l'expert et à sa capacité de performance dans la reconnaissance de structures complexes est de déterminer dans quelle mesure l'expert peut faire confiance à son fonctionnement intuitif.

L'intuition est un élément central dans la prise de décision basée sur l'expertise. Simon (1992) a proposé une définition concise de l'intuition : "La situation a fourni un indice : cet indice a permis à l'expert d'accéder à des informations stockées en mémoire, et ces informations fournissent la réponse. L'intuition n'est ni plus ni moins que la reconnaissance". Selon la définition de Simon, le processus par lequel l'infirmière pédiatrique reconnaît qu'un nourrisson peut être gravement malade n'est pas différent en principe du processus par lequel elle remarquerait qu'un ami a l'air fatigué ou en colère ou de la manière dont un petit enfant reconnaît qu'un animal est un chien et non un chat.

Donc, confronté à un problème relevant d'un domaine dans lequel il possède une expertise, l'expert repère des indices par le biais de l'observation. Ces indices permettent d'activer des expériences ressemblantes parmi un répertoire de situations déjà expérimentées et stockées en mémoire. L'expérience la plus proche est retenue tandis que les autres seront inhibées. Sur la base de cette expérience, la stratégie retenue est confrontée à la situation présente. Cette confrontation se fait par stimulation mentale pour prédire/simuler, si, sur la base des connaissances passées et présentes, la stratégie qui avait été appliquée par le passé mène au succès dans le cas présent. Si non, une modification/adaptation acceptable est mise en œuvre. Si aucune adaptation n'est possible, la stratégie est abandonnée et la seconde expérience mémorisée la plus proche est alors activée. La reconnaissance et l'interprétation des indices sont donc un élément central de l'intuition. De ce fait, la digitalisation de la pratique orthodontique soulève la question de la fiabilité des indices proposés par ces environnements désormais numériques.

Plus précisément, le processus de prise de décision basé sur l'expertise a été abordé par le courant « Heuristics and Biases » (HB) qui fait preuve d'un scepticisme envers l'intuition de l'expert. Ce courant met en avant les défaillances de l'intuition, notamment la trop grande confiance de l'expert en son intuition. Kahneman et Klein (2009) formulent donc les conditions

de développement d'une bonne intuition et les conditions dans lesquelles cette intuition peut s'avérer fiable.

Pour qu'une intuition fiable puisse être construite, il faut qu'un apprentissage ait lieu dans un environnement dit de « haute validité ». Un environnement valide est un environnement comportant des relations stables entre des indices objectivement identifiables et des événements ultérieurs ou des rétroactions.

Selon ces auteurs, en pratique, pour qu'une intuition puisse être jugée fiable, il est essentiel que l'expert ait conscience des biais et des limites de son intuition. Ces auteurs mettent notamment en avant l'illusion du sentiment de validité et le sentiment de confiance injustifié qui impacte l'expert. Il est donc indispensable pour l'expert de remette en cause son intuition.

Le principal biais (i.e. déviation dans le traitement cognitif d'une information) est dû au mécanisme d'ancrage. Ce biais d'ancrage modifie la perception des indices disponibles et amène à les évaluer de manière variable, en fonction des dernières expériences vécues. Il crée une subjectivité due à l'instant. Ainsi, sur la base des mêmes informations, mais à des occasions distinctes, la décision peut être différente. L'expert se doit d'être conscient de l'impact de cette incohérence sur la validité des jugements basés sur l'intuition. Ce biais peut notamment se manifester dans l'estimation des grandeurs. L'estimation de différentes grandeurs et plus particulièrement des distances est au cœur de la pratique orthodontique. Dans leur travaux, Kahneman et Klein (2009) concluent que l'évaluation de la qualité probable d'un jugement intuitif nécessite une évaluation de la prévisibilité de l'environnement dans lequel le jugement est fait. Il apparait donc indispensable d'évaluer la validité des intuitions (et plus largement des traitements cognitifs) que l'expert met en œuvre lorsqu'il est confronté à ces nouveaux environnements numériques avec lesquels il est désormais amené à travailler.

2-6 L'essentiel à retenir

L'expert peut être défini comme une personne très habile et bien informée dans un domaine particulier, dans lequel il obtient des résultats supérieurs à ceux obtenus par la majorité de la population.

Il existe différents domaines d'expertise : perceptive, cognitive et sensorimotrice.

Une des caractéristiques des mécanismes cognitifs de l'expert est sa capacité à chercher la solution en se référant à une situation similaire sauvegardée dans la mémoire à long terme. La théorie des chunks en détaille le processus. Elle établit que la performance des experts semble en partie imputable à la reconnaissance de certains patterns spécifiques, en procédant par chunking, ce qui permet de retenir plus d'informations, puisque les unités ne sont plus mémorisées individuellement mais au sein d'unités signifiantes plus vastes.

Dans ce modèle de chunking, l'importance de la perception est soulignée. Elle constitue la voie donnant accès à l'information destinée au cerveau.

Une expertise perceptive visuelle permet, alors que l'on dispose d'une expérience massive sur une catégorie d'objets, de réaliser des discriminations fines sur les catégories visuelles concernées.

L'expertise perceptive visuelle repose davantage sur des processus holistiques que sur des processus analytiques. La zone du visage fusiforme (FFA) est très importante dans le traitement holistique.

Un environnement de haute validité et stable est une condition indispensable pour le développement et la fiabilité de l'intuition. Un changement d'environnement nécessite donc une évaluation de la validité de l'expertise mise en œuvre dans ce nouvel environnement.

2-7 Conséquence logique : une thèse en sciences cognitives

Ainsi, au fil des trois premiers articles publiés dans des revues médicales à destination des orthodontistes, j'ai esquissé un argumentaire de l'apport des sciences cognitives à la transition numérique et digitale dans la pratique de l'orthopédie dento-faciale, puis, dans la dernière partie du chapitre, cette problématique se connecte à la perception de l'expert et sa prise de décision. L'intérêt de ce travail consiste dès lors dans l'étude de la manière dont la virtualisation des données cliniques du patient impacte le travail de l'expert qu'est l'orthodontiste.
Cette problématique est abordée sous deux aspects. Le premier, constituant l'axe central de la thèse, a trait à l'expertise perceptive ; comme cela a été développé dans le chapitre ci-dessus, cette dernière repose davantage sur des processus holistiques que sur des processus

analytiques. Ce paramètre transposé à l'utilisation des modèles numériques en 3D rend particulièrement importante l'estimation des distances des différentes variables de malocclusions sur les modèles numériques en 3D par l'orthodontiste. L'objectif des protocoles expérimentaux a donc été de mesurer la fiabilité de cette estimation, puis de la comparer à celles effectuées sur la base de modèles réels. Au préalable, dans le chapitre 3, nous avions passé en revue la bibliographie concernant l'estimation des distances en réalité virtuelle.

Le seconde approche explore la manière dont l'interprétabilité des réseaux de neurones convolutifs pourrait aider le praticien dans son diagnostic des dysmorphoses crânio-faciales, en lui apportant un nouveau regard. En effet, l'hypothèse défendue est que les zones d'une image où le réseau de neurones prête attention pour discriminer l'image peuvent constituer une source d'information précieuse pour le praticien dans sa compréhension des dysmorphoses crânio-faciales et dans sa prise de décision thérapeutique. L'exemple choisi pour illustrer cette méthode est l'impact de la rétrognathie mandibulaire sur l'architecture crânio-faciale étudiée à partir de la téléradiographie de profil.

3 : La réalité virtuelle

3-1 Définition de la réalité virtuelle

Aujourd'hui, les environnements virtuels sont utilisés dans de nombreux domaines d'application et secteurs d'activités. Même si une perception spatiale exacte n'est pas nécessaire pour toutes les applications, pour certaines, l'estimation correcte de la distance et de la taille des objets présentés est fondamentale. En effet, la perception de la distance influe sur la performance, le comportement et les interactions des utilisateurs des technologies de réalité virtuelle (Wann et Mon-Williams, 1996). C'est notamment le cas dans le domaine médical, où la réalité virtuelle est utilisée pour la visualisation de données anatomiques ou pour suivre et prévoir des interventions chirurgicales (Pensieri et Pennacchini, 2016). La perception de la distance est généralement répartie en deux catégories : la perception de la distance égocentrique et la perception de la distance exocentrique. La perception de la distance égocentrique peut se définir comme l'évaluation de la distance entre un objet et un observateur, tandis que la perception de la distance exocentrique est l'évaluation de la distance entre deux objets. Il est souvent considéré que la distance égocentrique est sous-estimée dans les environnements virtuels et que la distance exocentrique est surestimée dans ces mêmes environnements (Peillard, Argelaguet et al., 2019).

La réalité virtuelle (RV) peut se définir comme un outil permettant « à une personne (ou plusieurs) une activité sensorimotrice et cognitive dans un monde artificiel, créé numériquement, qui peut être imaginaire, symbolique ou une simulation de certains aspects du monde réel » (Fuchs et al., 2006). Cette technologie permet, entre autres, d'accéder et d'interagir avec des objets en 3D et des représentations spatiales en 3D nommées « environnement virtuel » (EV) (Mikropoulos et Natsis, 2011). Trois caractéristiques sont considérées comme essentielles aux technologies de RV : l'interactivité, l'immersion et le sentiment de présence (Radianti et al., 2020). L'interactivité correspond à la capacité de l'environnement à évoluer en fonction des actions de l'utilisateur en temps réel (Steuer, 1992). Un élément important de cette interactivité est l'accord entre le feedback renvoyé par la technologie et les commandes de l'utilisateur (Slater et Wilbur, 1997).
L'immersion représente la caractéristique technique du système selon laquelle il peut fournir "une illusion de réalité inclusive, étendue, environnante et vivante aux sens d'un participant humain" (Slater et Wilbur, 1997). Ainsi la capacité d'une technologie de RV à immerger son

utilisateur dépend de plusieurs facteurs liés aux méthodes de présentation de l'information, tels que : le nombre de canaux sensoriels utilisés pour présenter l'information à l'utilisateur, la résolution et la fidélité de chaque canal utilisé, le degré d'exclusion de la réalité physique et la largeur du champ de l'environnement présenté, ect (Slater et Wilbur, 1997). Ainsi, les technologies de RV peuvent être classées en fonction de leur capacité d'immersion. Par exemple, un casque de RV sera considéré comme une technologie immersive tandis qu'un ordinateur de bureau sera considéré comme une technologie non-immersive.

Le sentiment de présence comparé à l'immersion correspond à l'expérience subjective de se sentir présent au sein de l'EV (Nowak et Biocca, 2003). Il dépend de la perception que l'utilisateur possède de l'EV et non de l'environnement existant dans le monde physique (Steuer, 1992). Il y a donc, quant au sentiment de présence, une composante individuelle qui peut varier d'un individu à un autre.

Des facteurs propres à chaque individu sont ainsi déterminants pour ce qui est du sentiment de présence. Gilbert (2016) indique qu'un niveau d'immersion et une grande fidélité devraient augmenter le réalisme perçu et, par conséquent, induire un plus grand sentiment de présence à condition que le contenu dispensé en RV soit bien élaboré. Ainsi, le réalisme perçu par l'apprenant est un autre facteur ayant son importance dans les dispositifs utilisant la RV (Kalantari et Neo, 2020). En effet, le réalisme perçu serait associé au sentiment de présence, à l'engagement et au plaisir des apprenants lors de l'utilisation de la RV (McGloin et al., 2011 ; Ribbens et al., 2016 ; Weber et al., 2021).

Le réalisme perçu correspond à l'évaluation par l'utilisateur de la « précision avec laquelle une représentation narrative représente les personnes et les événements tels qu'on pourrait s'y attendre s'ils existaient ou se produisaient dans le monde réel, ou dans un monde hypothétique médiatique » (Busselle, 2020). Ainsi, le réalisme perçu dépend de la perception de l'utilisateur concernant : le degré de réalisme des objets, sons et scènes de l'EV, le degré de réalisme et de facilité des interactions avec l'EV ainsi que le degré de cohérence et de plausibilité des scénarios et des personnages présentés au sein de l'EV (Weber et al., 2021). Le réalisme perçu résulterait de l'interaction entre les attentes de l'utilisateur et les propriétés techniques de l'EV (immersion, interaction et fidélité) (Weber et al., 2021).

Enfin, la fidélité correspond au degré de similitude entre la situation virtuelle et la situation réelle (Hays et Singer, 1989). La fidélité d'un EV peut être physique ou fonctionnelle (Hamstra et al., 2014 ; Scerbo et Dawson, 2007). La fidélité physique correspond au degré de ressemblance entre les composantes pouvant être décrites physiquement dans l'environnement

simulé et celles présentes dans l'environnement réel, comme la couleur, les matériaux, les objets. La fidélité fonctionnelle correspond, quant à elle, à la similitude entre l'environnement simulé et l'environnement réel pour ce qui est de l'exécution de processus et de comportements de l'environnement concerné.

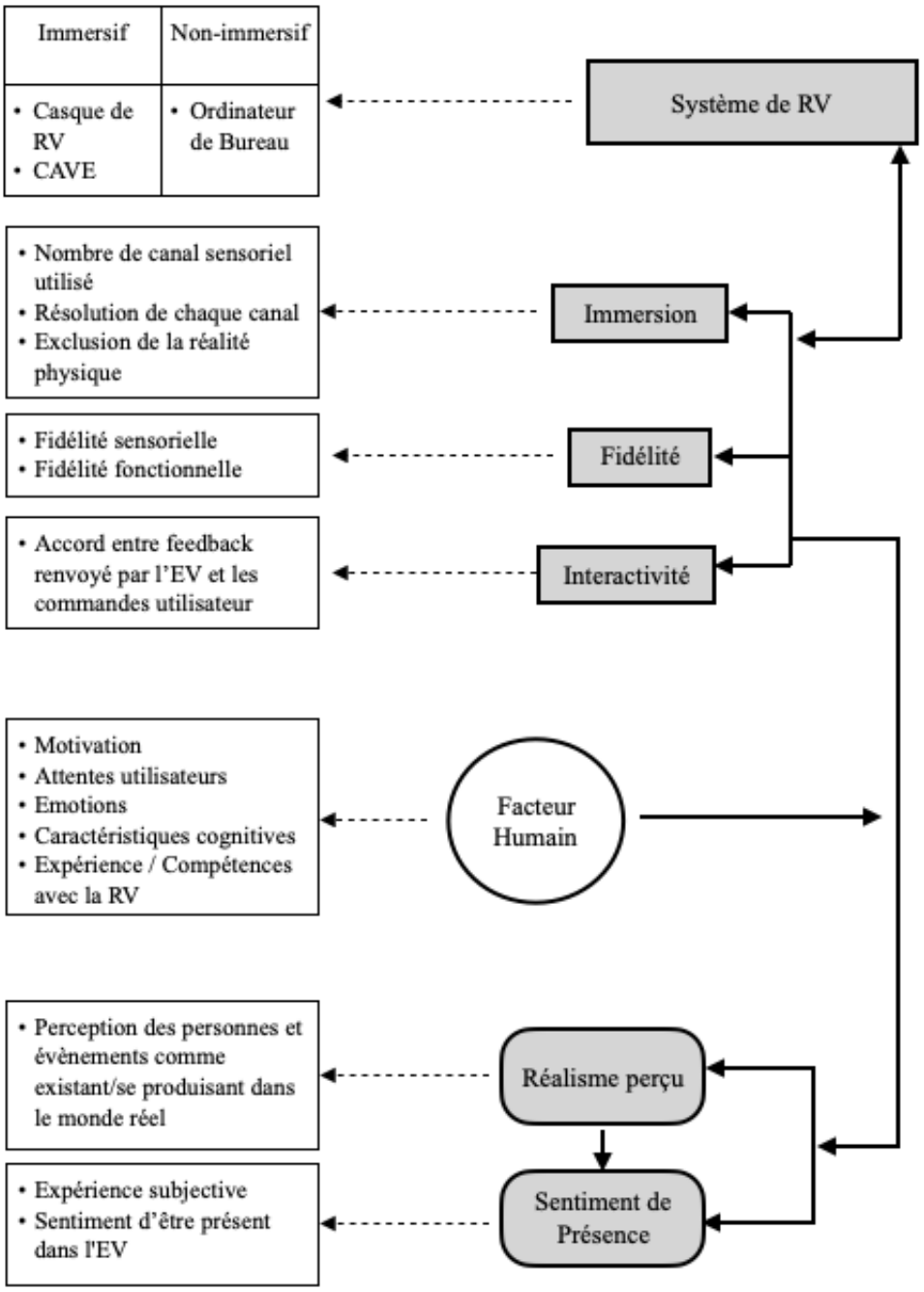


Figure 12 : Schéma des caractéristiques de la réalité virtuelle (ci-après RV).

Ce schéma (figure 12) présente les caractéristiques majeures des technologies de RV.
L'immersion, la fidélité et l'interactivité varient en fonction du type de système de RV utilisées.

Ces facteurs (immersion, fidélité, interactivité) associés aux caractéristiques de l'utilisateur (facteur humain), ont un impact sur le réalisme perçu et le sentiment de présence de l'utilisateur.

3-2 Perception de la distance et réalité virtuelle

Même si, pour certaines applications, les technologies de RV visent à offrir une perception de l'EV qui se rapproche le plus possible de la perception du monde réel, des différences d'estimation de distance existent entre situations réelles et virtuelles. En effet, la perception de la distance est un élément qui reste insatisfaisant au sein d'EV quelle que soit la technologie de présentation utilisée.

3-2-1 Perception des distances dans le monde réel

La lumière réfléchie par les objets traverse différents milieux avant d'atteindre la rétine, et il en résulte une image en 2D renversée. Par conséquent, notre perception visuelle en 3D est une construction de notre cerveau. Pour cela, le cerveau utilise des indices de profondeur permettant d'obtenir des informations sur les relations spatiales entre les objets et l'environnement.

Il existe plusieurs classifications relatives à ces indices de profondeur. Rousset (2018) indique que la classification de Cutting et Vishton (1995) est la plus utilisée. Deux types d'indices de profondeur peuvent alors être différenciés : les indices picturaux et les indices non-picturaux (Cutting et Vishton, 1995).

Les indices picturaux correspondent aux informations pouvant être identifiées sur des scènes statiques, tandis que les indices non-picturaux correspondent aux informations provenant du mouvement et du système oculomoteur binoculaire humain.

Cinq indices picturaux ont été recensés par Cutting et Vishton (1995) : l'occlusion, la taille relative, la densité de texture relative, l'angle d'incidence et la perspective aérienne (figure 13).

L'occlusion signifie que, lorsqu'un objet cache une partie d'un autre, l'objet partiellement caché est perçu comme étant plus éloigné de l'observateur. La taille relative renvoie au principe selon lequel plus un objet est éloigné, plus l'image sur notre rétine est petite, tandis que la densité de texture relative correspond à un phénomène de distance, selon lequel plus un objet est éloigné,

plus sa texture de surface paraît dense. L'angle d'incidence signifie qu'en augmentant la distance, la base des objets se rapproche de la ligne d'horizon. Enfin, la perspective aérienne indique que plus un objet est éloigné, plus il sera perçu difficilement, entouré d'une sorte de voile bleuté et floutant. Rousset (2018) ajoute à ces éléments la perspective linéaire, selon laquelle plus un objet est éloigné, plus les lignes parallèles de ses contours se rapprochent.

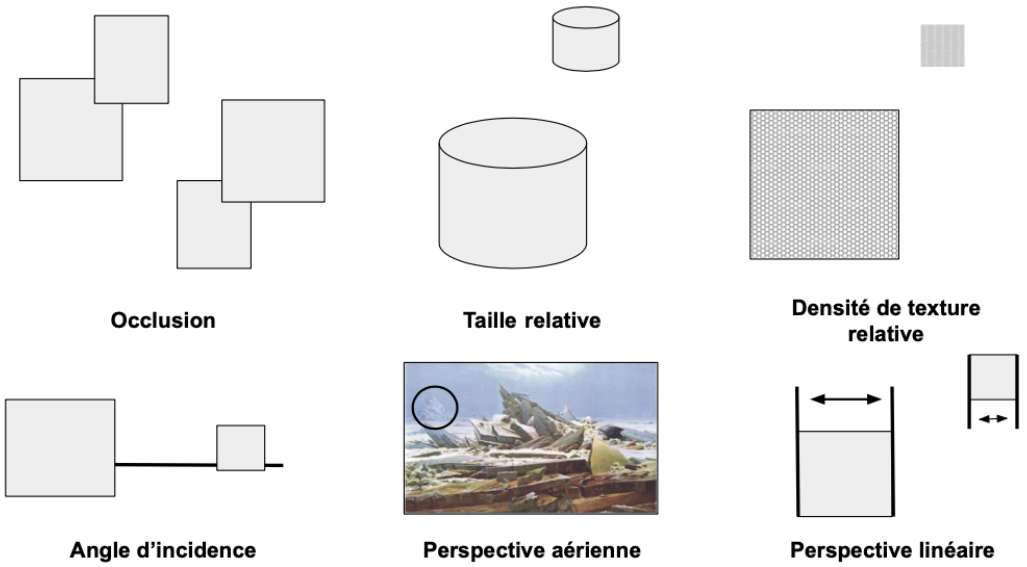


Figure 13 : Schéma des indices de profondeur picturaux d'après Cutting et Vishton (1995).

Quatre indices non-picturaux ont également été recensés (Cutting et Vishton, 1995) : la parallaxe du mouvement, la convergence, l'accommodation et la disparité binoculaire. La parallaxe du mouvement correspond au fait que, lors d'un mouvement, l'image d'un objet statique bouge simultanément. Ainsi, plus l'objet est éloigné, plus il bougera lentement. De plus, lorsqu'un objet est observé, les objets situés derrière lui auront un sens identique à celui du mouvement de l'observateur, tandis que ceux situés devant ce même objet auront un sens contraire à ce mouvement.

La convergence et l'accommodation permettent d'obtenir des informations sur la distance par la combinaison des informations proprioceptives des muscles qui rendent possibles ces phénomènes. La convergence indique que les yeux convergent grâce aux muscles extra-oculaires pour fixer un objet proche et l'accommodation signifie le changement de convexité du cristallin effectué par les muscles ciliés afin d'effectuer une mise au point sur un objet.

Enfin, la disparité binoculaire renvoie à la différence des images perçues par l'œil droit et par l'œil gauche. Il s'agirait de l'indice fournissant le plus d'informations sur la distance, notamment sur des distances proches de l'observateur.

De plus, en fonction de la gamme de distance à percevoir, plusieurs espaces ont pu être définis (Cutting et Vishton, 1995) : l'espace personnel, l'espace d'action et l'espace éloigné. En fonction de l'espace considéré, les différents indices picturaux et non-picturaux auront plus ou moins d'importance (figure 14).

Ainsi, dans l'espace personnel allant de l'observateur à 2 mètres de distance, les indices effectifs seront l'occlusion, la taille relative, la densité de texture relative, la parallaxe du mouvement, la convergence et l'accommodation, ainsi que la disparité binoculaire. Dans l'espace d'action allant de 2 à 30 mètres, les indices pertinents seront l'occlusion, la taille relative, la densité de texture relative, l'angle d'incidence, la parallaxe du mouvement et la disparité binoculaire. Enfin, dans l'espace éloigné, l'occlusion, la taille relative, la densité de texture relative, l'angle d'incidence et la perspective aérienne seront les seuls indices effectifs (Cutting et Vishton, 1995).

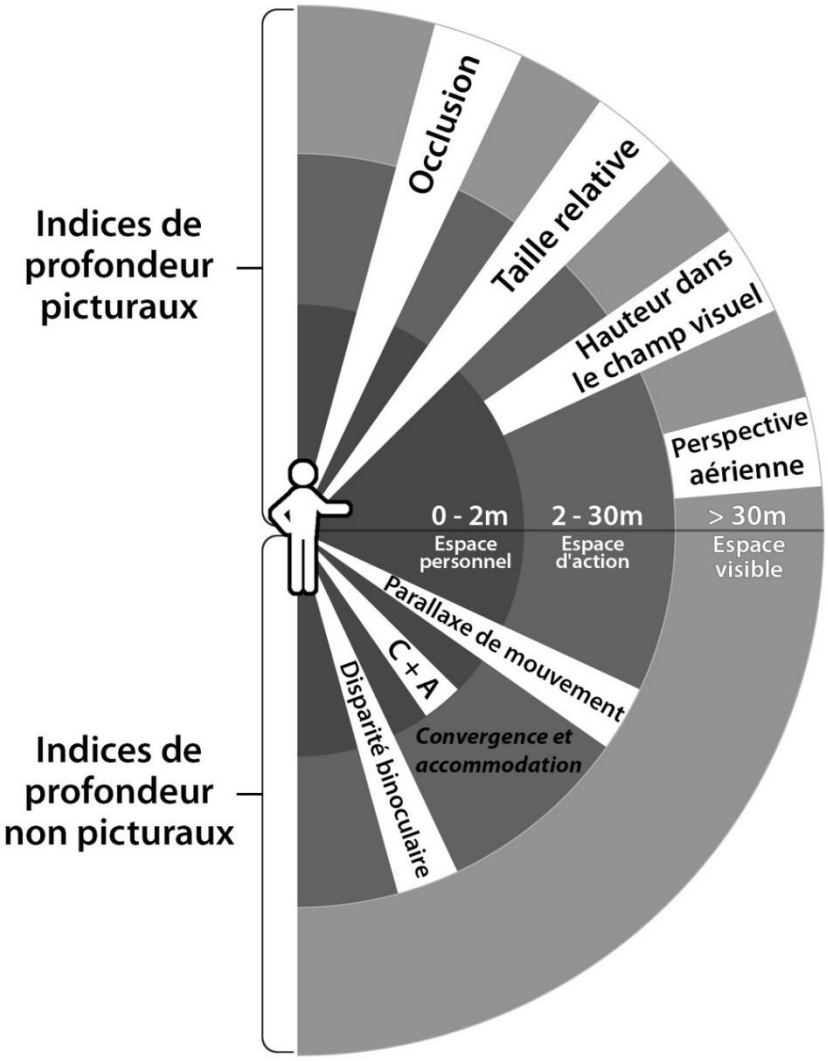


Figure 14 : Schéma des portées effectives des différents indices de profondeur inspiré de cutting et Vishton (1995).

La perception visuelle est aussi influencée par les facteurs environnementaux et individuels (Rousset, 2018). En effet, la taille de notre corps fournit un cadre de référence pour la perception de notre monde (Harris et al., 2015) ce qui impacte donc la perception des distances tout comme de la taille des objets. Van der Hoort et al. (2011) ont montré que la perception de la taille de notre corps influe sur la perception de la taille de l'environnement. Des images de corps artificiels de différentes tailles étaient montrées à des participants à travers des lunettes, donnant aux participants l'impression que le corps présenté était le leur. Par rapport à une taille de corps normal, un corps plus petit entraine une surestimation des objets dans

l'environnement, tandis qu'un corps plus grand entraine une sous-estimation des mêmes objets. De plus, l'environnement dans lequel la personne se situe a un impact sur la perception de la distance. Par exemple, le biais de contraction suggère une tendance à surestimer les petites distances et à sous-estimer les grandes lorsqu'une comparaison avec une distance intermédiaire est possible (Link et Poulton, 1991).

L'état psychologique et l'objectif de l'action ont également leur importance. La théorie de l'effort pour la perception des profondeurs indique que l'estimation de la distance est influencée par l'effort attendu pour atteindre et/ou pour voir un objet donné (Proffitt, 2006). Par exemple, Proffitt et al. (2003) ont montré que porter un sac lourd au cours d'une tâche liée à l'atteinte d'un objectif fait paraître la distance à parcourir comme étant plus grande. D'autres facteurs, comme l'âge, l'expérience avec la tâche concernée ou encore l'état du système visuel peuvent aussi rentrer en ligne de compte (Rousset, 2018).

Ainsi, notre perception des distances dans le monde réel est influencée par de nombreux facteurs. Elle dépend des informations renvoyées par les éléments présents dans la scène observée (indices picturaux), par les informations induites par le système binoculaire humain (indices non-picturaux) qui auront plus ou moins d'importances en fonction de la gamme de distance à estimer. Enfin, elle sera influencée par l'environnement et l'état psychologique de la personne estimant les distances. Tous ces facteurs ont une importance dans la perception des distances au sein d'environnements réels mais également d'environnements proposés par la réalité virtuelle.

3-2-2 Perception de la distance en réalité virtuelle

La perception des distances est erronée. Dans leur revue de la littérature portant sur la perception des distances au sein d'EV, Renner et al., (2013) indiquent que les distances séparant un objet d'un observateur (distances egocentriques) sont généralement sous-estimées. De plus, dans plusieurs études, les distances séparant deux objets (distances exocentriques) sont surestimées. Une première étude (Waller, 1999) a comparé la distance perçue entre 2 cubes présentés dans un EV simple à l'aide d'un casque de RV ou d'un ordinateur de bureau. Cette étude a montré que la distance était surestimée. D'autres auteurs, (Wartenberg et Wiborg, 2003) ont étudié la perception de la distance entre 2 points au sein de trois environnements : l'environnement réel, un EV présenté avec un système de CAVE et un EV présenté à l'aide d'un

ordinateur de bureau. Les distances ont été mal perçue quel que soit l'environnement présenté et plus les distances à estimer étaient grande plus les erreurs d'estimation augmentaient. Dans l'environnement réel et dans l'EV présenté dans un système de CAVE une prévalence de surestimation des distances était présente tandis que dans l'EV présenté avec un ordinateur de bureau les distance étaient autant surestimées que sous-estimées (selon les participants).

Par ailleurs, les distances perçues entre plusieurs éléments d'un objet complexe seraient également surestimées. Tourancheau et al. (2012) ont analysé l'estimation de la distance, et la perception de la profondeur par des radiologues et des étudiants en médecine par rapport à des tâches impliquant l'analyse spatiale d'un puzzle en 3D (figure 15) présenté à l'aide de 2 systèmes de RV stéréoscopique. Les tâches de perception de la profondeur ont été réussies par 90% des participants cependant, les distances entre les éléments du puzzle ont été surestimées au sein des deux systèmes de RV.

Figure 15 : Puzzle en 3D et son modèle virtuel utilisé dans l'étude de Tourancheau et al. (2012).

Ces études montrent que la perception des distances au sein de systèmes de réalité virtuelle est inexacte. De plus, l'erreur varie selon le type de distances à estimer : les distances sont généralement sous-estimées dans le cas de distances égocentriques et surestimées dans le cas de distance exocentrique ainsi que dans le cas d'estimation de distances au sein d'un même objet.

Renner et al., (2013) ont regroupé en 4 classes de facteurs les éléments pouvant impacter l'estimation des distances en RV : les méthodes d'évaluation, les facteurs techniques, la

composition de l'environnement et le facteur humain (figure 16). Nous allons donc détailler plus précisément chacune de ces classes.

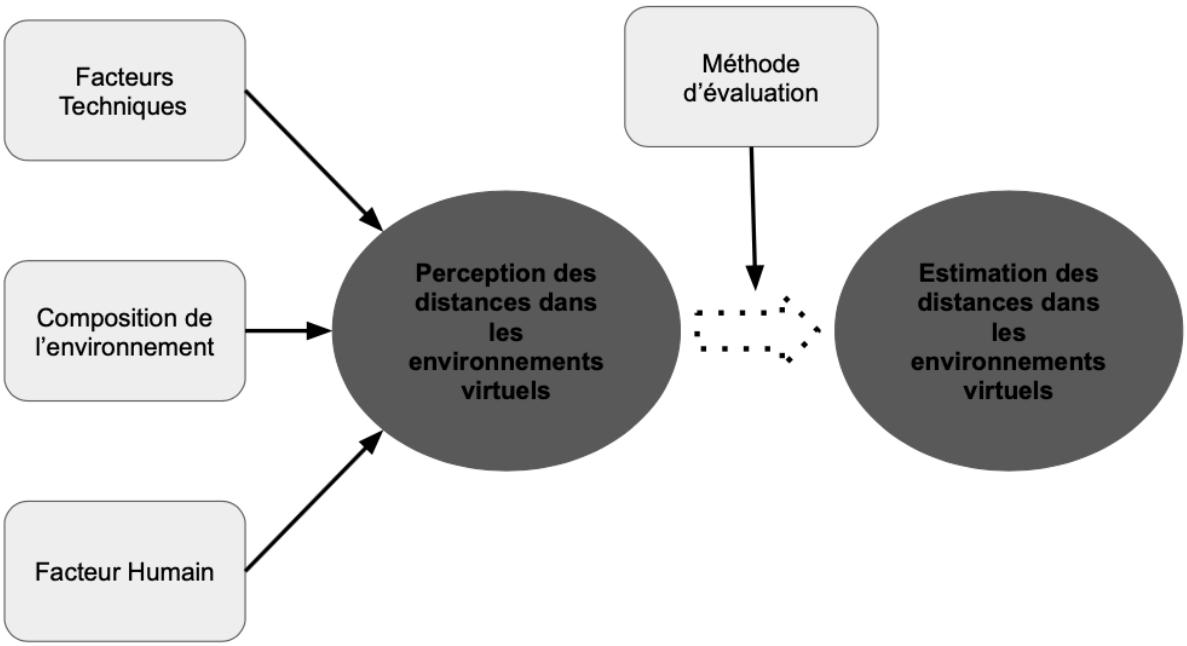


Figure 16 : Schéma des quatre groupes de facteurs influant sur la perception de la distance et les estimations des distances basées sur Renner et al. (2013).

Méthode d'évaluation

Plusieurs méthodes ont pu être utilisées dans la littérature pour relever les estimations des distances : l'estimation verbale, les évaluations basées sur le comportement moteur et l'ajustement de la perception. Chaque méthode a des avantages et des inconvénients, et c'est révélées avoir un impact sur l'estimation des distances en RV (Renner et al., 2013). L'estimation verbale consiste en l'énonciation des distances ou de la taille d'un objet avec une unité de mesure. Même si elle demeure la méthode d'évaluation la plus simple, elle est moins efficace que les méthodes basées sur le comportement moteur (Knapp et Loomis, 2003). L'une des méthodes se basant sur le comportement moteur est la marche aveugle, qui demande à l'utilisateur de se rendre à une cible les yeux bandées. Cette méthode est efficace pour les distances inférieures à 20 m, alors que les distances supérieures à cette limite sont grandement sous-estimées (Willemsen et Gooch, 2002). La triangulation constitue une autre méthode basée sur le comportement moteur (Fukusima et al., 1997). La triangulation par pointage demande à l'utilisateur de pointer une cible les yeux fermés en parcourant un chemin passant devant la

cible. La marche triangulée consiste à demander à l'utilisateur guidé par l'expérimentateur de se rendre vers une cible vue auparavant en suivant un chemin oblique. Cette méthode est performante pour ce qui est des longues distances (Ghinea et al., 2018). Une autre méthode est celle du jugement d'affordance. Elle consiste à estimer la taille ou la distance à partir des informations du corps de l'utilisateur et de ses potentielles actions (Stefanucci et al., 2015). Par exemple, on peut demander à l'utilisateur s'il est à même de saisir un objet avec sa main ou de toucher un objet sans bouger. Dans le cadre de la bissection, il est demandé à l'utilisateur de trouver le point central entre lui et un objet ou entre 2 objets (Meng et al., 2006). Enfin, l'ajustement de la perception évalue la perception de la distance en demandant à l'utilisateur de déplacer un objet vers une position cible (Ghinea et al., 2018).

Les différentes méthodes utilisées ont donc un impact sur l'estimation des distances. En effet, même si l'estimation verbale demeure la plus simple, elle serait moins performante que les méthodes basées sur le comportement moteur ou l'ajustement de la perception. Il est alors nécessaire d'identifier la méthode la plus efficace en fonction de la nature de la tâche à effectuer et du type de distance à estimer.

Facteurs Techniques

Comme nous l'avons vu, d'une façon générale, les distances égocentriques sont sous-estimées alors que les distances exocentriques sont surestimées. Les facteurs techniques (facteurs liés aux dispositifs de RV utilisés) peuvent cependant affecter l'estimation des distances.

Renner et al. (2013) ont trouvé qu'une sous-estimation de la distance égocentrique était présente quel que soit le niveau d'immersion des systèmes de RV. Cependant, cela ne signifie pas que l'immersion n'a aucun effet sur l'estimation de la distance. En effet, la sous-estimation de la distance égocentrique est supérieure avec un grand écran (non-immersif) qu'avec un système de CAVE (immersif) (Klein et al., 2009). L'immersion exerce également un effet sur la perception des distances exocentriques. Des études ont montré que les distances exocentriques étaient plus largement surestimées avec un ordinateur de bureau qu'avec un système immersif (Naceri et al., 2010 ; Waller, 1999). L'immersion serait donc un facteur améliorant l'estimation des distances en RV.

De plus, la présence de feedback améliore la perception des distances égocentrique et exocentrique (Makin et al., 2019 ; Renner et al., 2013 ; Waller, 1999). Waller et al. (1999) ont montré qu'un feedback sur l'erreur d'estimation permettait d'améliorer la perception de la distance egocentrique. De plus, dans une tâche où les participants devaient évaluer la distance les séparant d'un cube virtuel avec leur main, il a été montré qu'un feedback haptique de retour de force ou la combinaison d'un feedback haptique avec retour de force et vibration améliore l'expérience d'utilisation du système ainsi que l'estimation de la distance egocentrique (Makin et al., 2019).

Enfin, les indices non-picturaux provenant du système binoculaire (convergence, accommodation et disparité binoculaire) sont impactés par les technologies de RV et peuvent donc ne pas renvoyer d'informations pertinente sur la distance.

Dans les systèmes de RV stéréoscopiques (permettant la disparité binoculaire), l'accommodation s'effectue sur le plan de l'écran (Huckauf, 2005). A contrario, le point de convergence dépend de la parallaxe des objets présentés, qui peuvent apparaître devant (parallaxe négative), sur (parallaxe nulle) ou derrière l'écran (parallaxe positive) (Bruder et al., 2016 ; Renner et al., 2013) (figure 17). Ainsi, un conflit est susceptible de se produire entre l'accommodation et le point de convergence. Il a été reporté que, lorsque ce conflit a lieu, c'est l'accommodation qui est priorisée en tant qu'indice de profondeur (Swenson, 1932). Ainsi les systèmes de RV stéréoscopiques pourraient entraîner des informations de profondeur contradictoires, notamment en cas de parallaxe négative ou positive. Par conséquent, la perception de la distance serait mal orientée par l'accommodation (Bruder et al., 2016 ; Drascic et Milgram, 1996 ; Renner et al., 2013).

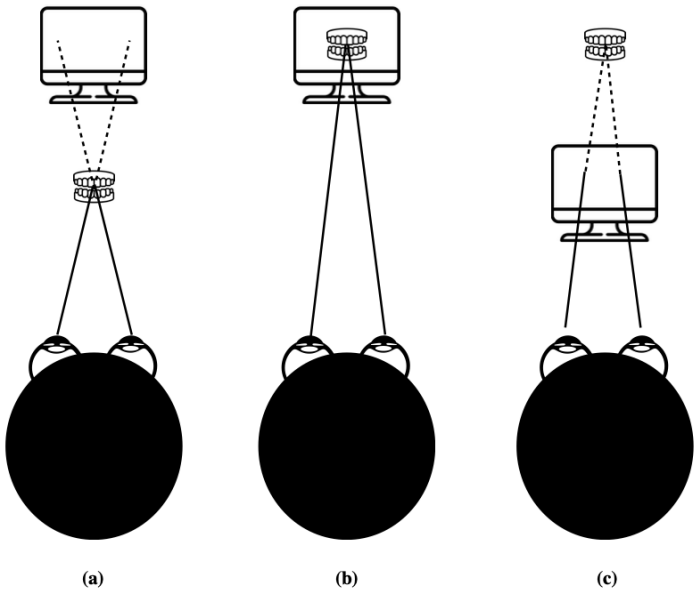


Figure 17 : Schéma des différentes parallaxes des objets : (a) parallaxe négative, (b) parallaxe nulle et (c) parallaxe positive.

Concernant la disparité binoculaire, les informations renvoyées par celle-ci ne seraient utilisables que dans certains systèmes de RV. En effet, le fait de manquer d'information sur la profondeur par la vision binoculaire dans les systèmes de RV non stéréoscopique entraine une baisse de la performance dans l'estimation des distance (Wainman et al., 2018). Cet effet serait d'autant plus marqué dans l'analyse d'objets proches de l'utilisateur (Renner et al., 2013).

Par conséquent, le type de dispositif (immersif ou non immersif, stéréoscopique ou non stéréoscopique) impacte l'estimation des distances d'une part directement par le niveau d'immersion et d'autre part par l'impact sur les indices visuel non-picturaux. L'absence de feedback serait aussi un facteur diminuant la performance d'estimation des distances au sein de système de RV. Concernant la qualité graphiques, Renner et ses collaborateurs (2013) indiquent que l'influence de celle-ci sur l'estimation des distances demeure discutée. En effet, une étude a montré que la qualité des graphiques n'impacte pas l'estimation de la distance en RV (Thompson et al., 2004), tandis que, selon une autre étude, la qualité des graphiques a un effet sur l'estimation de la distance (Kunz et al., 2009). Cependant, ces facteurs seuls ne suffisent pas à expliquer la mauvaise perception des distances au sein d'un EV, et la composition de l'environnement virtuel joue également un rôle important.

Facteurs de composition de l'environnement virtuel

La composition (nombre d'éléments présents, nature de ces éléments, position de ces éléments) de l'EV joue également un rôle dans la perception des distances. Renner et ses collaborateurs (2013), en passant en revue la littérature sur le sujet, montrent que l'ajout d'indices de profondeur non picturaux (l'occlusion, la taille relative, la densité de texture relative, l'angle d'incidence et la perspective aérienne) provenant des différents éléments de la scène présentée améliore la perception de la distance egocentrique, ce qui peut expliquer la sous-estimation dans les EV simples (ne possédant pas beaucoup d'éléments) mais pas dans les EV complexes (présentant un grand nombre d'éléments). Cependant, Waller (1999) a trouvé que les indices de profondeur picturaux avaient un faible impact sur la surestimation des distances exocentriques. Il convient toutefois de noter que l'EV utilisé dans cette étude était simple ce qui peut expliquer le fait que les indices picturaux présents aient eu un faible impact. Le nombre d'indices picturaux présent serait un des facteurs majeurs exerçant un rôle dans la perception des distances au sein d'EV.

De plus, toutes les caractéristiques des objets ne seraient pas importantes pour l'estimation des distances. Une étude a montré que les indices ou caractéristiques visuels ayant pour objectif de guider une action (orientation des objets, parallaxe du mouvement, distribution des ombres et motifs géométriques de l'objet) diminuent la sous-estimation de la distance égocentrique, tandis que les indices visuels relatifs à l'identification d'un objet (reflet, couleur, forme) n'ont pas d'impact sur l'estimation de la distance égocentrique en RV (Ahn et al., 2021). La position des éléments par rapport à l'utilisateur a aussi son importance. En effet, il a été montré que plus un objet est présenté sur un coté du champ visuel de l'observateur, plus la distance est sous-estimée (Peillard, Thebaud et al., 2019). Les caractéristiques des objets ayant pour objectif de guider l'action ainsi que leurs positions sont donc des facteurs impactant l'estimation des distances.

L'espace dans lequel est effectué l'estimation de la distance (espace personnel [0-2 mètres], espace d'action [2-30 mètres], espace éloigné [plus de 30 mètres]) aurait lui aussi son importance. Ahn et al. (2020) ont pu observer une sous-estimation de la distance égocentrique dans l'espace d'action et l'espace éloigné, et une surestimation de cette distance dans l'espace personnel. Ces résultats sont en accord avec ceux d'Alexandrova et al. (2010). Cependant, la

question de l'estimation de la distance égocentrique dans l'espace personnel demeure débattue. En effet, une étude a montré une sous-estimation de la distance dans l'espace personnel (Napieralski et al., 2011), tandis qu'une autre étude a observé, pour certains sujets une surestimation et pour d'autres une sous-estimation dans l'espace personnel (Armbrüster et al., 2008). Les distances seraient donc globalement sous-estimées dans l'espace d'action alors qu'elles seraient soit surestimées soit sous-estimées dans l'espace personnel.

Par ailleurs, la présence d'un avatar représentant l'utilisateur au sein de l'EV permettrait d'améliorer la perception de la distance (Ebrahimi et al., 2018 ; Renner et al., 2013 ; Ries et al., 2008) à condition que l'utilisateur ait un sentiment de présence, c'est-à-dire à condition qu'il perçoive l'avatar comme étant son propre corps et non comme un objet différent. En effet, il a été montré qu'après le calibrage de l'avatar, l'estimation de la distance égocentrique s'améliore et que l'estimation de cette distance s'améliore à mesure qu'augmente le réalisme de l'avatar (Ebrahimi et al., 2018). La taille de l'avatar aurait aussi un impact sur l'estimation de la taille des objets au sein d'EV. Un avatar plus grand que notre corps entraine une sous-estimation de la taille des objets, tandis qu'un avatar plus petit que notre corps entraine une surestimation de la taille des objets (Weber et al., 2020).

Enfin, si l'EV est une réplique virtuelle correcte d'un environnement connu et que l'utilisateur a conscience qu'il s'agit d'une réplique, cela induit une bonne estimation de la distance (Combe et al., 2021 ; Interrante et al., 2006 ; Renner et al., 2013). De plus, utiliser une réplique virtuelle comme EV de transition vers un EV inconnu améliore l'estimation de la distance et le sentiment de présence dans l'EV inconnue (Steinicke et al., 2009).

Ainsi, la composition de l'environnement virtuel joue un rôle majeur dans la perception des distances. La possibilité de disposer d'indices visuels picturaux est un élément important dans l'estimation des distances**.** De plus, l'espace dans lequel est effectué l'estimation des distances joue aussi un rôle important. Enfin, l'ajout d'avatar ou encore le fait que l'EV soit une réplique virtuelle sont des facteurs améliorant l'estimation des distances. Une dernière classe d'éléments susceptibles d'affecter l'estimation des distances au sein d'EV est constituée par le facteur humain.

Facteurs humains

Concernant le facteur humain, Renner et ses collaborateurs (2013) rapportent dans leur revue de la littérature des preuves indirectes indiquant que le sentiment de présence améliore l'estimation de la distance en RV. De plus, ces auteurs ajoutent que les différences interindividuelles dans l'estimation des distances en dehors du sentiment de présence perçue par l'utilisateur viendraient entre autres de l'expérience avec la RV (Renner et al., 2013).

Par ailleurs, les facteurs psychologiques influent eux aussi l'estimation de la distance en RV. Il a, par exemple, été montré que si un objet parait menaçant, l'objet sera perçu comme étant plus proche (Takahashi et al., 2013).

Les capacités visuo-spatiales jouent également un rôle dans la perception de la profondeur dans les modèles virtuels. En effet, de faibles capacités visuo-spatiales entrainent une baisse de la performance dans l'étude de modèle virtuel alors que des capacités visu-ospatiales fortes ou faibles n'impactent pas la performance sur des modèles physiques (Wainman et al., 2020).

En résumé, l'estimation des distances au sein de dispositifs de RV est un problème complexe mettant en jeu de nombreux facteurs comme le niveau d'immersion, la présence ou l'absence de feedback, la richesse de l'environnement et des objets présentés ou l'état psychologique et la familiarité de l'observateur avec la RV.

Comme nous l'avons vu le niveau d'immersion joue également un rôle, et selon le degré d'immersion certains facteurs auront plus ou moins d'impacts sur l'estimation des distances en RV.

3-2-3 Evaluation des distances au sein de dispositifs immersifs

Concernant la sous-estimation de la distance égocentrique dans des EV immersifs, ce sont surtout les caractéristiques des casques de RV qui ont été étudiés. Renner et al. (2013) ont constaté que la combinaison de la réduction du champ visuel induit par les casques de RV, de la masse du casque, du temps d'inertie et de la sensation de porter un casque était susceptible de contribuer à la sous-estimation de la distance egocentrique. Cependant, ces facteurs non

combinés n'influent pas sur l'estimation des distances. Ces résultats sont en accord avec ceux obtenus par (Combe et al., 2021), qui ont montré que le poids seul du casque de RV n'influençait pas l'estimation pour des distances courtes.

De plus, avec les casques de ce type, le champ visuel virtuel appelé « champ visuel géométrique » (CVG) doit être identique au champ visuel d'affichage (CVA) (figure 18) (Steinicke et al., 2011). Cet ajustement peut s'avérer difficile car les valeurs de CVA indiquées par le fabricant diffèrent des valeurs réelles. Ainsi la non-correspondance entre le CVA et le CVG entraîne un agrandissement ou une réduction de l'image, impactant l'estimation de la distance égocentrique (Renner et al., 2013, Steinicke et al., 2011). Ces résultats sont en accord avec ceux de Waller (1999) indiquant que le mauvais calibrage CVA/CVG était un facteur de la surestimation de la distance égocentrique. Cependant, les auteurs ont observé que la distance pouvait être sous-estimée même quand un calibrage minutieux était assuré. La divergence CVA/CVG n'est donc qu'un facteur parmi ceux diminuant la perception de la distance en RV (Renner et al*.,* 2013).

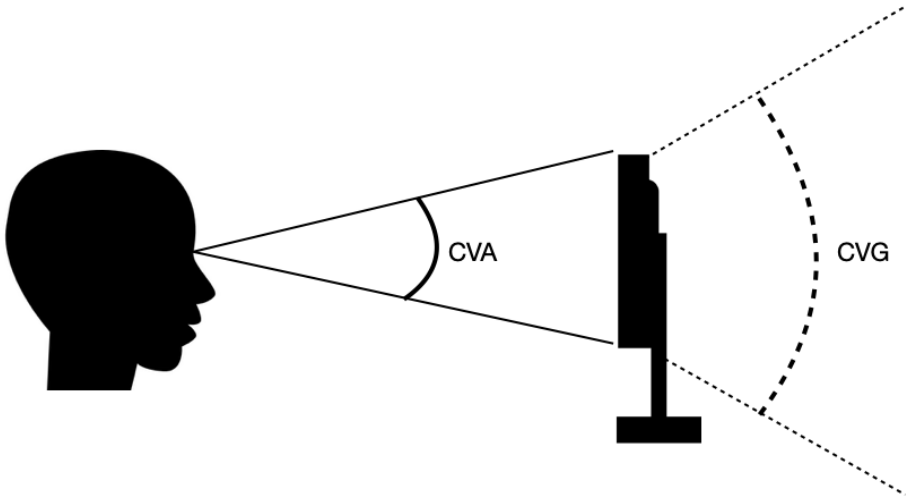


Figure 18 : Schéma du champ visuel d'affichage (CVA) et du champ visuel géométrique (CVG)

Par ailleurs, l'utilisation de casques de RV pour l'évaluation de la distance exocentrique entrainerait une compression de la distance dans le plan de la profondeur mais pas dans le plan horizontal (Geuss et al., 2012). Enfin, concernant les casques de RV, il a été montré qu'ils

rendaient possible une meilleure estimation de la distance que les systèmes CAVE pour des distances allant de 7 à 15 mètres (Ghinea et al., 2018).
Ainsi, dans les EV immersif proposé à l'aide de casque de RV plusieurs facteurs techniques entre en compte dans la mauvaise estimation des distances : la réduction du champ visuel, la masse du casque, le temps d'inertie, la sensation de porter un casque et le calibrage CVA/CVG. Il convient de noter que la perception des distances est également erronée dans les EV immersif proposé par d'autres technologies comme les systèmes CAVE.

Les systèmes CAVE sont des environnements de projection immersifs (l'EV est projeté sur plusieurs parois d'une pièce). Il a été observé que ce type de technologie de RV exerce une influence sur l'estimation de la distance. En effet, la distance égocentrique est généralement sous-estimée dans les EV présentées au sein des systèmes CAVE, et surtout pour ce qui concerne la distance verticale (Bruder et al., 2016 ; Ghinea et al., 2018 ; Marsh et al., 2014). Les objets seraient aussi perçus comme étant plus petits (Marsh et al., 2014).
Les systèmes CAVE seraient plus performants que les casques de RV pour l'estimation des distances égocentriques entre 2 et 3 mètres ainsi qu'entre 5 et 7 mètres (Ghinea et al., 2018). La possibilité de se mouvoir au sein de ces CAVE améliore la perception de la distance en permettant à l'utilisateur de changer de point de vue virtuel (Marsh et al., 2014). Dans les grands systèmes CAVE pouvant atteindre 10m d'espace de mouvement, l'estimation optimale de la distance égocentrique s'effectuerait lorsque l'utilisateur se trouve entre 6 et 7 mètres de l'écran (Bruder et al., 2016). Enfin, dans un système CAVE, les utilisateurs peuvent se servir des éléments de l'environnement physique pour améliorer leur estimation de la distance. Marsh et al. (2014) ont observé que les utilisateurs parviennent à une meilleure estimation de la distance lorsque les objets virtuels se situent à un des angles du système CAVE.

Quel que soit le système RV immersif employé, la perception des distances est erronée. Les casques de RV, même s'ils biaisent la perception des distances à cause de leurs caractéristiques (réduction du champ visuel, calibrage CVA/CVG…), seraient plus performants que les systèmes CAVE pour des distances supérieures à 7 mètres et inversement. Les systèmes CAVE, quant à eux, offrent notamment comme avantage de permettre à l'utilisateur de changer de point de vue en se déplaçant ainsi que de pouvoir se servir d'élément physique du système pour améliorer la perception des distances virtuelles.

3-2-4 Évaluation de la distance au sein de dispositifs non-immersifs

Des études ont également porté sur l'estimation de la distance au sein d'environnement non-immersifs c'est à dire un EV dans lequel « l'utilisateur se retrouve face à un simple écran et à certains dispositifs qui lui permettent d'interagir avec l'EV » (Chellali, 2009).
Alors que la distance exocentrique est généralement surestimée au sein d'environnements virtuels immersifs, une étude a démontré qu'une sous-estimation et une surestimation peuvent survenir de manière équivalente au sein des EV non immersif (Wartenberg et Wiborg, 2003). Cependant, selon une autre étude, la distance exocentrique était surestimée quel que soit le type de dispositif concerné, immersif ou non (Waller, 1999). Selon cette dernière étude, la surestimation de la distance entre deux objets est plus élevée avec un ordinateur de bureau qu'avec un casque de RV. De plus, Watenberg et Wilborg (2003) ont observé que le taux d'erreur augmente à mesure que la distance à estimer devient grande au sein d'EV, effet amplifié par le fait d'interagir avec un environnement non-immersif.
Ainsi, dans les EV non-immersif l'estimation de la distance est erronée. Cette erreur serait par ailleurs plus grande avec un environnement non-immersif comparé à un environnement immersif.

Une des explications possibles de cette mauvaise estimation des distances au sein d'EV non-immersifs pourrait être que les indices de profondeur non-picturaux (disparité binoculaire, accommodation et convergence) communiquent des informations de profondeur relatives à la distance entre l'utilisateur et l'écran et non pas entre l'utilisateur et les objets eux-mêmes (Stefanucci et al., 2015).
C'est notamment le cas du conflit entre l'accommodation et le point de convergence. Le point de convergence dépend des objets pouvant être présentés sur ou derrière l'écran tandis que l'accommodation s'effectue sur la surface de l'écran (Huckauf, 2005 ; Renner et al., 2013). Des informations contradictoires sont alors renvoyées biaisant l'estimation des distances.
De plus, l'absence d'information renvoyée par la vision binoculaire sur les éléments présentés à l'écran serait aussi un facteur diminuant la performance d'estimation de la distance dans les EV non-immersifs. Wainman et al. (2018) ont comparé l'effet de la stéréoscopie dans une tâche d'apprentissage de l'anatomie du pelvis féminin nécessitant une analyse spatiale sur modèle physique ou sur modèle virtuel. Durant la phase d'apprentissage, les participants ont eu soit l'œil non-dominant cacher (éliminant la vision binoculaire) soit une vision binoculaire. Lorsque

la vision binoculaire était disponible, le modèle physique a permis une meilleure performance que le modèle virtuel. Cependant, sans vision binoculaire, les performances sur le modèle physique ont diminué tandis qu'elles n'ont pas changé pour le modèle virtuel. De plus, il n'y avait pas de différence significative entre le modèle physique et le modèle virtuel dans la condition sans vision binoculaire. Ces résultats indiquent que la performance dans l'analyse spatiale d'éléments présentés à l'aide d'un ordinateur, est diminuée par l'absence d'information renvoyée par la vision binoculaire relative à ces éléments.

Ainsi, les informations contradictoires provenant de l'accommodation et du point de convergence ainsi que l'absence d'informations provenant de la disparité binoculaire sur les éléments présentés au sein de l'EV sont des facteurs diminuant l'estimation des distances au sein d'EV non-immersifs.

En l'absence d'informations pertinentes provenant des indices de profondeur non-picturaux pour évaluer la distance dans les EV non-immersifs, les études se sont centrées sur les autres facteurs permettant l'estimation de la distance au sein de ces environnements. Les recherches ont alors porté sur les indices de profondeur picturaux (informations relatives à la profondeur et à la distance pouvant être identifiées sur des scènes statiques) (Cutting et Vishton, 1995) et leurs possibles combinaisons au sein d'environnement non-immersifs. L'absence de plusieurs indices visuels picturaux peut donner lieu à une mauvaise interprétation de la distance. Par exemple, si un petit carré proche de l'observateur est présenté avec un grand carré éloigné, le petit carré semble plus éloigné que le grand carré, car seule la taille relative donne un indice sur la distance et la profondeur entre les deux objets. Cet effet peut être annulé en ajoutant des indices de perspective au sein de l'environnement concerné (figure 19) (DeLucia, 1991).

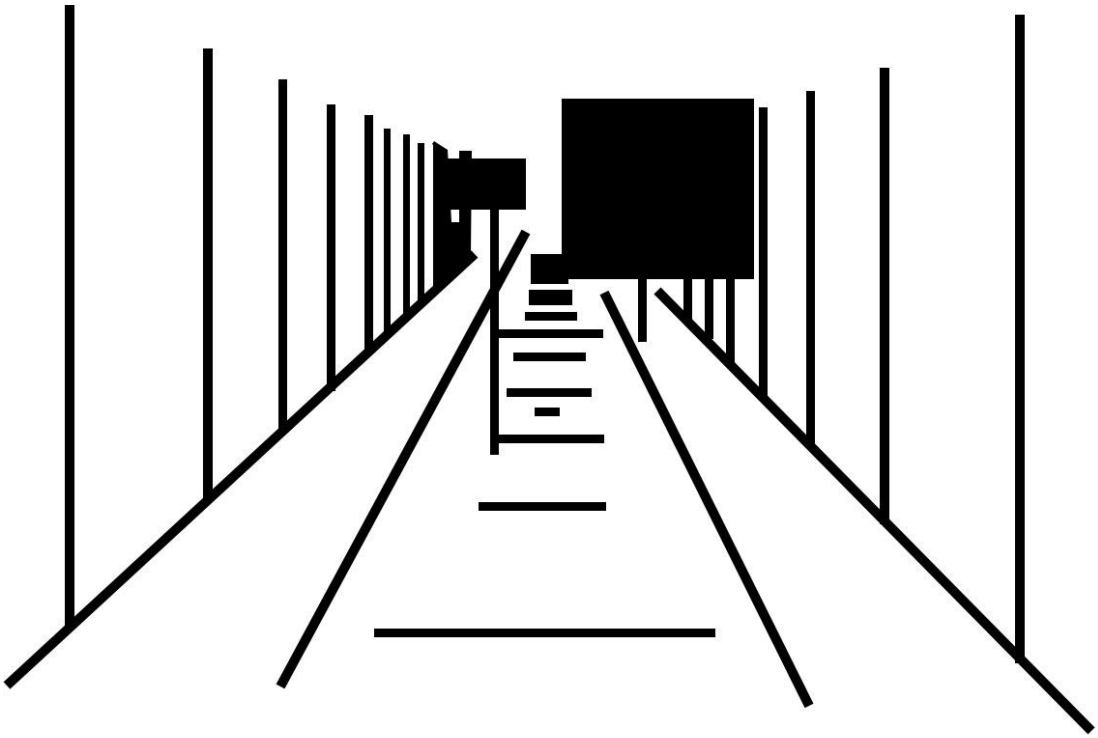


Figure 19 : Exemple d'indice de perspective améliorant l'estimation des distances d'après Delucia (1991).

Dans leur revue de la littérature portant sur l'estimation des distances au sein d'environnements non-immersifs, Macredie et Morar (2000) ont relevé que l'éclairage, la texture et la perspective linéaire étaient des facteurs ayant peu d'impact sur l'estimation de la distance, tandis que l'occlusion, la taille relative seraient les facteurs ayant le plus grand impact sur l'estimation de la distance. Néanmoins, les effets des différents indices de profondeur et de leur combinaison demeurent une question complexe.

Une première étude (Morar et al., 2002) a démontré que l'ajout de texture au sein des objets présentés améliorait la perception de la distance egocentrique seulement en présence de parallaxe du mouvement. De plus, cette même étude a établi que l'ajout d'indices de taille relative et de parallaxe du mouvement combinée, soit à des indices d'éclairage, soit à la texture des objets, améliorait l'estimation de la distance exocentrique. Une autre étude (Gerig et al., 2018) a comparé 3 compositions d'environnements dans le contexte d'une tâche d'atteinte de cible : un environnement minimal, où seule la taille de la cible à atteindre était connue, un environnement moyen où des indices de perspectives aériennes, de perspectives linéaires, de parallaxes du mouvement, d'occlusions, d'ombrages et de textures étaient ajoutés, et un environnement complet similaire à l'environnement moyen, comprenant en plus des objets

dont la taille relative était connue. Il n'y avait pas de différence significative quand a l'estimation de la distance pour atteindre la cible entre les environnements minimal et moyen ; cependant, l'environnement complet a entrainé une réduction du temps d'atteinte de la cible par rapport aux deux autres environnements. Les auteurs émettent l'hypothèse que l'ajout de l'objet connu a augmenté le sentiment de présence et/ou la motivation des participants, permettant de rendre les autres indices visuels utilisables améliorant ainsi la perception de la distance vis-à-vis de la cible à atteindre.

L'effet des différents indices de profondeurs picturaux et de leurs combinaisons sur l'amélioration de l'estimation des distances dans les EV non-immersifs reste une question ouverte. De plus, d'autre facteurs, notamment la position des objets dans l'environnement (Kjelldahl et Prime, 1995 ; Morar et al., 2002) ainsi que la nature de la tâche à effectuer ont un impact sur l'effet des différents indices visuels (Wanger et al., 1992). Kjelldahl et Prime (1995) ont étudié la perception de la distance entre deux objets dans une tâche de bissection (tâche ou l'utilisateur doit trouver le point central entre lui et un objet ou entre 2 objets) en faisant varier l'alignement des objets (horizontal ou en profondeur), la texture des objets et l'éclairage. Les auteurs ont observé que la perception de la distance était meilleure lorsque les objets étaient alignés horizontalement plutôt qu'en profondeur. De plus, l'éclairage avait un impact sur l'estimation de la distance, tandis que la texture n'avait aucun effet. Par ailleurs, Wanger et ses collaborateurs (1992) ont étudié les effets de l'ombrage, de la perspective linaire, de la parallaxe du mouvement et de la texture des objets sur la perception de la profondeur à l'aide de 3 types de tâches. Une première tâche de positionnement – où le sujet devait placer une sphère au milieu de deux sphères – nécessitant d'accéder à la position des objets dans l'espace. Seul l'ombrage et la perspective linéaire ont eu un effet positif sur cette tâche. Une deuxième tâche de rotation – où un cube devait être placé dans la même orientation qu'un autre cube – demandant donc de suivre l'orientation des objets dans l'espace. La parallaxe du mouvement et l'ombrage ont exercé une influence positive sur cette tâche, tandis que la perspective linéaire exerçait une influence négative. Enfin, une tâche d'échelonnage – où le participant devait faire correspondre la taille d'une sphère à celle d'une autre sphère – demandant à l'utilisateur d'avoir la position et la taille des objets. L'ombrage, la perspective linéaire, la parallaxe du mouvement et la texture des objets ont eu un effet positif sur cette tâche.

Ainsi, en fonction de la position des objets et de la tâche nécessitant une estimation de la distance ou de la taille des objets, certains indices de profondeur ont plus d'importance pour l'estimation des distances que d'autres.

Le type de matériel utilisé pour interagir avec l'EV impacte également l'estimation de la distance au sein d'environnements non-immersifs. Pour interagir avec les EV non-immersifs, un clavier et une souris sont couramment utilisés. Cependant, cette méthode ne serait pas optimale pour la perception de la distance. En effet, la manipulation d'objet en 3D s'effectue la plupart du temps à l'aide d'un curseur en 2D (traditionnellement une flèche pouvant se déplacer seulement en 2 dimensions) contrôlé grâce à une souris. Ce dispositif entrainerait un conflit entre l'espace en 2D du curseur et l'espace en 3D du contenu, en plus de ne donner aucun indice sur la profondeur des objets sélectionnés (Jauregui et al., 2012). Ainsi, des méthodes de curseur en 3D (utilisation d'une torche ou d'une main en 3D virtuelle) qui ont pu être utilisées, ont montré une amélioration de la perception de la profondeur. Néanmoins, elles diminueraient les performances de sélection par rapport à un curseur en 2D traditionnel (Jauregui et al., 2012). De plus, les commandes basées sur le clavier et la souris ne seraient pas non plus optimales. Kulik et al. (2009) ont comparé l'utilisation d'une interface bimanuelle (où la navigation du point de vue s'effectue à l'aide d'un joystick manipulé avec la main non dominante tandis que la sélection et la manipulation géométrique d'objets virtuels s'effectue à l'aide d'une souris maniée avec la main dominante) à une interface traditionnelle (clavier, souris) dans une tâche de navigation nécessitant une estimation des distances egocentriques et exocentriques. Les auteurs ont observé que l'interface bimanuelle entraine moins d'erreurs que l'interface traditionnelle.

L'estimation des distances au sein d'EV non-immersifs demeure un problème complexe. En l'absence d'informations provenant d'indices non picturaux sur la distance au sein d'EV non-immersifs, les éléments composant la scène virtuelle, leur position ainsi que les méthodes d'interaction sont les facteurs majeurs ayant une influence sur l'estimation des distances. Cependant, même en présence de ces éléments, les EV non-immersifs sont associés à une mauvaise perception de la distance. Des techniques de visualisation ont alors pu être mises en place pour améliorer la perception des formes, de la profondeur et de la distance d'objets complexes en 3D présentés sur un ordinateur de bureau.

Ces techniques ont pu être utilisées dans le domaine médical, notamment pour améliorer la perception spatiale de modèles angiographiques (Heinrich et al., 2021 ; Preim et al., 2016 ; Ropinski et al., 2006). Dans leur revue de la littérature portant sur les techniques améliorant la perception spatiale au sein de modèles en 3D dans le domaine médical, Preim et al. (2016) distinguent 2 types de techniques de visualisation : les techniques de visualisation basées sur la physique et les techniques de visualisation illustratives.

3-2-4-1 les techniques de visualisation basées sur la physique

Les techniques de visualisation basées sur la physique imitent les effets physiques induits par le transport de la lumière. Ainsi, les techniques d'ombrage qui utilisent les ombres induites par le transport de la lumière sont des facteurs revenant souvent dans la littérature comme ayant une importance dans la perception des distances au sein d'un environnement non-immersif (Díaz et al., 2017 ; Heinrich et al., 2021 ; Preim et al., 2016 ; Ropinski et al., 2006). Une étude a comparé plusieurs techniques d'ombrage pour la compréhension des distances d'un objet complexe et a montré que toutes les techniques d'ombrage améliorent la perception des distances au sein d'un objet virtuel (Diaz et al., 2017). D'autres techniques de visualisation basées sur la physique et améliorant la perception de la distance sont les techniques de perspectives aériennes (Heinrich et al., 2021 ; Preim et al., 2016 ; Ropinski et al., 2006). Elles consistent à créer, relativement aux éléments éloignés, un effet de brouillard qui augmente avec la distance (figure 20).

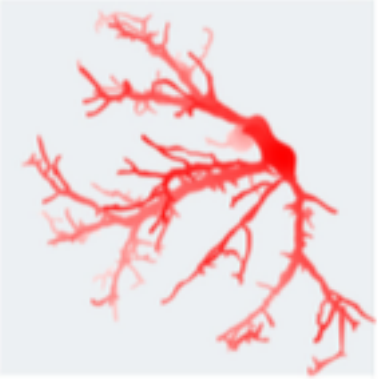


Figure 20 : Technique de perspective aérienne (Heinrich et al., 2021). Dans cette figure utilisée dans l'étude de Heinrich et al., (2021) les éléments plus profonds sont plus flous que les éléments proches.

3-2-4-2 les techniques de visualisation illustratives

Les techniques illustratives ne consistent pas, quant à elles, à imiter des effets physiques du monde réel, mais à accentuer des aspects importants pour la perception de la distance au sein d'objets en 3D et à en supprimer d'autres, afin de diriger l'attention de l'utilisateur (Preim et al., 2016). La technique illustrative la plus couramment utilisée pour la perception des profondeurs de modèles vasculaires en 3D est la technique de profondeur pseudo-chromatique (Lichtenberg et Lawonn, 2019). Cette technique consiste à donner un code couleur aux éléments d'un objet en fonction de leur distance (figure 21) ce qui induit une amélioration de la perception de la distance (Heinrich et al., 2021 ; Ropinski et al., 2006).

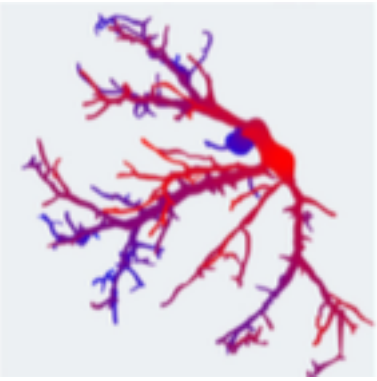


Figure 21 : Technique de profondeur pseudo-chromatique (Heinrich et al., 2021). Dans cette figure tirée de l'étude de Heinrich et al. (2021) un gradient de couleur allant du bleu pour les

éléments profonds au rouge pour les éléments moins profonds est utilisé pour coder la profondeur.

Une autre technique consiste à renforcer l'occlusion en accentuant les limites des éléments constituant les objets en 3D (figure 22). Son efficacité en matière d'amélioration de la perception de la distance et du volume a été démontrée (Preim et al., 2016 ; Ropinski et al., 2006). Cependant, l'épaisseur des éléments ne doit pas être affectée par ce renforcement des contours (Ropinski et al., 2006).

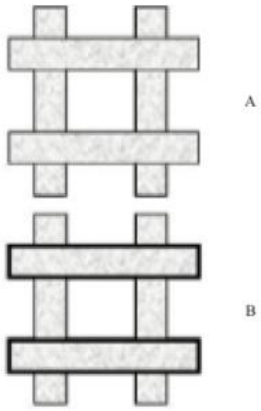


Figure 22 : Principes des techniques de renforcement de l'occlusion d'après Ropinski et al., 2006. Dans cette figure illustrant le principe des techniques de renforcement de l'occlusion les contours des éléments moins profonds de l'objet B sont renforcer tandis que les contours des éléments profonds restent identiques.

Pour ce qui est des novices, les techniques de profondeur pseudo-chromatique se sont avérées plus efficaces que les techniques de renforcement de l'occlusion en matière de perception de la distance au sein de modèles angiographiques en 3D ; en revanche, les deux techniques se sont avérées pareillement efficaces pour ce qui est des experts (Kersten-Oertel et al., 2014). Enfin, l'ajout de textures hachurées sur la surface des objets a induit une amélioration de la perception de la forme des objets, notamment quand l'ajout est associé à des techniques d'ombrages (Preim et al., 2016).

Ces différentes techniques (ombrage, perspectives aériennes, profondeur pseudo-chromatique et renforcement de l'occlusion) même si elles modifient la nature des objets se sont révélées

être efficace pour améliorer l'analyse spatiale d'objets 3D complexes présentés au sein d'un EV non-immersif.

Outre l'estimation des distances, qui, comme nous venons de le voir, semble moins performante lorsque qu'elle est réalisée dans un environnement virtuel, un autre paramètre important à prendre en compte et qui pourrait être en lien avec les difficultés rencontrées dans ce type d'environnement est la notion de charge mentale.

3.3 Réalité virtuelle, charge mentale et complexité de la tâche

La charge mentale est un facteur identifiécomme ayant un effet sur les performances au sein d'EV. Ce facteur fait référence aux ressources cognitives mobilisées à un moment donné pour mener à bien une tâche (Oviatt, 2006 ; Parasuraman, 2003). La charge mentale augmente avec la difficulté de la tâche alors que la performance diminue (Chanquoy et al., 2007 ; Jaquess et al., 2017). Différentes études ont alors analysé l'effet produit par la charge mentale sur l'exécution d'une tâche au sein d'un EV. Une première étude (Mubarrat et al., 2020) a comparé la réalisation d'une tâche demandant aux participants de déplacer des objets d'une position initiale à une position cible dans un EV avec feedback haptique et dans un environnement réel. Même si la charge mentale diminuait avec le nombre d'essais, elle était plus importante au sein de l'EV que dans l'environnement réel. Koumaditis et al. (2020) ont comparé la charge mentale dans la résolution d'un puzzle en 3D, virtuel ou physique. Une meilleure performance était observée en présence de l'objet physique. Par ailleurs, les auteurs ont trouvé que, dans les deux conditions, une hausse de la charge mentale était associée à une baisse de la performance. De plus, une autre étude a montré que la manipulation d'un objet physique entraine une meilleure performance que la manipulation d'un objet virtuel (Winther et al., 2020). Ces résultats confirment le fait que l'exécution d'une tâche manuelle au sein d'un EV augmente la charge mentale par rapport à l'exécution de cette même tâche avec des objets physiques.

Par ailleurs, la charge mentale a également été étudiée dans le domaine de l'anatomie, par le biais de la comparaison entre modèles physiques et modèles virtuels. Dans leur revue de la littérature portant sur l'utilisation de modèles physiques dans l'apprentissage de l'anatomie, Chan et Cheng (2011) ont trouvé que l'utilisation de modèles physiques diminue la charge mentale par rapport à d'autres modalités d'apprentissage, dont les modèles virtuels. Une étude (Wismer et al., 2018) a porté sur la charge mentale lors de l'apprentissage de l'anatomie du

cerveau à l'aide de la saturation en oxygène régional. Cette méthode mesure le taux d'oxygène dans le cerveau, et une hausse de ce niveau d'oxygène serait, selon les auteurs, associée à une hausse de la charge mentale (Peck et al., 2014). Les résultats ont alors montré une hausse de l'oxygène dans l'hémisphère droit lors de la condition en RV non présente dans la condition avec modèle physique, suggérant une hausse de la charge mentale. Une autre étude (Khot et al., 2013) a démontré que l'apprentissage de l'anatomie du système pelvien féminin était meilleure avec des modèles physiques qu'avec des modèles virtuels. Les auteurs expliquent alors que la différence de performance est due à une hausse de la charge mentale en RV par rapport aux modèles physiques.

Même si les tâches réalisées dans des EV entrainent une charge mentale plus grande que celles réalisées dans des environnements réels, les EV se sont révélés efficaces pour baisser la charge mentale d'une tâche réalisée dans un environnement réel après avoir été réalisée dans un EV (Barré et al., 2019 ; Sakib et al., 2020).

Les recherches se sont aussi concentrées sur les facteurs influençant la charge mentale au sein d'EV. Parmi les facteurs étudiés, le niveau d'immersion a des effets qui restent débattus. Plusieurs études menées sur le niveau de charge mentale en fonction de l'immersion dans des tâches de laparoscopie, ont démontré que la condition immersive entraîne une charge mentale plus faible que la condition non-immersive (Gómez-Gómez et al., 2015 ; Sakata et al., 2017; Zhang et al., 2021). Des résultats similaires ont été obtenus dans une tâche qui consistait à couper des fruits apparaissant à l'écran (Lum et al., 2018). Cependant, dans une tâche de calcul, la hausse du niveau d'immersion était associée à une hausse de la charge mentale (Jost et al., 2020). Riley et Kaber (1999), quant à eux, n'ont pas observé un quelconque effet du niveau d'immersion sur la charge mentale dans une tâche de navigation spatiale.

Un autre facteur revenant souvent dans la littérature est la complexité de la tâche. La complexité d'une tâche correspond à l'agrégation des caractéristiques de la tâche influant sur la performance (Liu et Li, 2012). Ainsi, le nombre d'éléments composant la tâche jouera un rôle dans sa complexité (Liu et Li, 2012 ; Wood, 1986), qui dépendra du nombre d'alternatives possibles permettant d'établir une prise de décision (Johnson et Bruce, 1998). De plus, selon March et Simon (1958), la complexité varie en fonction de la capacité de l'exécutant : les tâches sont plus ou moins complexes en fonction des capacités de l'individu qui les exécute. Enfin, Jacko et Ward, (1996) ont estimé que la complexité d'une tâche influait sur la charge mentale et sur la performance des participants. Dans les EV, la complexité de la tâche est aussi associée à la charge mentale. Plusieurs études ont montré que la charge mentale augmentait au fur et à

mesure que la complexité augmentait (Chao et al., 2017 ; Das et al., 2020 ; J. Li et Kim, 2021). Par exemple, Li et Kim (2021) ont évalué la charge mentale lors de l'exécution d'un puzzle simple et d'un puzzle complexe en RV. Les auteurs ont montré que la charge mentale était plus importante avec le puzzle complexe qu'avec le puzzle simple.
Enfin, les habiletés spatiales des utilisateurs dans des tâches nécessitant une analyse spatiale (Levinson et al., 2007) et le niveau d'expérience avec la RV (Moro et al., 2017) seraient des facteurs impactant la charge mentale au sein d'EV.

Malgré leur impact sur les ressources cognitives et une mauvaise estimation de la distance au sein des dispositifs immersifs et non immersifs, les technologies de RV sont utilisées dans de nombreux domaines médicaux, en y apportant des bénéfices réels.

3-4 Application de la réalité virtuelle dans les spécialités médicales

3-4-1 Utilisation de la réalité virtuelle en médecine

La médecine est un des principaux domaines d'application de la réalité virtuelle (L. Li et al., 2017), par exemple à des fins de diagnostic ou de planification thérapeutique. En effet, les progrès de l'imagerie médicale et de la puissance de calcul ont rendu les images simulées beaucoup plus réalistes et beaucoup plus rapides à générer, permettant ainsi la reconstruction d'entités anatomiques tridimensionnelles et offrant aux praticiens la possibilité d'interagir directement avec cette imagerie (L. Li et al., 2017). De nombreux domaines de la santé sont concernés. Dans le secteur médical, Pensieri e Pennacchini (2016) ont identifié 4 domaines d'application majeurs de la RV : la RV en tant qu'interface de communication, la psychothérapie, la planification chirurgicale et la formation médicale.

Dans le milieu médical, des applications de RV ont pu être développées en tant qu'outils de communication au sein du personnel médical. Le service digital « Visible patient » développé par Soler et al. (2014) est un service permettant la modélisation en 3D de l'anatomie d'un patient à partir de l'imagerie médicale et incluant différents outils de visualisation et de planning chirurgical à même d'optimiser les actions de tous les acteurs impliqués. Agus et al. (2007) ont développé une interface de communication combinant des informations de données médicales en 2D (texte, imagerie médicale en 2D) avec une application de visualisation et de

manipulation de reconstruction anatomique en 3D de chaque patient permettant de faciliter les discussions diagnostiques et la planification préopératoire des anévrismes de l'aorte abdominale.

Les applications de RV ont aussi pu être utilisées comme outils de prévention ou pour fournir des informations sur les soins de santé afin d'améliorer la communication entre patients et cliniciens. Par exemple, un EV intégrant des quizz, des jeux, des ressources web et des séminaires en direct a pu être développé à des fins de prévention en matière de santé sexuelle (Kamel Boulos et Toth-Cohen, 2009).

La RV est également utilisée en raison de sa capacité à induire des réponses émotionnelles et à réduire la distinction entre la réalité virtuelle et la réalité grâce au sentiment de présence dans de nombreuses psychothérapies (Pensieri et Pennacchini, 2016). Les techniques utilisant l'imagerie motrice progressive et la RV peuvent aider à gérer la peur de bouger ou à réaliser un mouvement qui serait normalement trop douloureux (Baldominos et al., 2015). De plus, dans le cas de la douleur du membre fantôme (la sensation, très souvent douloureuse, que le membre manquant est toujours présent), la présentation du membre manquant à l'aide de la RV immersive permet la réduction de la douleur, une meilleure acceptation de la perte du membre et, ainsi, une meilleure adaptation du cerveau à cette perte (Collins et al., 2018). La RV s'est aussi montrée efficace dans la récupération motrice et dans l'amélioration de la démarche et de l'équilibre chez des personnes ayant subi un accident vasculaire cérébral ou atteintes de la maladie de Parkinson (Amirthalingam et al., 2021). Le traitement de troubles anxieux tels que les phobies s'est révélé efficace à l'aide de la RV (A. Li et al., 2011), avec la création d'un espace interactif dans lequel le patient est confronté à sa situation ou à son objet phobique.

Par ailleurs, dans le domaine de la chirurgie, une télécommande permet d'agir sur un modèle virtuel au moyen d'un robot virtuel : les actions sont d'abord réalisées dans l'environnement virtuel avant d'être transmises à l'opérateur réalisant l'opération, permettant ainsi de tester la procédure avant de la réaliser sur le patient (Satava et Jones, 1997).

Les modèles anatomiques en 3D offerts par la RV ont aussi pu être utilisés à des fins de diagnostic préopératoire et de planification chirurgicale. Les parties anatomiques pertinentes pour la chirurgie peuvent être identifiées sur des coupes en 2D obtenues grâce à l'imagerie médicale. Même si ces coupes contiennent toutes les informations nécessaires, les chirurgiens peuvent avoir du mal à percevoir les relations entre les structures avant l'intervention (Soler et

al., 2014). Les modèles en 3D ont alors montré des avantages par rapport à la visualisation standard de coupe (Agus et al., 2007 ; R. A. Kockro et al., 2016 ; Lamadé et al., 2000 ; Rousian et al., 2011 ; Soler et al., 2014). Dans le domaine de la chirurgie hépatique, Lamadé et al. (2000) ont montré que l'utilisation de modèles en 3D améliore la compréhension des relations tridimensionnelles de la tumeur avec les arbres vasculaires intra-hépatiques. Les auteurs ont comparé l'évaluation par des cliniciens de l'appartenance de la tumeur à un segment du foie et la proposition d'opérer sur des coupes ou des modèles en 3D. Les modèles en 3D ont permis une amélioration significative de la localisation de la tumeur et de la proposition de résection. Soler et al. (2014) ont testé, de 2009 à 2013, chez 719 cas cliniques, l'utilisation de la modélisation et de la visualisation en 3D du foie en tant que moyens facilitant l'estimation du volume du foie et de la tumeur à des fins de planification chirurgicale. L'application de RV a permis une représentation précise et fiable du volume du foie et du volume tumoral, ouvrant ainsi de bonnes perspectives pour le diagnostic et la mise en place de procédures chirurgicales. La chirurgie de l'anévrisme nécessite elle aussi une compréhension spatiale précise de l'anatomie vasculaire et de son environnement. Une application de RV consistant en la visualisation d'arbres vasculaires en 3D à l'aide d'un dispositif stéréoscopique a été testée pour la planification préopératoire, de 2006 à 2015, chez 105 patients (R. A. Kockro et al., 2016). Cette application a permis une hausse de la compréhension spatiale de l'anatomie vasculaire individuelle par le chirurgien ainsi qu'une bonne planification préopératoire avec une anticipation des difficultés et complications potentielles. De plus, la planification à l'aide du dispositif de RV a permis de très bons résultats cliniques avec des taux de réussite des opérations équivalents à ceux des cohortes de références. Un système de RV s'est aussi montré performant pour la visualisation et la quantification de volumes tridimensionnels échographiques (Rousian et al., 2011).

Ainsi, les modèles anatomiques visualisés à l'aide de la RV se sont montrés efficaces dans l'amélioration de la compréhension spatiale dans un but de diagnostic préopératoire. Cependant, les reconstructions en 3D ne sont pas encore utilisées de manière routinière dans les hôpitaux. D'une part, les médecins sont traditionnellement formés à extraire les informations nécessaires à partir d'images de coupe en 2D, et d'autre part, certains dispositifs de RV stéréoscopique peuvent entrainer de la fatigue et de l'inconfort visuels (Tourancheau et al., 2012).

Il existe également diverses applications dédiées à l'apprentissage et à la formation du personnel médical basées sur les simulateurs chirurgicaux et l'exploration logicielle du corps humain (Ackerman, 1998).

En effet, la visualisation en 3D de différentes parties du corps humain sous différents points de vue permet une meilleure compréhension des principes physiologiques et anatomiques ainsi que des relations entre les structures pouvant être difficilement étudiées, y compris par la dissection de cadavres (Pensieri et Pennacchini, 2016). Ainsi, les technologies de réalités virtuelles ont pu être utilisées, seules ou en complément des méthodes traditionnelles, pour favoriser une meilleure connaissance de l'anatomie. Elles ont été perçues par les apprenants comme plus engageantes et plus motivantes, propres à augmenter le plaisir dans l'apprentissage (Azer et Azer, 2016). De plus, elle se sont avérées très utiles pour améliorer la compréhension spatiale des relations anatomiques par les étudiants (Azer et Azer, 2016 ; Birbara et al., 2020 ; R. Kockro et al., 2015 ; Sakellariou et al., 2009). Dans un premier travail, Sakellariou et al. (2009) ont étudié les avantages de la RV pour l'apprentissages anatomique du canal inguinal par des étudiants en médecine. Le canal inguinal est une partie du corps humain reconnue comme étant extrêmement complexe à appréhender par les étudiants, notamment pour ce qui est des relations spatiales entre ses différents éléments. Cette étude a alors montré que les étudiants avaient du mal à se construire une représentation en 3D du canal inguinal à partir d'images de coupe en 2D et que l'apprentissage en RV permet une meilleure compréhension spatiale des interrelations entre les éléments du canal inguinal, par rapport aux méthodes traditionnelles. Dans une autre étude, Krockro et al. (2015) ont montré des résultats d'apprentissage similaires entre un système de RV stéréoscopique et un diaporama pour l'apprentissage de l'anatomie du troisième ventricule par des étudiants en deuxième année de médecine. Cependant, le système de RV s'est avéré meilleur que le diaporama pour ce qui est de l'efficacité, du plaisir et de la compréhension spatiale du troisième ventricule. Le niveau d'immersion aurait aussi un impact sur l'apprentissage de l'anatomie. Birbara et ses collaborateurs (2020) ont comparé l'apprentissage en RV de l'anatomie du crâne avec un système de RV stéréoscopique (immersif) et avec un ordinateur de bureau (non-immersif) chez des étudiants et des enseignants. Même si les étudiants ont perçu le dispositif immersif comme étant plus utile, celui-ci était associé à un plus grand inconfort et à une plus grande charge mentale que l'ordinateur de bureau. De plus, les enseignants ont, quant à eux, perçu le dispositif non-immersif comme étant plus utile.

Néanmoins, l'utilisation de la RV seule pour l'apprentissage de l'anatomie demeure débattue. En effet, certaines études ont montré que l'apprentissage en RV serait moins efficace que l'utilisation de modèles physiques (Khot et al., 2013 ; Preece et al., 2013) ou la dissection de cadavres (Codd et Choudhury, 2011 ; Petersson et al., 2009), notamment à cause du manque d'interaction tactile (Petersson et al., 2009). L'utilisation de la dissection de cadavres pour l'apprentissage de l'anatomie par les étudiants en médecine étant en baisse à cause du coût et de l'expertise nécessaires (Gartner, 2003 ; Korf et al., 2008), la RV semble être une méthode efficace pour compléter l'apprentissage de l'anatomie, notamment au niveau de la compréhension spatiale, importante pour la pratique clinique (Sakellariou et al., 2009).

Les simulateurs en RV ont aussi montré leur efficacité dans le domaine médical. Par exemple, un simulateur en RV avec retour haptique pour l'entrainement à l'élimination de caries a montré son efficacité chez des novices et chez des étudiants en prosthodontie (Eve et al., 2014).

Le tableau 1 présente une synthèse des différents domaines d'application de la réalité virtuelle en médecine et de ses apports en fonction du domaine concerné.

Domaine d'application	Exemple d'utilisation	Apports
Communication	Interface de communication joignant différentes données dont des modélisations en 3D des patients (Agus et al., 2007 ; Soler et al., 2014)	Facilitation des discussions diagnostiques et de la planification préopératoire
	Application de prévention (Kamel Boulos & Toth-Cohen, 2009)	Amélioration des connaissances médicales des patients
Psychothérapie	Traitement de la douleur du membre fantôme (Collins et al., 2018)	• Acceptation de la perte du membre • Réduction de la douleur
	Rééducation post-AVC ou pour les patients atteints de la maladie de Parkinson (Amirthalingam et al., 2021)	• Récupération motrice • Amélioration de la démarche et de l'équilibre
	Traitement des phobies (Li et al., 2011)	• Réduction de l'anxiété face à l'objet de la phobie ou la situation phobique
Chirurgie	Simulation de la chirurgie sur un modèle virtuel avec un robot virtuel (Saltava & Jones, 1997)	Tester la procédure avant la réalisation sur un patient
	Utilisation de reconstructions en 3D de l'anatomie du patient à des fins de diagnostic et de planification préopératoire (Agus et al., 2007 ; R. A. Kockro et al., 2016 ; Lamadé et al., 2000 ; Rousian et al., 2011 ; Soler et al., 2014)	• Amélioration de la compréhension spatiale de l'anatomie du patient par le chirurgien • Planification préopératoire efficiente • Anticipation des difficultés et des complications potentielles
Formation	Apprentissage de l'anatomie avec la visualisation en 3D de différentes parties du corps (Azer & Azer, 2016 ; Birbara et al., 2020 ; R. Kockro et al., 2015 ; Sakellariou et al., 2009)	• Hausse de l'engagement, de la motivation et du plaisir d'apprentissage • Amélioration de la compréhension spatiale des structures anatomiques
	Entrainement aux opérations chirurgicales (Ackerman, 1998 ; Eve et al., 2014)	• Permet aux apprenants de faire l'expérience de leur choix et de ses effets dans un environnement sans risque • Apprentissage efficient

Tableau 1 : Utilisation de la RV en médecine.

3-4-2 Réalité virtuelle et orthodontie

L'utilisation de la RV dans les champs de l'orthopédie dento-faciale, de l'orthodontie et de la chirurgie orthognathique est un reflet de ses différentes applications en médecine – des applications à visée tant éducative que clinique. En effet, la réalité virtuelle est amenée à jouer

un rôle majeur dans la formation initiale et continue des praticiens. L'acquisition en 3D des différents tissus de la région de la tête et du cou fournit une plate-forme réaliste pour l'entraînement en orthopédie faciale et dento-faciale (A. Ayoub et Pulijala, 2019). Par exemple, une nouvelle approche de réalité virtuelle a été introduite et validée pour l'étude céphalométrique sur téléradiographie de profil assistée (permettant au praticien d'observer tout ce qu'il ne peut pas observer directement par l'observation de la bouche du patient) (Medellín-Castillo et al., 2016). Vingt et un orthodontistes ont effectué une série d'études de cas en utilisant l'analyse céphalométrique numérique activée par le sens haptique. Ils ont montré qu'en fournissant un sens du toucher, les erreurs dans l'analyse céphalométrique ont été réduites et le repérage est devenu plus faisable et plus intuitif. Dans un autre exemple, Voxel Man Simulator (Wu et al., 2014) a été utilisé dans la procédure d'apicoectomie virtuelle (consistant à enlever la racine d'une dent et à la remplacer par une petite obturation) et a révélé que, sur 53 étudiants en médecine dentaire qui ont entrepris une apicoectomie virtuelle, 51 avaient une appréciation positive de l'impact de la simulation virtuelle en tant que modalité supplémentaire dans l'enseignement dentaire.

Les applications cliniques sont également nombreuses. Elles peuvent concerner la planification virtuelle (Kwon et al., 2018). Mais elles sont surtout utilisées à des fins de diagnostic. En effet, différentes techniques ont été développées pour capturer des données sur les tissus durs et mous de la sphère maxillo-faciale afin de produire des modèles virtuels en 3D pour l'analyse et la planification chirurgicales. Ces techniques ont permis de pallier les inconvénients des photographies et radiographies en 2D. Quatre principaux types de systèmes d'imagerie en 3D ont été utilisés pour capturer les structures dentaires et oro-faciales, notamment la tomographie volumique numérisée par faisceau conique (Cone Beam Computed Tomography ou CBCT)[2], ou le scanner à lumière structurée[3] et la stéréophotogrammétrie[4] (A. F. Ayoub et al., 2007).

Parmi ces applications, une des plus significatives étant donnés son impact et le développement qu'elle connaît, est le modèle numérique tridimensionnel des arcades dentaires. En effet, les

[2] Reconstructions 2D et/ou 3D réalisées à partir de données récupérées par un générateur de rayon X émettant un faisceau de forme conique (Benjelloun et al., 2012)

[3] Reconstructions 3D élaborées à partir de données récupérées par une lumière dont l'intensité est spatialement modulée (par exemple la lumière provenant d'une source lumineuse devant laquelle on a placé une grille) (Bieri et al., 2003)

[4] Reconstructions 3D élaborées par la mise en correspondance de deux images d'un même objet prise de deux position différentes (Nguyen, 2018)

modèles des arcades dentaires revêtent une grande importance pour les orthodontistes (Marcel, 2001). Ils sont un élément essentiel du dossier de diagnostic et sont utilisés pour documenter l'état d'origine, planifier le traitement et mesurer ses effets.

Jusqu'à récemment, les moulages en plâtre issus d'une empreinte des arcades dentaires (figure 1.a des modèles en plâtre) étaient le seul moyen de réaliser des modèles tridimensionnels pour représenter avec précision les arcades dentaires et les malocclusions[1]. Les modèles numériques en 3D sont récemment devenus une alternative dématérialisée (Kravitz et al., 2014). Ces modèles sont des enregistrements électroniques des dents et de l'occlusion d'un patient à partir d'une caméra optique intra-orale (figure 23, modèles numériques en 3D). L'empreinte optique des arcades dentaires est traitée par un logiciel qui dessine une projection tridimensionnelle, obtenu à partir de fichiers STL (STL pour Stereo-Lithography[5]) sur l'écran de l'ordinateur. Les fonctionnalités du logiciel permettent à l'observateur de changer de vue et de faire pivoter ou incliner l'image des modèles, ensemble ou individuellement. Le logiciel permet également de mesurer et de stocker les circonférences de l'arcade, le diamètre des dents etc.

Les modèles numériques offrent les avantages du stockage électronique et de l'accès à partir d'un emplacement distant, réduisant les risques de perte ou d'endommagement des modèles. Cette technologie peut être incluse dans le champ de la RV non immersive et présente deux caractéristiques essentielles :

- Le praticien peut manipuler de façon virtuelle les empreintes à travers l'écran. Il existe donc une interaction en temps réel qui permet à l'utilisateur de changer les propriétés de l'environnement avec lequel il interagit (Ellis, 1994 ; Lanier 1988 ; Ouramdane et al., 2009).

- À travers le diagnostic et la projection thérapeutique qu'effectue le praticien, une activité cognitive se déploie dans l'espace créé numériquement par les modèles en 3D (Burdea et Coiffet, 1993).

Le passage de modèles en plâtre à des modèles numériques constitue une « révolution » dans la pratique professionnelle des orthodontistes. Cependant cette transition numérique a évidemment comme corolaire, des changements importants en termes de diagnostic et de suivis thérapeutique. Professionnellement les diagnostics reposent, en très grande partie, sur de l'estimation de distance (fonctionnement holistique de l'expert). De ce fait, ces estimations qui étaient réalisées à partir de modèles en plâtre sont aujourd'hui réalisées à partir de modèles

[5] Le format de fichier STL est le format de transmission de données standard de l'industrie du prototypage rapide. Un fichier STL stocke l'information de la surface brute de modèle 3D avec de petits triangles.

numériques. Or, très peu d'études se sont intéressées aux éventuelles différences entre les estimations réalisées dans ces 2 situations. Mieux comprendre les différences entre 2 situations constitue pourtant un enjeu majeur pour l'avenir de cette spécialité de santé, tant sur les plans de la formation, du diagnostic que du traitement et du suivi thérapeutique.

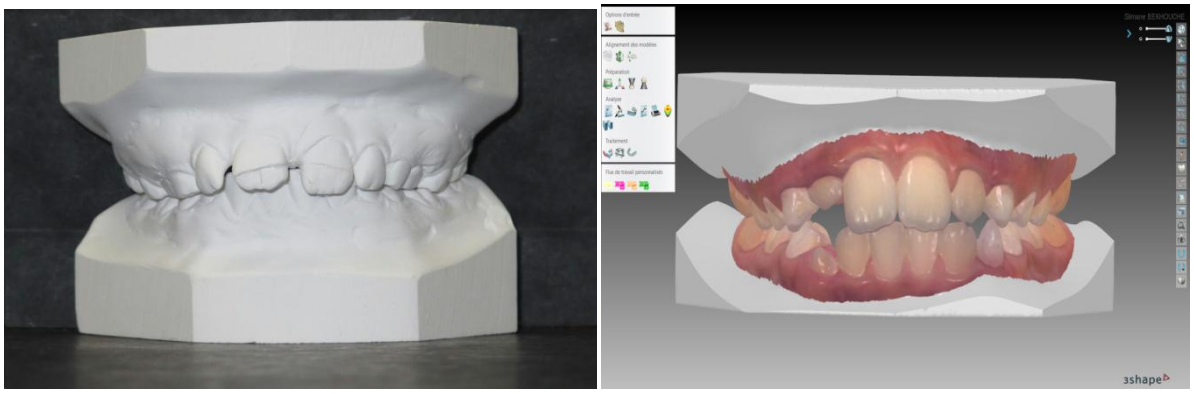


Figure 23 : Modèle d'étude orthodontique en plâtre et en 3D numérique.

3-5 Conclusion

Ainsi nous pouvons conclure, qu'au vu des nombreuses données de la littérature spécialisée, la perception des environnements virtuels est altérée par rapport à celle des environnements réels, alors que les technologies de RV sont utilisées à profit dans de nombreux domaines médicaux. Par exemple, une des problématiques observées est liée au fait que la manipulation utilise un dispositif avec 2 degrés de liberté, à la manière d'une souris d'ordinateur, pour interagir avec un environnement 6 degrés de liberté Ce gouffre séparant le monde réel du monde virtuel limite notre capacité à tirer le meilleur parti des informations dont nous disposons. Réussir une stratégie d'augmentation nécessite la connexion de l'intuition de l'expert aux flux de travail numérique.

Dans les champs de l'orthopédie dento-faciale et de l'orthodontie, il est fondamental que le praticien soit à même d'avoir une perception des modèles virtuels aussi proche que possible de la réalité clinique. Ce changement d'interface pose la question de l'interprétation de relations intra- et interarcades opérée par l'orthodontiste à travers un écran d'ordinateur et a partir desquelles il va prendre des décisions thérapeutiques comme celles des extractions ou non de prémolaires pour l'alignement de la denture. Les travaux menés dans le cadre de cette thèse ont pour objectif d'étudier la perception que le praticien peut avoir à travers des modèles

virtuels et de mener une réflexion sur l'optimisation de l'interface praticien-nouvelles technologies en orthopédie dento-faciale. Ces travaux serviront de base de réflexion pour la conception d'une interface plus immersive, gage d'une meilleure connexion entre les nouvelles technologies et l'orthodontiste.

3-6 Essentiel à retenir

Les technologies de réalité virtuelle permettent d'accéder et d'interagir avec des représentations spatiales en 3D nommées « environnement virtuel ». Des caractéristiques de la réalité virtuelle telles que l'immersion et le sentiment de présence vont affecter la perception de ces environnements virtuels.

La réalité virtuelle a pu être utilisée dans de diverses applications médicales : comme interface de communication, comme outil de formation, comme outil de thérapie ou pour la simulation d'intervention chirurgicale. Dans le cas de la planification préopératoire ou de l'apprentissage, l'utilisation de la RV a montré une hausse de la compréhension spatiale des structures anatomiques de la part des cliniciens et des étudiants par rapport aux méthode traditionnelles.

Pour la perception de la distance, le cerveau utilise des indices visuels de profondeur pour identifier les relations spatiales entre les objets et l'environnement.

La perception de la distance au sein d'un environnement virtuel diffère de celle opérée dans le monde réel. Alors que la distance égocentrique est généralement sous-estimée, la distance exocentrique est généralement surestimée. De plus, les objets sont perçus comme plus petits.

La mauvaise estimation de la distance au sein des environnements virtuels dépendrait du type de technologie utilisée, de la composition de l'environnement, de facteurs individuels et de la méthodes utilisée pour évaluer la distance.

Dans les environnements non-immersifs, soit les environnements virtuels présentés à l'aide d'un ordinateur, le type de moniteur ainsi que les méthodes d'interaction vont affecter l'estimation de la distance. Les indices visuels de profondeur y jouent un rôle majeur dans cette estimation et leurs effets dépendent de leurs positions au sein de l'environnement et de la

tâche visuelle à accomplir. L'effet de ces indices et de leurs combinaisons demeure néanmoins une question complexe.

De nombreuses techniques et outils auxiliaires ont pu être développés pour améliorer l'estimation de la distance au sein des environnements virtuels non-immersifs.

La charge mentale est un facteur identifié par la littérature scientifique comme ayant un effet sur les performances aussi bien dans des environnements virtuels que réels.

La médecine est un des principaux domaines d'application de la réalité virtuelle avec quatre domaines majeurs d'application : la RV en tant qu'interface de communication, la psychothérapie, la planification chirurgicale et la formation médicale.

L'utilisation de la RV dans les champs de l'orthopédie dento-faciale, de l'orthodontie et de la chirurgie orthognathique est un reflet de ses différentes applications en médecine. Parmi ces applications, une des plus significatives, étant donnés son impact et le développement qu'elle connaît, est le modèle numérique tridimensionnel des arcades dentaires.

Dans les champs de l'orthopédie dento-faciale et de l'orthodontie, il est fondamental que le praticien soit à même d'avoir une perception des modèles virtuels, notamment en termes d'estimation des distances, aussi proche que possible de la réalité clinique.

4 : Estimation des distance en 3D par l'orthodontiste utilisant des modèles numériques

Cette étude intitulée Estimation of Distances in 3D by Orthodontists Using Digital Models a été publiée dans Applied Sciences. Le manuscrit est intégré en Annexe (Annexe1)

Référence bibliographique de l'article :

Makaremi, M., & N'Kaoua, B. (2021). Estimation of Distances in 3D by Orthodontists Using Digital Models. *Applied Sciences*, *11*(18), 8285.

4-1 Objectif de l'étude

Compte tenu du caractère récent de l'apparition de l'orthodontie numérique, peu d'études se sont penchées sur l'analyse de l'interface praticien-nouvelles technologies numériques. Parmi ces travaux l'étude pilote de (Dunbar et al., 2014) réalisée en 2014 a analysé l'influence de la perception des modèles digitaux sur l'établissement du plan de traitement orthodontique.

Les auteurs ont cherché à déterminer si la planification du traitement orthodontique était reproductible quand il était réalisé avec des données numériques en comparaison avec un examen clinique ou ayant recours à des données standard sur papier.

Quatre orthodontistes non familiers avec les modèles en 3D ont pris part à l'étude, menée sur 27 patients. Il s'agissait d'établir un plan de traitement selon les options définies : absence de traitement / traitement orthodontique / traitement chirurgical.

Le premier mois, les orthodontistes devaient choisir une option de traitement des 27 patients au moyen d'un examen clinique seul.

Le deuxième mois, ils devaient définir une option de traitement en utilisant des données diagnostiques papier.

Enfin, le troisième mois, ils devaient se prononcer à l'aide de données numériques.

Les résultats des analyses statistiques ont révélé que :

- La reproductibilité inter-observateurs du choix du plan de traitement était plus faible pour le modèle numérique ;
- La reproductibilité intra-observateurs variait de bonne à médiocre

Cette étude démontre l'importance du format des informations de diagnostic dans la planification du traitement. De plus, les avis entre les observateurs divergent encore plus

lorsqu'il s'agit d'établir une thérapeutique à partir de données numériques.

Avec l'avènement des empreintes 3D, il paraît légitime de s'interroger sur l'interprétation des données numériques par l'orthodontiste : avons-nous une estimation correcte des dimensions réelles d'un objet lorsque nous nous tenons face à un écran ?
Des études ont montré la précision des estimations des mesures de distances sur des modèles en plâtre (song et al., 2015). En revanche, de nombreux praticiens témoignent des difficultés rencontrées lorsqu'ils analysent des modèles virtuels numériques. Et comme indiqué dans notre revue de la littérature (Chapitre 3) les estimations de distances réalisées dans des environnements virtuels, y compris dans des environnements non-immersifs, sont moins fiables que celles réalisées sur des objets réels. Dans ce contexte, notre étude a pour objectif de déterminer la précision des estimations de distances réalisées sur des modèles numériques, par des orthodontistes.

4-2 Matériel et méthode

4-2-1 Matériel

Le protocole expérimental a consisté à recueillir des estimations de mesures effectuées par des orthodontistes, puis à valider la fiabilité de ces mesures
Les mesures ont été estimées par 31 orthodontistes dont les caractéristiques étaient les suivantes : tous les praticiens ont commencé leur spécialisation en orthopédie dento-faciale et orthodontie au sein des universités françaises. La répartition de l'échantillon par sexe était de 19 femmes et 12 hommes, l'âge moyen des participants était de 37 ans (±3,7). Le nombre d'années d'exercice de la profession a été réparti comme suit : 0 à 5 ans d'exercice professionnel (42 %), 5 à 10 ans (32 %) et plus de 10 ans (26 %). L'objectif de cette recherche été expliqué aux praticiens, qui ont donné leur accord. Les participants ont été recrutés par courriel envoyé à tous les cabinets d'orthodontie de la région. La participation était volontaire.

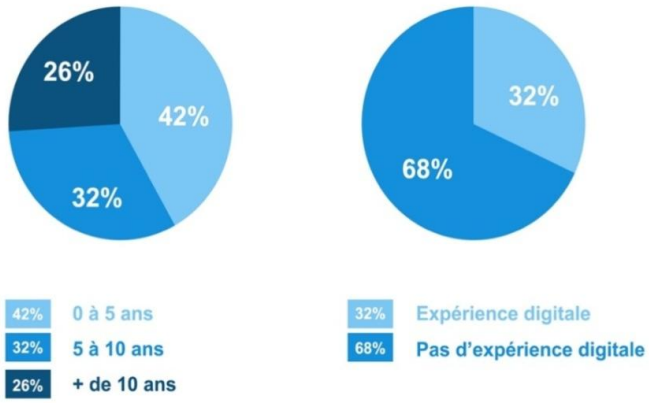


Figure 24 : Identification de la population participant à l'étude. Le schéma de gauche indique le nombre d'années d'expérience des praticiens, et le schéma de droite l'absence ou la présence d'expérience digitale.

Les estimations des distances ont été réalisées sur cinq modèles numériques 3D (obtenus à l'aide du logiciel Insignia®), issus de 5 cas cliniques (figure 24). Le choix de ces 5 cas était basé sur l'objectif de permettre aux praticiens de faire des estimations sur un panel varié de malocclusions orthodontiques comprenant tous les types de dysmorphies dento-maxillaires[6] dans les trois dimensions de l'espace. Les caractéristiques des différentes malocclusions sont rapportées dans le tableau 2.

Cas n°1	Classe I avec un encombrement modéré de l'arc
Cas n°2	Classe II avec supraclusion prononcée des incisives
Cas n°3	Classe II avec un surplomb significativement augmenté
Cas n°4	Classe II avec infraclusion des incisives
Cas n°5	Classe II avec une malposition des incisives mandibulaires

Tableau 2 : Présentation des 5 cas cliniques utilisés dans notre étude. Arc : ensemble anatomique formé par la disposition architectonique des dents et de leurs alvéoles.

[6] Différence entre la somme des diamètres mésio-distaux des dents d'une arcade, dite « espace nécessaire » et la longueur de cette arcade, dite « espace disponible ».

Supraclusion : excès de recouvrement des incisives et éventuellement des canines mandibulaires par leurs antagonistes maxillaires. Infraclusion : maloclusion localisée, caractérisée par l'insuffisance de contact dentaire due à l'éloignement vertical des dents antagonistes.

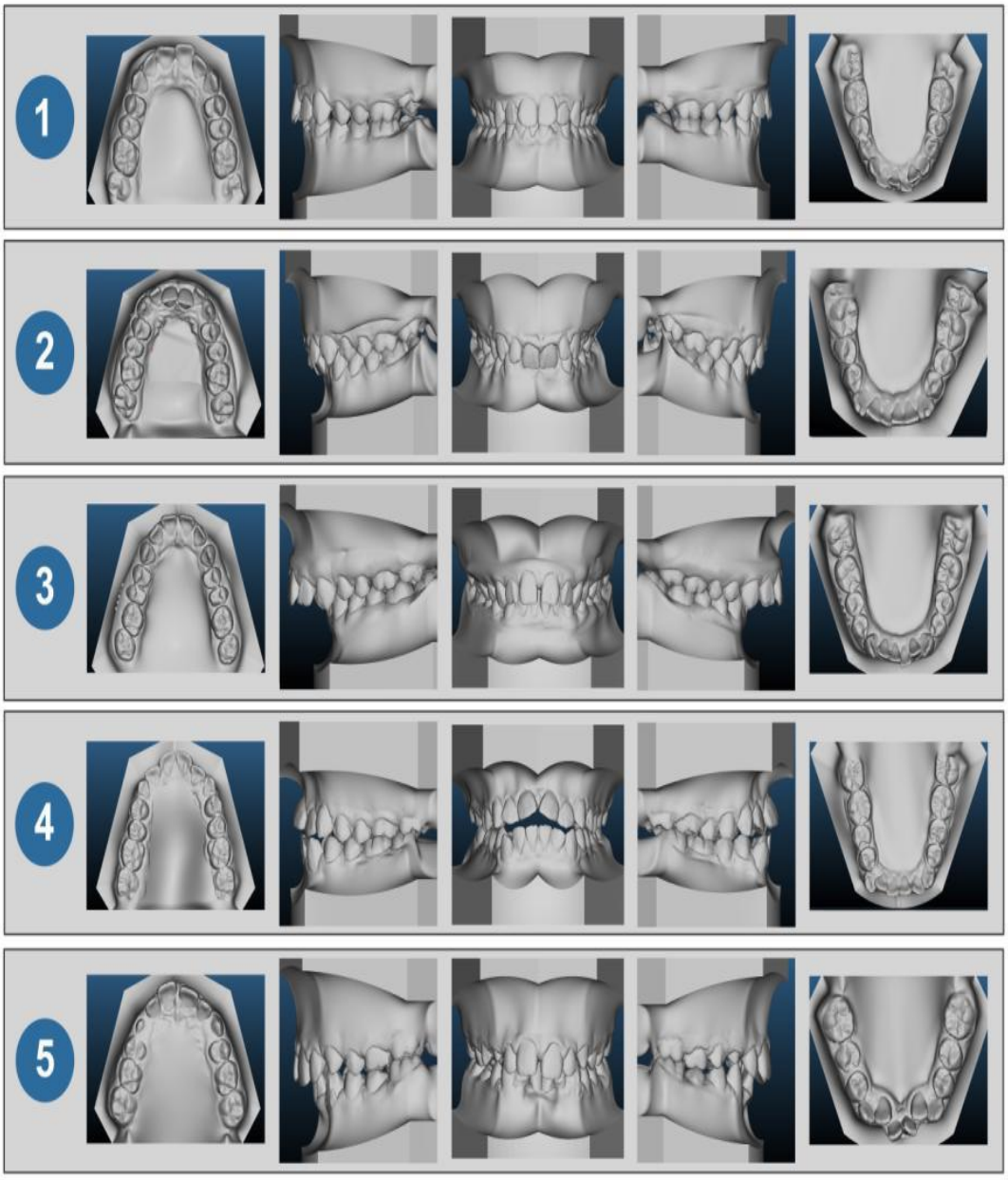


Figure 25 : Les 5 cas cliniques utilisés pour l'estimation de la distance.

Les modèles numériques ont été affichés sur le même ordinateur, par le même opérateur et dans les mêmes conditions pour chacun des 31 praticiens. Les participants avaient la possibilité

de changer l'orientation de la vue du modèle. Le recueil des estimations des distances s'est effectué de la manière suivante : après avoir observé les modèles virtuels de chaque cas clinique sur l'écran, le praticien donne verbalement ses estimations de distance pour chacune des 7 variables (7 distances à estimer) à l'opérateur qui les inscrit dans un tableau. Ces estimations ont été réalisées sur les cinq modèles virtuels mentionnés (5 * 7 = 35 estimations par orthodontiste), sans possibilité de faire des mesures directes (il ne s'agissait que d'estimations). Les variables sélectionnées étaient similaires à celles de l'étude de (Song et al., 2015).

Ces variables (distances à estimer) étaient les suivantes (figure 25) :

- L'encombrement de l'arcade mandibulaire : l'espace nécessaire au bon alignement des dents dans l'arcade (reconstitution volumétrique, mesure intra-arcade) ;
- Courbe de Spee : la courbe formée par la projection dans un plan sagittal des cuspides[7] vestibulaires des dents mandibulaires (dimension verticale, mesure intra-arcade) ;
- Symétrie antéro-postérieure des premières molaires maxillaires définitives (dents 16 et 26) en prenant comme référence leur face mésiale (dimension sagittale, mesure intra-arcade) ;
- Distance intermolaire de 16/26 (dimension horizontale, mesure intra-arcade) : distance séparant la cuspide mésiobuccale des premières molaires maxillaires définitives(dents 16 et 26) ;
- Classe canine droite selon la classification d'Angle (Campbell et Goldstein, 2021) basée sur le rapport mésiodistal des dents (dimension sagittale, mesure inter-arcade) ;
- Le surplomb de 11/41 – l'écart évalué sur le plan occlusal entre la face vestibulaire des incisives mandibulaires et les bords occlusaux des incisives centrales maxillaires (dimension sagittale, mesure inter-arcade) ;
- Alignement de 11/41 – Alignement des incisives mandibulaires par leurs maxillaires opposés (dimension verticale, mesure inter-arcade).

[7] Eminence de forme pyramidale située sur la face broyeuse des dents, en nombre variable selon le type de dent.

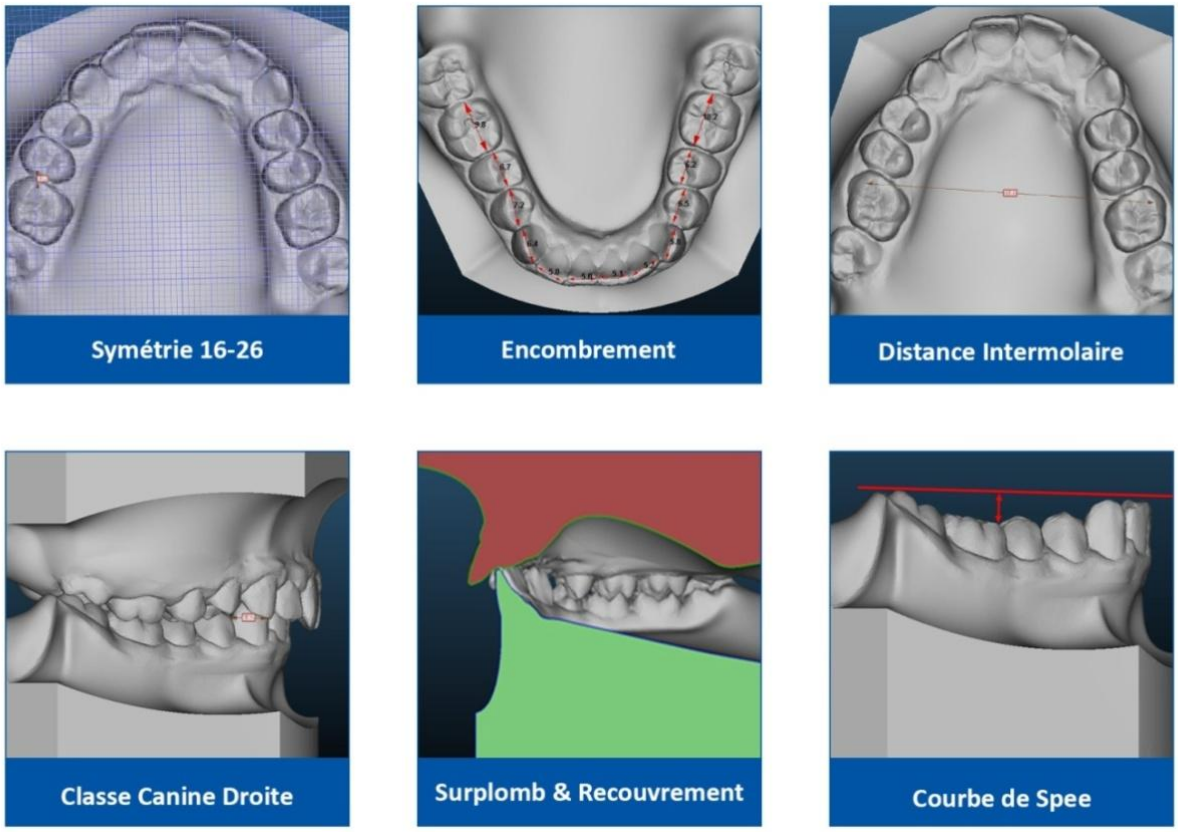


Figure 26 : Présentation des sept variables à estimer.

La fiabilité et la précision des mesures effectuées à l'aide de différents logiciels ont déjà été démontrées (Rossini et al., 2016 ; Kumar et al., 2015 ; Reuschl et al., 2016). Les mesures réelles ont été calculées pour chaque cas par le même opérateur à l'aide du logiciel Insignia® puis vérifiées par un deuxième opérateur.

Pour chacune des 7 variables, la différence en valeur absolue entre l'estimation de la mesure par le praticien et la mesure de la variable réalisée à l'aide du logiciel Insignia® a été déterminée (Figure 26).

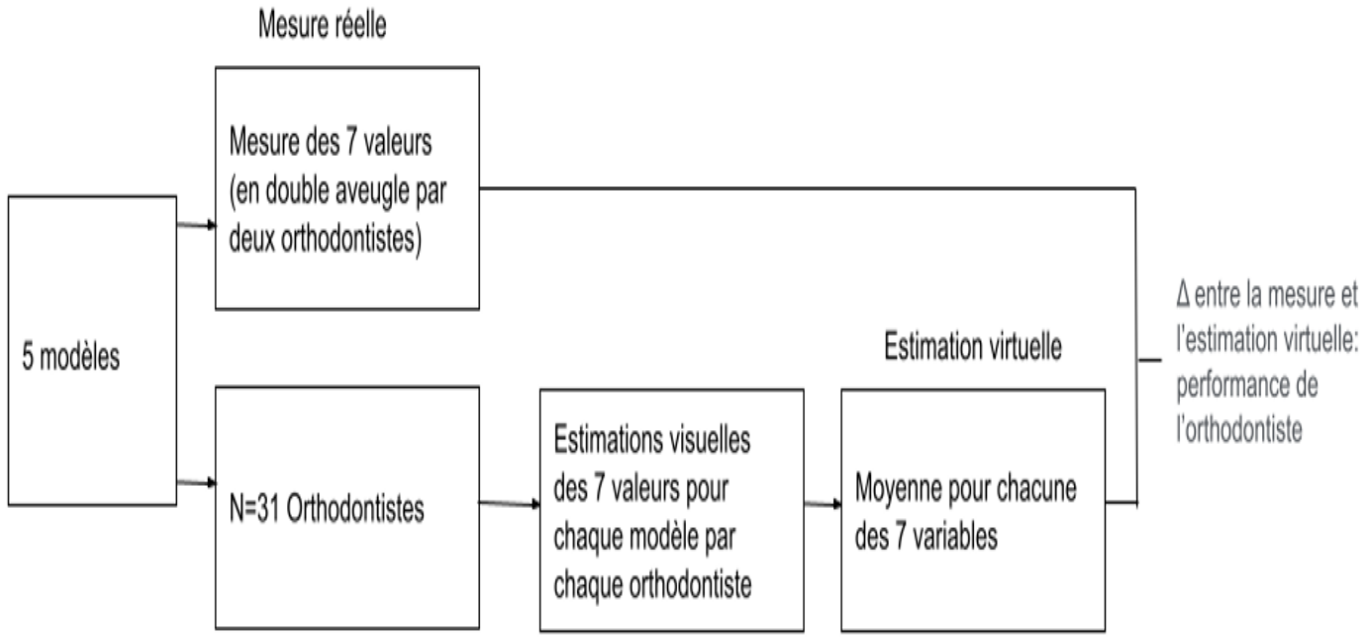


Figure 27 : Organigramme de la procédure.

4-2-2 Analyses statistiques

Les analyses ont été réalisées à l'aide de tests t sur l'échantillon unique (one Sample t-Test). Ce test permet généralement de comparer la moyenne d'un échantillon pour une variable précise (ici la moyenne des estimations de l'orthodontiste) avec une moyenne de référence (ici la distance réelle mesurée à l'aide du logiciel insignia).

La méthode employée pour déterminer la performance pour chacune des variables compte 5 étapes :

Etape 1 : Dans un premier temps, on détermine, pour chaque cas et pour les 31 orthodontistes, l'écart en valeur absolue entre l'estimation de la mesure par le praticien et la mesure de la variable pour chacun des 5 cas effectuée au moyen du logiciel : c'est la performance individuelle par cas et par variable. On obtient alors 5 mesures par orthodontiste et par variable

Etape 2 : Pour chacune des variables et pour chaque orthodontiste, on détermine la performance globale qui est la moyenne des 5 valeurs précédemment obtenue.

Etape 3 : Est aussi précise que les mesures effectuées au moyen du logiciel. Le choix s'est donc porté sur un test de moyenne à échantillon unique appelé *One Sample t-Test*, la moyenne de chaque estimation étant comparée à une moyenne théorique de référence.

Etape 4 : Si le test à échantillon unique avec valeur de référence égale à 0 s'avère non

significatif, on effectuera également un test à échantillon unique à partir des mêmes données, mais avec valeur de référence égale à 1, afin d'estimer si la population des praticiens arrive à estimer la variable au millimètre près.

4-3 Résultats

Les résultats obtenus sont présentés dans le Tableau 3.

	Cas n°1	Cas n°2	Cas n°3	Cas n°4	Cas n°5
Encombrement	x1: -3 x2: -2.73 σ: 1.43	x1: 0.5 x2: -2.35 σ: 3.21	x1: -3 x2: -3.55 σ:1.26	x1: -3.5 x2: -4.44 σ: 1.5	x1: 2.5 x2: -4.60 σ: 7.47
Distance intermolaire	x1: 36 x2: 38.10 σ: 7.82	x1: 37 x2: 38.6 σ: 8.22	x1: 36.5 x2: 38 σ: 8.73	x1: 40 x2: 36.6 σ: 8.11	x1: 40 x2: 41 σ: 8.36
Courbe de Spee	x1: 2 x2: 2.42 σ: 0.81	x1: 1.5 x2: 1.66 σ: 0.90	x1: 2.5 x2: 2.47 σ: 0.86	x1: 2 x2: 1.03 σ: 1.07	x1: 2.5 x2: 2.2 σ: 0.87
Symétrie 16/26	x1: 3 x2: 1.52 σ: 1.80	x1: 1 x2: 0.32 σ: 0.97	x1: 1 x2: 0.11 σ: 1.19	x1: 0 x2: -0.32 σ: 0.77	x1: 1.5 x2: 0.19 σ: 1.60
Classe de la canine droite	x1: 6 x2: 5.45 σ: 1.47	x1: 1 x2: 3.24 σ: 2.57	x1: 8 x2: 7.34 σ: 2.19	x1: -4 x2: -2.75 σ: 3.08	x1: 2 x2: 2.95 σ: 1.36
Surplomb	x1: 1.5 x2: 1.45 σ: 1.23	x1: 3.5 x2: 4.32 σ: 1.81	x1: 14 x2: 10.94 σ: 4.47	x1: 3 x2: 2.97 σ: 1.04	x1: 4.5 x2: 5.08 σ: 1.96
Recouvrement	x1: 7 x2: 6.27 σ: 1.60	x1: 5 x2: 3.52 σ: 1.80	x1: 5 x2: 9.55 σ: 1.11	x1: -5 x2: 4.18 σ: 2.05	x1: 4 x2: 3.63 σ: 0.97

Tableau 3 : Moyennes et écarts types pour chaque modèle. x1 représente la valeur obtenue via le logiciel de mesure, x2 la moyenne des mesures et σ l'écart type.

Sur la base des résultats du test t sur échantillon unique, les 7 variables se répartissent en 3 groupes :

1) Variables pour lesquelles les estimations n'étaient pas significativement différentes des valeurs réelles ; dans ce cas, la valeur estimée n'était pas significativement différente de la valeur réelle ;

2) Variables pour lesquelles les estimations étaient significativement différentes de la valeur réelle 0 mais pas de la valeur 1 ; dans ce cas, l'erreur d'estimation était supérieure à 0 mais inférieure à 1 cm ;

3) Variables pour lesquelles l'erreur d'estimation était significativement différente de la valeur 1 ; dans ce cas, l'erreur d'estimation était supérieure à 1 cm (tableau 4). Ce tableau 4 (colonne puissance) montre également la puissance de chaque résultat pour un risque de 20 %. Les valeurs des colonnes indiquent la taille minimale de l'échantillon pour que la conclusion à H0 soit suffisamment solide. Cette analyse de puissance montre que les échantillons de taille sont suffisants pour tous les résultats significatifs, à l'exception de la variable de classe canine droite pour une valeur théorique de 0 (mais pour cette même variable, l'échantillon de taille est suffisant pour une valeur théorique de 1).

- Variables pour lesquelles les estimations n'étaient pas significativement différentes des valeurs réelles. Ce fut le cas pour les trois variables suivantes :

Le surplomb ($t(30) = -1,60$; $p = 0,11$) ;

La courbe de Spee ($t(30) = -1,33$; $p = 0,19$) ;

La distance intermolaire ($t(29) = 0,39$; $p = 0,69$).

- Variables pour lesquelles les estimations étaient significativement différentes de la valeur réelle 0 mais pas de la valeur 1. C'était le cas pour les trois variables suivantes :

La symétrie 16/26 : valeur = 0 ($t(30) = -9,65$; $p \leq 0,01$), mais avec une valeur test = 1 ($t(30) = -0,66$; $p = 0,51$) ;

Le recouvrement : valeur test = 0 ($t(30) = -5,40$; $p < 0,01$), mais avec une valeur test = 1 ($t(30) = 0,64$; $p = 0,52$) ;

La classe canine droite : valeur test = 0 ($t(30) = 3,19$; $p = \leq 0,01$), mais avec une valeur test = 1 ($t(30) = -0,51$; $p = 0,61$).

- Variables pour lesquelles l'erreur d'estimation était significativement différente de la valeur 1. Ce n'était le cas que pour la variable d'écart mandibulaire : valeur test = 0 (t (30) = -10,74 ; p 0,01), valeur test = 1 (t (30) = 6,23 ; p 0,01).

t=0	Moyenne	Ecart type	t	Sig.	Puissance (0.8)
Encombrement	-2.23	1.15	-10.74	≤ 0.01	5
Classe canine droite	0.62	1.08	3.19	≤ 0.01	48
Courbe de Spee	-0.13	0.58	-1.33	0.19	313
Distance intermolaire	0.56	7.79	0.39	0.69	3049
Recouvrement	-0.65	0.66	-5.40	≤ 0.01	17
Surplomb	-0.34	1.20	-1.60	0.11	197
Symétrie 16/26	-0.93	0.53	-9.65	≤ 0.01	6

t=1	Moyenne	Ecart type	t	Sig.	Puissance (0.8)
Encombrement	2.25	1.11	6.23	≤ 0.01	4
Classe canine droite	0.92	0.84	-0.51	0.61	14
Recouvrement	0.80	0.44	0.64	0.52	5
Symétrie 16/26	0.93	0.53	-0.66	0.51	6

Tableau 4 : Résultats du test à un échantillon unique pour l'ensemble les variables. Ce tableau montre la valeur moyenne, l'écart type, la valeur du test t et la signification pour chaque variable. Le tableau du haut présente les résultats avec 0 cm comme valeur de référence. Le tableau du bas présente les résultats avec 1 cm comme valeur de référence (seules les variables qui se sont avérées significatives avec la valeur 0 sont testées avec la valeur 1).

4-4. Discussion

Les résultats de l'analyse des variances effectuée sur la moyenne des trois distances estimées ont montré que les distances étaient surestimées tant dans les environnements réels que virtuels. Ces résultats sont cohérents avec les résultats trouvés par d'autres auteurs. Plusieurs

recherches ont montré une surestimation de la distance d'encombrement mandibulaire lorsqu'elle est estimée par des orthodontistes visualisant des modèles réels (Johal et Battagel, 1997 ; Naish et al., 2016 ; Wallis et al., 2014). Makaremi et N'Kaoua, (2021) ont montré une surestimation des distances d'encombrement mandibulaire estimées par 31 orthodontistes sur des modèles virtuels. De plus, plusieurs études se sont intéressées à la perception des distances exocentriques (distances entre deux objets, dans notre cas entre deux éléments d'un même objet) au sein d'environnement virtuel non-immersif. Par exemple, Waller (1999) a trouvé que la distance exocentrique au sein de dispositifs non-immersif. Cependant, Wartenberg et Wiborg (2003) ont démontré que la distance exocentrique était généralement surestimée au sein d'environnements virtuels immersifs, tandis qu'une sous-estimation et une surestimation peuvent survenir de manière équivalente au sein des EV non immersif. De plus, ces auteurs ont observé qu'une mauvaise estimation des distances était amplifiée par le fait d'interagir avec un environnement non-immersif.

Nos résultats montrent que les distances étaient moins bien perçues avec les modèles virtuels qu'avec les modèles en plâtre. En effet, les distances étaient davantage surestimées dans la condition RV que dans la condition réelle.
Une explication possible réside dans l'effet de différents indices de profondeur non picturaux sur l'estimation de la distance. Ces indices correspondent, entre autres, aux informations du système oculomoteur binoculaire humain qui permet la perception de la profondeur. Ainsi, la disparité binoculaire et l'accommodation ont une grande importance dans la perception et l'estimation des distances, notamment dans l'espace personnel allant de l'observateur à 2 mètres (Cutting et Vishton, 1995).

Les modèles en plâtre permettent un accès naturel à divers indices de profondeur non picturaux, alors que dans le cas des modèles virtuels, présentés ici à l'aide d'un écran d'ordinateur, ces indices donnent des informations de profondeur relatif à la distance entre l'utilisateur et l'écran et non pas entre l'utilisateur et les objets eux-mêmes (Stefanucci et al., 2015). La disparité binoculaire désigne la différence entre les images perçues par l'œil droit et l'œil gauche (Cutting et Vishton, 1995). Wainman et al, (2018) ont montré, dans une tâche nécessitant une compréhension spatiale de l'anatomie pelvienne féminine, l'importance de cet indice dans la différence de performance entre les modèles physiques et virtuels. Dans cette étude, les participants avaient l'œil non dominant caché (éliminant la vision binoculaire) ou non

lors de l'estimation de la distance sur des modèles virtuels ou physiques. Lorsque la vision binoculaire était disponible, le modèle physique conduisait à une meilleure performance que le modèle virtuel. Cependant, sans vision binoculaire, la performance sur le modèle physique a diminué alors qu'elle n'a pas changé pour le modèle virtuel (pas de différence significative entre les modèles physiques et virtuels dans la condition sans vision binoculaire). Ces résultats indiquent que la performance en analyse spatiale est diminuée par le peu d'information fournie par la vision binoculaire sur les modèles virtuels par rapport aux modèles physiques.

De plus, en situation virtuelle, l'accommodation se fait sur le plan constitué par l'écran de l'ordinateur plutôt que sur les objets virtuels eux-mêmes, ce qui diminuerait l'acuité visuelle en profondeur (Huckauf, 2005). Par conséquent, la perception de la distance des modèles virtuels serait mal orientée par l'accommodation (Bruder et al., 2016 ; Drascic et Milgram, 1996 ; Renner et al., 2013).
Ainsi l'absence d'informations provenant de l'accommodation et de la disparité binoculaire sur les éléments présentés au sein de l'EV sont donc des facteurs ayant pu diminuer l'estimation des distances au sein de l'EV non-immersif.

D'autres facteurs, tels que le type d'apprentissage effectué par les participants ou leurs capacités visuospatiales peuvent également être impliqués dans la diminution de la performance sur les modèles virtuels par rapport aux modèles physiques. En effet, il a été démontré que des capacités visuospatiales faibles entraînent une diminution de la performance sur les modèles virtuels alors que des capacités visuospatiales fortes ou faibles n'ont pas d'impact sur la performance sur les modèles physiques (Wainman et al, 2020). Par ailleurs, Wainman et al. (2018) ont analysé l'effet du type d'apprentissage (apprentissage réalisé sur modèle physique vs modèle virtuel) du bassin féminin sur une tâche nécessitant une analyse spatiale sur un écran d'ordinateur. Ces auteurs ont montré que l'apprentissage sur un modèle physique conduisait à de meilleures performances que l'apprentissage sur un modèle virtuel. L'effet du type de formation (les professionnels ont été formés soit sur des modèles réels, soit sur des modèles virtuels, soit sur les deux types de modèles au cours de leurs études) et des capacités visuospatiales pourrait être étudié dans de futures études.

De plus, dans notre étude, les participants ont utilisé pour interagir avec l'EV non-immersifs, un clavier et une souris. Cependant, cette méthode ne serait pas optimale pour la perception de la distance. En

effet, la manipulation d'objet en 3D effectué à l'aide d'un curseur en 2D (une flèche pouvant se déplacé seulement en 2 dimensions) contrôlé grâce à une souris entrainerait un conflit entre l'espace en 2D du curseur et l'espace en 3D du contenu, en plus de ne donner aucun indice sur la profondeur des objets sélectionnés (Jauregui et al., 2012).

Enfin, une dernière source d'explication est que les tâches effectuées sur un modèle virtuel entraînent une charge mentale plus élevée que celles effectuées sur un modèle réel. En effet, une augmentation de la charge mentale est associée à une diminution de la performance (Chanquoy et al., 2007 ; Jaquess et al., 2017). Or, l'un des résultats de notre étude est que la charge mentale était plus élevée lorsque les estimations étaient réalisées sur un écran d'ordinateur que lorsqu'elles étaient réalisées sur les modèles en plâtre. Ce résultat est en accord avec différentes études qui ont montré une augmentation de la charge mentale dans les tâches réalisées en RV par rapport à ces mêmes tâches réalisées dans l'environnement réel (Koumaditis et al., 2020 ; Stevens et al., 2016). En effet, plusieurs auteurs montrent que l'interaction avec un VE, de par sa nature complexe, pourrait augmenter la charge cognitive (Fraser et al., 2012 ; Huang et al., 2020). Koumaditis et al. (2020) ont par ailleurs observé que l'augmentation de la charge mentale était associée à une diminution de la performance dans une tâche de résolution de puzzle 3D réalisée en RV.

Ainsi, dans notre étude, pour compenser l'absence d'indices visuels non picturaux dans la condition du modèle virtuel, les praticiens peuvent avoir dû compter sur leurs compétences visuospatiales et/ou d'autres types de traitement cognitif pour appréhender le modèle virtuel et estimer les distances dans le modèle, ce qui a entraîné une augmentation de la charge mentale et une diminution de la performance. Une fois encore, ce type d'hypothèse doit être étudié plus avant.

Concernant l'effet de la complexité du modèle, nos résultats indiquent que plus le niveau de complexité augmente, plus la charge mentale augmente tant pour les modèles en plâtre que pour les modèles virtuels. Ces résultats sont en accord avec la littérature montrant qu'une charge mentale élevée est associée à une complexité élevée de la tâche (Chao et al., 2017 ; Das et al., 2020 ; Jacko et Ward, 1996 ; J. Li et Kim, 2021). Par exemple, Haji et al. (2015) ont montré que la charge mentale était plus élevée dans une tâche complexe que dans une tâche simple lors d'une simulation physique d'un nœud pour fixer un tube thoracique. Di Stasi et al. (2016),

quant à eux, ont constaté que la charge mentale augmentait avec le niveau de complexité de la tâche dans différents exercices de laparoscopie réalisés en RV. Nos résultats montrent que c'est également le cas dans les tâches d'estimation de distance réalisées en réel ou en virtuel, où l'augmentation de la complexité des modèles à estimer augmente la charge mentale. Nos résultats montrent également que l'augmentation de la complexité est associée à une augmentation de la charge mentale mais à une amélioration des performances d'estimation de distance (pour l'encombrement mandibulaire et l'inter-canine). Or, l'augmentation de la charge mentale est généralement associée à une baisse des performances (Chanquoy et al., 2007 ; Jaquess et al., 2017). Il faut cependant noter que certaines études ont montré que la relation charge mentale/performance est plus complexe et peut avoir une forme de U inversé (Babiloni, 2019 ; Basahel et al., 2010 ; Bruggen, 2015). La performance peut augmenter avec la charge mentale jusqu'à atteindre un seuil à partir duquel elle diminue lorsque la charge mentale augmente.

Comme indiqué ci-dessus, nos résultats montrent que l'augmentation de la complexité des modèles (virtuels et réels) était associée à de meilleures estimations de distance pour l'encombrement mandibulaire et la distance inter-canine alors qu'aucune différence d'estimation en fonction de la complexité de la tâche n'a été trouvée pour la distance inter-molaire. Un tel résultat pourrait alors s'expliquer par les questions d'estimation de distance réalisées dans notre étude. En effet, l'encombrement mandibulaire et la distance inter-canine ont un rôle important dans le diagnostic et donc dans le traitement qui sera mis en place alors que la distance inter-molaire est une condition qui ne se rencontre pas dans la pratique professionnelle. Des auteurs ont montré que les objectifs de la tâche médiatisent la relation entre la performance et la complexité de la tâche. Des objectifs importants dans des tâches complexes peuvent améliorer la performance par des processus motivationnels et cognitifs qui conduisent au développement de stratégies efficaces (Campbell, 1988 ; Wood et al., 1987). Ainsi, dans notre étude, l'estimation de la distance dans des cas complexes peut être associée à des enjeux de chirurgie plus sévère ou de prise en charge ultérieure potentiellement plus impactante. Ces enjeux plus importants lors de l'estimation de modèles complexes pourraient être à l'origine d'une attention accrue lors de la prise d'information de distance sur les modèles en réel comme en virtuel, améliorant ainsi les performances. Une telle hypothèse reste évidemment à tester dans des études ultérieures, notamment en termes d'impact de l'enjeu thérapeutique sur la précision des réponses d'estimation en environnement réel et virtuel. Une

autre source d'explication possible de cet effet de complexité sur les performances d'estimation de distance peut être en termes de niveau d'expérience. Mais notre étude n'a pas mis en évidence de lien entre le niveau d'expérience et le niveau de précision de l'estimation de distance (quelle que soit la condition d'estimation considérée), même si, comme indiqué dans la partie résultats (section 3.2.4), les résultats des corrélations sont à prendre avec précaution. De futures études prenant plus systématiquement en compte le niveau d'expérience permettront également de préciser le lien entre niveau de complexité, niveau d'expérience et estimation des distances en situation réelle et virtuelle (par exemple en comparant différents groupes de praticiens ayant des niveaux d'expérience différents).

Enfin, d'autres études complémentaires pourraient également permettre de comparer les estimations de distance à l'aide de différents dispositifs de RV, du moins immersif (comme celui que nous avons utilisé) au plus immersif (affichage monté sur la tête, par exemple).

4-5. L'essentiel à retenir

Les modèles numériques 3D d'arcades dentaires peuvent être considérés comme une réalité virtuelle non immersive caractérisée par un faible degré d'interaction. Ces modèles sont aujourd'hui largement utilisés pour réaliser des diagnostics, mais l'estimation des distances par l'utilisateur (orthodontiste) réalisée sur ce type de modèle numérique, et qui constitue un facteur déterminant dans la prise de décision thérapeutique, n'a pas encore été étudiée.

C'est dans ce cadre que nous avons mené notre étude. Nos résultats d'estimations de distances à partir de modèles numériques ont montré que 4 des 7 variables testées étaient associées à un manque de précision dans l'estimation des distances, la taille de l'arc mandibulaire étant la variable la moins bien estimée.

Bien qu'aucune étude n'ait été réalisée dans le domaine de l'orthodontie, le manque de précision dans l'estimation des mesures effectuées à partir de modèles virtuels 3D correspond à ce qui a été décrit dans les études relatives aux estimations de distance dans les environnements numériques.

D'autres études sont nécessaires pour mieux comprendre les variables (liées aux professionnels, aux équipements, au type de modèles etc.) susceptibles d'affecter l'estimation des distances et, par conséquent, la prise de décision thérapeutique. De telles études sont fondamentales notamment pour améliorer la formation des professionnels à l'utilisation des nouveaux outils numériques. Ces objectifs peuvent être mieux traités par une approche multidisciplinaire

(Makaremi, 2016), afin de contribuer à maintenir l'expertise et la perception du praticien au centre de la prise de décision, assurant ainsi une transition numérique réussie.

Les conclusions de cette première étude nous ont servi de base pour mener une étude complémentaire plus approfondie. L'étude suivante compare donc des modèles réels à des modèles virtuels, tout en faisant varier la complexité des modèles à estimer. La charge mentale est évaluée dans toutes les situations et l'effet de l'expertise (nombre d'année d'expérience professionnelle) sur les précisions d'estimations (en réel et en virtuel) est également déterminé.

5 : Performance et charge mentale de l'orthodontiste dans l'estimation des distances sur modèles réels et virtuels.

Cette étude intitulée Estimation of distances within real and virtual dental models as a function of task complexity. Est en cours de soumission dans une revue scientifique, le manuscrit de soumission étant intégré aux annexes (Annexe2).

5-1 : Contexte et intérêt de l'étude

L'étude présentée dans le chapitre 3 ainsi que les données de la littérature spécialisée sur l'estimation des distances en réalité virtuelle semblent confirmer que la numérisation des modèles d'étude orthodontiques impacte l'estimation des variables intra et interarcades par le praticien. Menée à partir de l'analyse de notre première étude, ainsi que de données bibliographiques, cette deuxième étude a pour objectif d'approfondir la compréhension des conséquences que ce changement de paradigme implique pour l'orthodontiste.
Dans le cadre de ce deuxième protocole, les estimations ont été réalisées de façon concomitante sur des modèles présentés de façon aléatoire, en mode réel et virtuel, ce qui permet d'objectiver de manière beaucoup plus pertinente la différence entre ces 2 situations dans l'estimation des différentes variables.

Trois variables ont été sélectionnées :

- L'encombrement de l'arcade mandibulaire a été choisie compte tenu des conclusions de la première étude, qui l'a identifiée comme la variable dont l'estimation par les praticiens est la plus biaisée,
Des mesures linéaires simples entre deux points A et B bien définis ont également retenues comme variables à estimer. Il s'agit des distances intercanine et intermolaire : la distance intercanine est la distance transversale rectiligne mesurée entre le sommet des pointes canines droite et gauche, alors que la distance intermolaire est la distance transversale rectiligne joignant le sommet des cuspides mésio-vestibulaires des premières molaires permanentes.

- De plus, dans la pratique orthodontique, les modèles à estimer pour prendre des décisions thérapeutiques sont de complexité variable. Or, à notre connaissance, aucune étude ne s'est

centrée sur l'effet de la complexité des taches sur les éventuelles erreurs d'estimations des distances réalisées en situation réelle ou virtuelle. La complexité des modèles a donc également été prise en compte dans notre étude.

- Enfin, la charge mentale est couramment considérée, au sein de la littérature spécialisée, comme un facteur en lien avec la complexité de la tâche et donc susceptible d'impacter les performances au sein d'environnements virtuels ou réels. L'étude de la charge mentale lors de l'estimation des variables est une donnée complémentaire tout à fait pertinente, qui a donc été incluse dans notre étude.

Ainsi, l'objectif de notre étude est de comparer l'estimation des distances réalisée, lors du diagnostic, par les praticiens sur des modèles réels et virtuels, en prenant en compte la complexité des modèles proposés.

5-2 Matériel et méthode

5-2-1 Sujets

Un total de 50 orthodontistes (31 femmes et 19 hommes), âgés de 36 ans en moyenne (σ = 12,84 ; min = 23 ; max = 63) ont participé à cette expérience. Une partie des participants étaient en cours de spécialisation à l'université (n = 23 ; première année = 7 ; deuxième année = 7 ; troisième année = 9), et une autre partie étaient diplômés chirurgiens-dentistes et spécialisés en orthopédie dento-faciale (n = 27). Les lieux de formation des participants étaient les universités de Bordeaux (n = 31), de Rennes (n = 13), de Toulouse (n = 3), de Nantes (n = 1), de Montpellier (n = 1) et de Paris (n = 1). L'expérience clinique de chaque participant a également été demandé.

6 conditions expérimentales étaient proposées aux sujets. Nous avons comparé des modèles virtuels à des modèles réels, et, pour chaque condition (réelle vs. Virtuel) 3 niveaux de complexité des modèles (facile, moyen, difficile) étaient proposés. Tous les participants étaient soumis à l'ensemble des 6 conditions.

		N	%
Genre	Hommes	19	38
	Femmes	31	62
Age	Moyenne	36	
	Ecart type	12.84	
Expérience	Interne 1ere année	7	14
	Interne 2éme année	7	14
	Interne 3éme année	9	18
	Diplômé	27	54
Lieu de formation	Bordeaux	31	62
	Rennes	13	26
	Toulouse	3	6
	Nantes	1	2
	Montpellier	1	2
	Paris	1	2
Utilisation des modèles numériques en pratique quotidienne	OUI	34	68
	NON	16	32
Méthode d'apprentissage (Pour les internes N=23)	Modèles réels	13	56.5
	Modèles virtuels	10	43.5

Tableau 5 : Tableau descriptif de la population.

5-2-2 Matériel

Pour réaliser cette étude nous avons utilisé des modèles d'arcades dentaires et des questionnaires que les participants ont remplit après avoir effectué les tâches proposées. L'ordre de présentation de chaque modèle, l'ordre de présentation des niveaux de complexité ainsi que les modèles présentés en mode réel et virtuel ont été choisis aléatoirement en amont par tirage au sort par les participants.

5-2-2-1 Modèles d'arcade dentaire

En préambule à notre recherche, nous avons identifié des modèles d'arcades dentaires de complexité différente (trois niveaux de complexité ont été retenus). Dans ce contexte, nous avons combiné les paramètres intraarcades des trois indices utilisés dans la classification de la difficulté de cas en orthodontie (Bellot-Arcis et al., 2012) :

- Le DHC (dental health component) : composante morphologique de l'IOTN (Indice of Orthodontic Treatment Need) ;
- Le PAR Index (Peer Assessment Rating Index) ;
- Le DAI Index (Dental Aesthetic Index).

Pour des raisons méthodologiques, chaque modèle devait satisfaire aux critères d'inclusion suivants : denture adolescente, adulte jeune ou adulte stable et complète, absence d'agénésie ou d'inclusion, et absence de dents extraites.

Cela nous a permis de répartir les modèles dentaires en trois catégories selon le degré de difficulté :

- « Facile » : encombrement antérieur seul, avec décalage des points de contact de 1 à 2 mm, sans anomalie de forme au niveau de l'arcade (distances intercanine et intermolaire normales), ni anomalie dentaire ;
- « Moyen » : encombrement antérieur engendrant un décalage des points de contact de 2 à 4 mm, avec rotation ou ectopie d'une dent par secteur ; asymétrie inférieure à 2 mm et/ou forme d'arcade pathologique ;
- « Difficile » : encombrement antérieur et moyen entrainant un décalage des points de contact de plus de 4 mm et/ou des positions ectopiques et des rotations de plusieurs dents par secteur ; asymétrie supérieure à 2 mm et/ou forme d'arcade pathologique.

La recherche a été effectuée parmi 2500 modèles numériques en 3D réalisés à l'aide de la caméra *Trios3*® de *3Shape*.

Après pré-sélection, 64 cas ont été retenus. Ces 64 cas ont été segmentés et analysés avec le logiciel *Orthoanalyzer*® pour définir la mesure des paramètres pris en compte dans notre étude (encombrement mandibulaire, distance intercanine, distance intermolaire) et chaque cas a été

catégorisé en tant que « facile », « moyen » ou « difficile ». Pour chaque niveau de difficulté, 2 modèles ont été sélectionnés pour leur intérêt clinique et leur similarité (soit 6 modèles au total).

Par la suite, les 6 modèles mandibulaires ont été imprimés à échelle réelle, à l'aide d'une imprimante de technologie DLP (Digital Light Processing) : la NextDent 5100® de 3D*Systems* et via le logiciel 3D *SprintBasic*®. Des duplicatas en plâtre de ces modèles ont été réalisés après réalisation d'empreintes à l'alginate.

Les modèles virtuels en 3D ont été enregistrés au format *dicom* sur un seul et même ordinateur afin qu'ils puissent être présentés dans les mêmes conditions expérimentales aux participants, qui auront la possibilité de les manipuler à leur guise à l'aide d'une souris en 2D sur le logiciel *OrthoViewer®*.

Les mesures réelles d'encombrement, de distance intercanine et de distance intermolaire, déjà réalisées via le logiciel *Orthoanalyzer®*, ont été de nouveau réalisées sur les modèles en plâtre à l'aide d'un pied à coulisse numérique et de fil de laiton (mesure traditionnelle). Cette opération a été effectuée par deux opérateurs distincts.

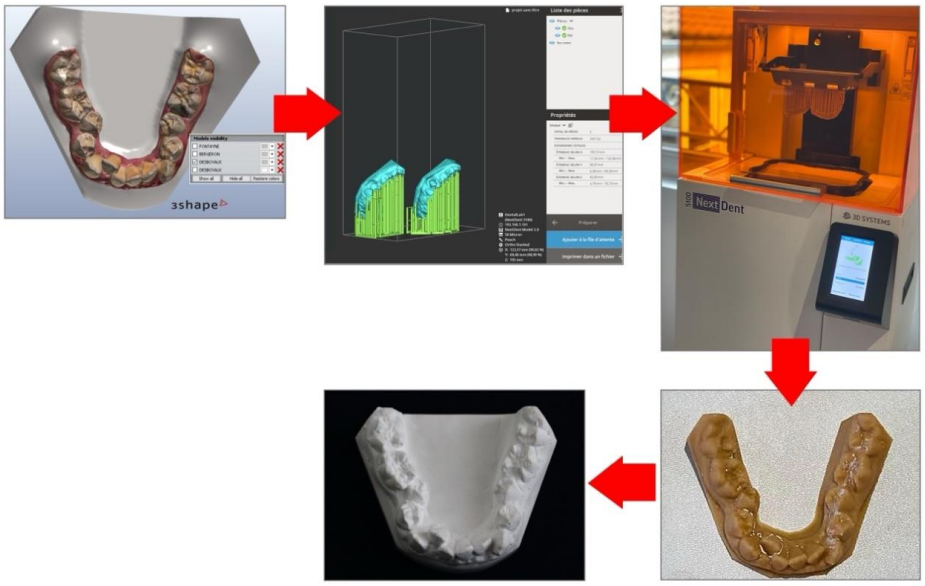


Figure 28 : Processus de fabrication des modèles en plâtre à partir des empreintes numériques des arcades mandibulaires.

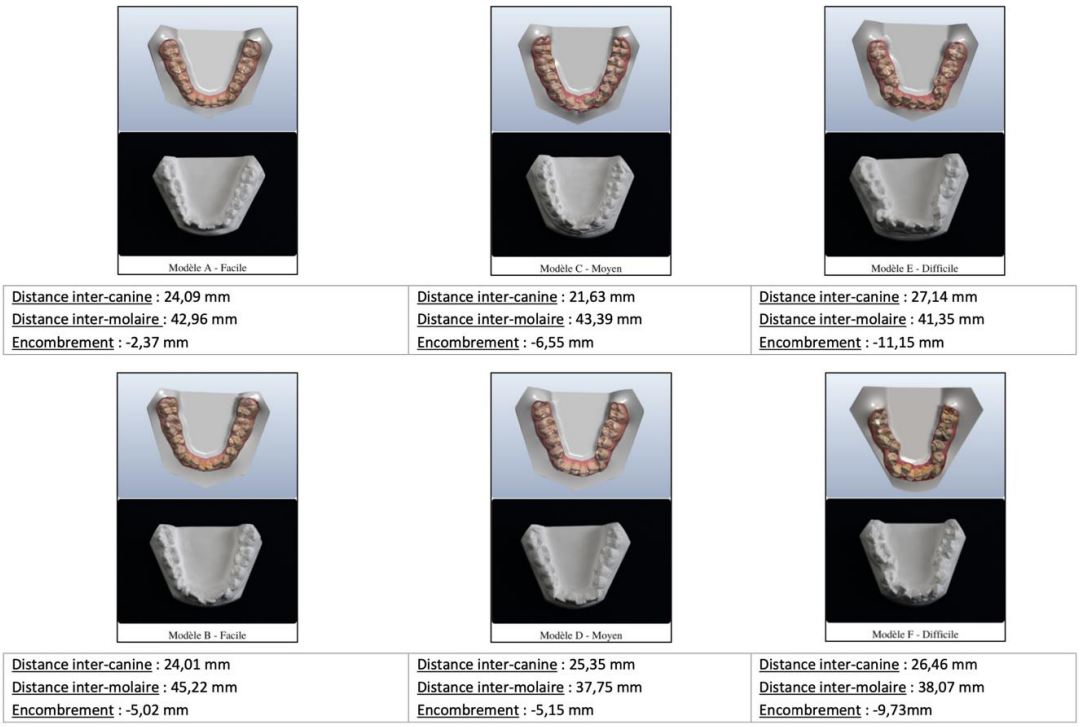


Figure 29 : Modèles mandibulaires inclus dans l'étude.

5-2-2-2 Questionnaires

Divers questionnaires ont été soumis à chaque participant.

Données démographiques

Avant la réalisation des tâches, un questionnaire de données démographiques a été rempli par les participants (appelés à spécifier leur genre, leur âge, l'année où ils ont débuté leur carrière comme orthodontistes, leur qualité d'internes le cas échéant, leur lieu de formation/spécialisation, l'utilisation de modèles virtuels et/ou réels au cours de leur formation/pratique).

Rapport avec l'outil informatique

Un questionnaire a été élaboré visant à recueillir les expériences et les compétences informatiques des participants.

Charge mentale

Afin de vérifier a posteriori la pertinence du classement en « facile », « moyen » et « difficile » réalisé avant l'expérimentation, chaque participant avait pour tâche d'évaluer la charge mentale associée à chacun des modèles dont il devait estimer les distances. Le NASA-TLX est un questionnaire utilisant une technique de mesures subjectives, c'est-à-dire qui repose sur le jugement de la personne qui réalise la tâche ou l'action. Il a été développé dans les années 1980 par le Human Performance Group du NASA-Ames Research Center, CA (Sandra Hart). Le NASA-TLX est composé de 6 dimensions : demande mentale, demande physique, demande temporelle, performance, effort et frustration. Ces dimensions sont chacune présentées sous la forme d'une échelle de Likert (allant de 0 à 20), ce qui permet d'obtenir un score pour chaque dimension et un score global par participant. Le questionnaire a été présenté dans sa version simplifiée (Byers, Bittner et Hill, 1989).

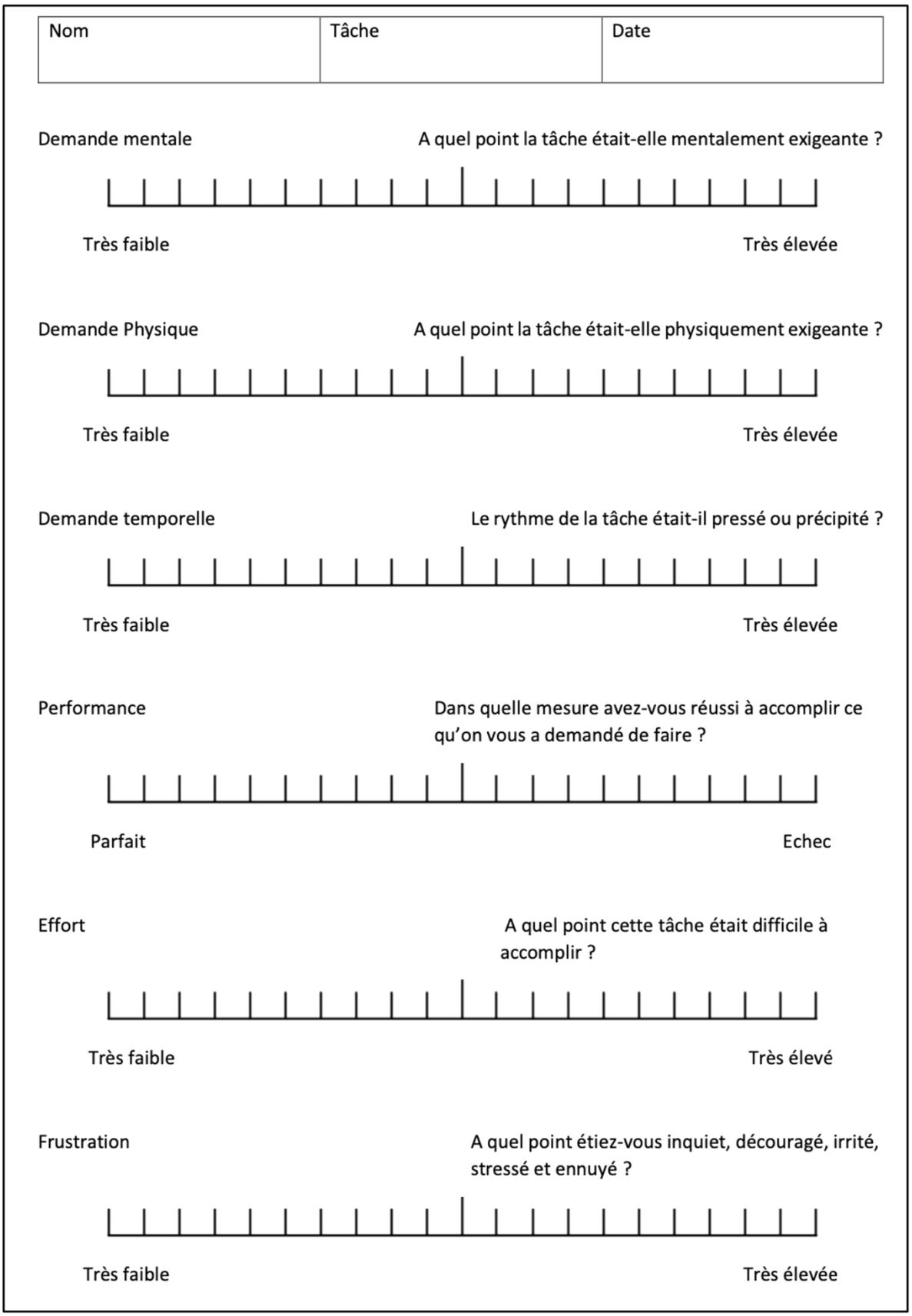


Figure 30 : Questionnaire de charge mentale – NASA-TLX (Byers, Bittner et Hill, 1989).

5-2-3 Procédure

Pour chaque participant, une session se divisait en trois phases. La première phase consistait en une « présentation » comportant le recueil du consentement par signature et l'explication détaillée de la tâche aux participants. Ensuite, il était demandé aux participants de remplir un questionnaire concernant leur expérience avec l'outil informatique. Ce questionnaire était présenté sur une tablette, sous forme de Google Forms.

Pendant la deuxième phase, les participants ont réalisé les estimations des distance demandées. L'ordre de présentation de chaque modèle, l'ordre de présentation des niveaux de complexité ainsi que les modèles présentés en mode réel ou virtuel avaient été choisis aléatoirement en amont.

Il a été demandé aux participants d'estimer trois variables quantitatives sur chaque modèle, tant en mode réel (en plâtre) qu'en mode virtuel (sur un écran d'ordinateur). Ces variables étaient l'encombrement mandibulaire, la distance intercanine et la distance intermolaire. Pour éviter toute interprétation dans la mesure des variables par les praticiens, une fiche explicative avec la définition précise de chacune des trois mesures à effectuer est remise au participant en amont de la tâche. Les estimations des modèles virtuels étaient réalisées sur le même ordinateur par l'ensemble des participants.

Les deux premières valeurs (encombrement mandibulaire et distance intercanine) rentrent dans le cadre des diagnostics couramment réalisés par les professionnels. La distance intermolaire, en revanche, est une condition contrôle qui n'est jamais estimée en pratique.

Ces valeurs s'expriment en millimètres et peuvent être arrondies au millimètre près.

La figure ci-dessous présente la fiche explicative distribuée aux participants avant l'expérience.

L'encombrement est un paramètre intra-arcade, évaluable à l'arcade maxillaire et à l'arcade mandibulaire. Il entre en compte dans l'appréciation de la dysharmonie dento-maxillaire (DDM). Dans cette étude, nous vous demandons la réalisation d'une estimation d'encombrement uniquement à **l'arcade mandibulaire** et sur les 10 dents antérieures (de la face mésiale de 36 à la face mésiale de 46) en partant du principe que :

Encombrement = ED – EN

Avec ED : espace disponible = périmètre d'arcade de mésial 36 à mésial 46
Et EN : espace nécessaire = somme des diamètres mésio-distaux des 10 dents antérieures des 35 à 45

Si l'Espace Nécessaire est supérieur à l'Espace Disponible (**EN>ED**), l'encombrement est dit **négatif** : **MANQUE DE PLACE.**

Si l'Espace Disponible est supérieur à l'Espace Nécessaire (**ED>EN**), l'encombrement est dit **positif** : **EXCES DE PLACE**.

Quelle que soit la valeur estimée, pensez à préciser le signe « + » (si excès de place) ou « - » (si manque de place). Cette valeur est exprimée en millimètres et peut être arrondie au ½ millimètre près (*exemple de réponse : -3,5mm d'encombrement*).

La **distance inter-canine** est la **distance transversale rectiligne mesurée entre le sommet des pointes canines droite et gauche**.
La **distance inter-molaire** correspond à la **distance transversale rectiligne joignant le sommet des cuspides mésio-vestibulaires des 1ères molaires permanentes**. Dans cette étude, nous vous demandons l'estimation de la distance inter-canine mandibulaire et de la distance inter-molaire mandibulaire. Ces valeurs s'expriment en millimètres et peuvent être arrondies au millimètre près.

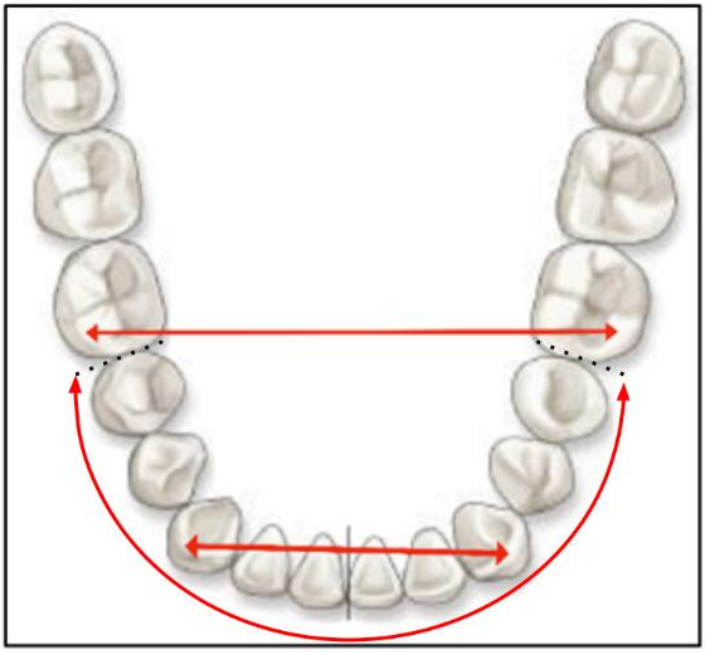


Figure 31 : Fiche explicative – mesure.

Après avoir estimé les distances sur un modèle dentaire, il a été demandé aux participants de remplir le questionnaire NASA-TLX. Ils ont donc répondu six fois au questionnaire. La figure 32 résume le déroulement de l'ensemble du protocole.

Cette méthodologie peut être considérée comme croisée et contrôlée par tirage au sort des modèles et de l'ordre, avec des sujets en aveugle quant au niveau de difficulté et aux valeurs réelles qu'ils ont à estimer. Le protocole est mis en place par un operateur unique et entrainé.

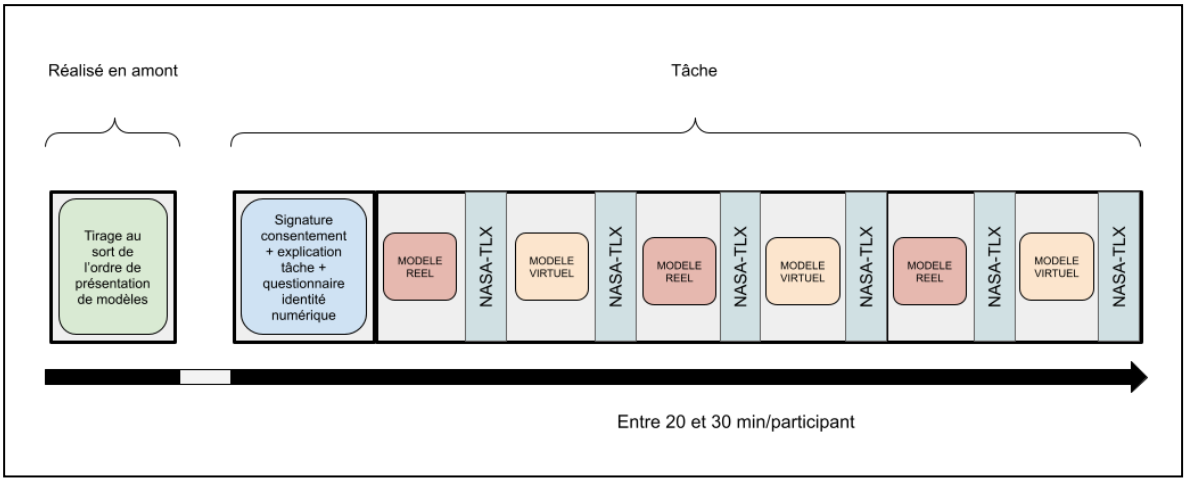


Figure 32 : Présentation schématique du protocole expérimental.

5-2-4 Analyse statistique

L'analyse des résultats a été réalisées à l'aide d'ANOVA à 2 facteurs.

Dans ces ANOVA, les deux facteurs considérés, le facteur « type de présentation » et le facteur « complexité », constituaient deux facteurs en mesures appariées. Le facteur « type de présentation » comportait 2 modalités (modèles réels et modèles virtuels), il s'agissait donc d'un facteur principal à deux modalités (réel – virtuel). Le facteur « complexité » était un facteur ordinal à trois modalités (facile – moyen – difficile). La variable dépendante (quantitative) était l'estimation de la distance. Avant la réalisation de chaque ANOVA, les échantillons ont été soumis au test de

normalité (Shapiro test) ainsi qu'au test d'homogénéité des variances (Test de Levene). Ces analyses ont indiqué que toutes les données suivaient bien des lois normales et que les variances étaient homogènes.
Dans toutes les situations observées, la variable dépendante correspondait à un indicateur = valeur (distance) estimée par le praticien – valeur (distance) réelle. La distance était surestimée lorsque l'indicateur était supérieur à 0, et sous-estimée lorsque l'indicateur était inférieur à 0.

Ces tests statistiques (ANOVA) ont permis de recueillir 3 types d'information :

1. Des informations sur le premier facteur : est-ce que les distances sont évaluées de manière différente selon la condition (réelle ou virtuelle) ?
2. Des informations sur le second facteur : est-ce qu'il y a des différences d'estimation selon le niveau de complexité ?
3. Des informations sur l'interaction entre les deux facteurs.

La charge mentale sera ici utilisée comme condition contrôle. Une ANOVA à 2 facteurs (facteur « présentation » et facteur « niveau de difficulté ») permet, dans un premier temps, de discuter le classement « facile », « moyen », « difficile » réalisé préalablement à l'étude, et dans un deuxième temps, de façon plus générale, d'évaluer la charge mentale selon le niveau de complexité et le caractère virtuel ou réel du modèle estimé.

5-3 Résultats

Les estimations de distances réalisées sur des modèles réels et virtuels ont été comparées dans les 3 situations distinctes que sont l'encombrement mandibulaire, la distance intercanine et la distance intermolaire. Comme déjà indiqué, les 2 premières situations correspondent à des estimations réalisées par le praticien dans le cadre du diagnostic. La 3ème situation est une situation contrôle qui ne rentre pas habituellement dans une prise d'information en vue d'une situation de soin.

5-3-1 Charge mentale

A la fin de chaque estimation, les participants étaient invités à remplir un questionnaire de charge mentale (NASA-TLX).

Une analyse à 2 facteurs appariés (facteur « présentation » – réel vs virtuel – et facteur « complexité » – facile, moyen, difficile –) a été réalisée avec la mesure de charge mentale comme variable dépendante.

Les résultats ont montré un effet principal du type de présentation (réel vs. virtuel) [F (1, 49) = 16.778, p < .05], ainsi qu'un effet principal du niveau de complexité [F (2, 98) = 5.331, p < .05]. En revanche, nous n'avons pas observé d'effet pour l'interaction entre le type de présentation et le niveau de complexité [F (2, 98) = 1.342, p > .05].

Les statistiques descriptives (Tabeau 6) ont montré que la charge mentale était plus importante en mode virtuel (M = 53.26) qu'en mode réel (M = 47.91) pour les 3 niveaux de complexité. Et pour les deux situations (réelle et virtuelle), le score de charge mentale était plus faible dans la condition « facile » (M = 48.35), augmentait dans la condition « moyenne » (M = 50.32), et atteignait son niveau le plus haut en condition difficile (M = 53.08) (voir figure 33).

	Facile		Moyen		Difficile	
	M	σ	M	σ	M	σ
REEL	44.70	17.57	47.22	19.02	51.38	20.00
VIRTUEL	51.50	18.62	53.10	17.47	54.76	18.92

Tableau 6 : Statistiques descriptives de la charge mentale

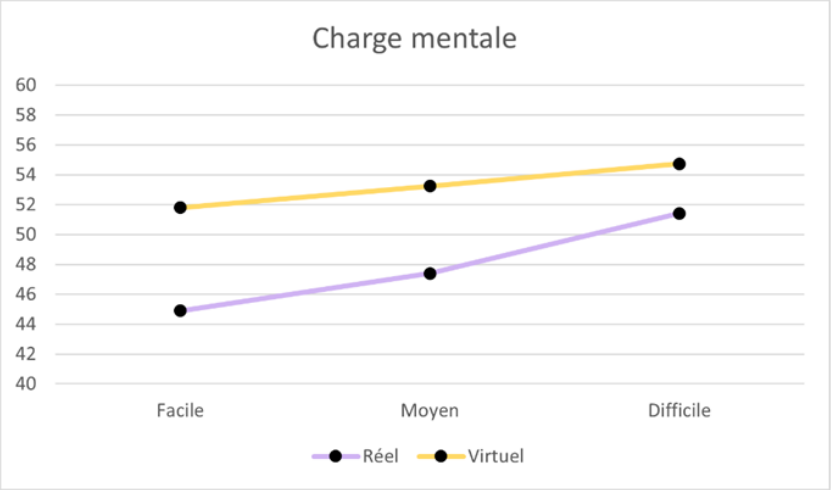


Figure 33 : Représentation graphique de la charge mentale selon le niveau de complexité

de la tâche (M = moyenne ; σ = écart-type).

Les résultats indiquent que la charge mentale de l'estimation des distances sur des modèles orthodontiques est significativement plus élevée sur modèles virtuels que sur modèles réels. Ils indiquent également que la charge mentale augmente avec la difficulté des cas, et ce quel que soit le support utilisé.

5-3-2 Estimations des distances

Nous avons tout d'abord réalisé une analyse globale de variances en moyennant les scores obtenus pour les 3 situations d'estimation (encombrement mandibulaire, distance intercanine et distance intermolaire). L'analyse montrait un effet principal du type de présentation (réel – virtuel) [F (1, 49) = 36.61, p < .05], un effet principal des niveaux de complexité [F (2, 98) = 47.476, p < .05], ainsi qu'une interaction entre type de présentation et niveaux de complexité [F (2, 98) = 12.744, p < .05]. Les statistiques descriptives (Tableau 6) ont montré que, globalement, les distances étaient surestimées quelle que soit la situation. Mais, la surestimation était plus importante en mode virtuel (M = 3.37) qu'en mode réel (M = 2.67). Pour ce qui est des niveaux de complexité, la surestimation était plus importante dans la condition « facile » (M = 3.58), diminuait dans la condition « moyenne » (M = 3.23), et atteignait son niveau le plus bas dans la condition « difficile » (M = 2.44). De plus, l'interaction était due au fait que les différences entre mode réel et mode virtuel (des surestimations plus importantes ont été observées avec des modèles virtuels) disparaissaient en situation « difficile » (voir figure 34).

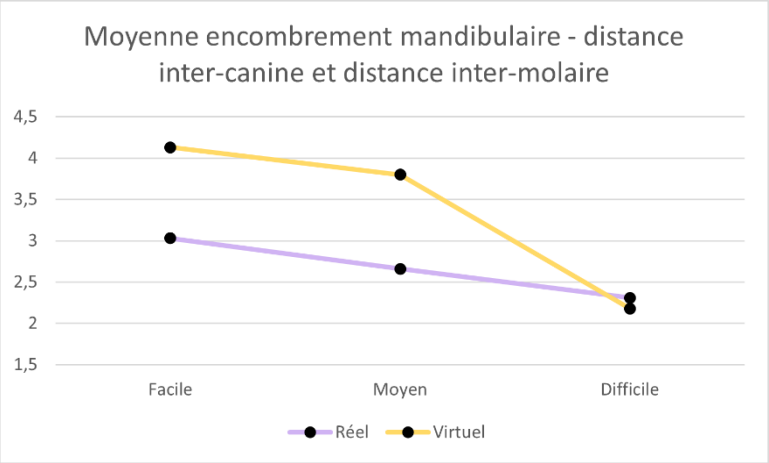


Figure 34 : Représentation graphique de la moyenne des estimations de l'encombrement mandibulaire, de la distance intercanine et de la distance intermolaire.

5-4-2-1 Encombrement mandibulaire

Les résultats indiquent que les distances sont surestimées tant sur les modèles réels et que sur les modèls virtuels, mais que cette surestimation est plus importante sur les modèles virtuels. Par ailleurs, plus la complexité des cas augmente, moins cette surestimation est importante, ce qui signifie que la performance augmente avec la complexité.

L'analyse de variances a montré un effet principal du type de présentation (réel – vs. virtuel) [F (1, 49) = 7.662, p < .05], ainsi qu'un effet principal des niveaux de complexité [F (2, 98) = 5.139, p < .05]. En revanche, nous n'avons pas observé d'interaction entre le type de présentation et le niveau de complexité [F (2, 98) = 0.951, p > .05].

Les statistiques descriptives (Tableau 7) ont montré que l'encombrement mandibulaire était surestimé tant en mode réel qu'en mode virtuel, mais que la surestimation était plus importante en mode virtuel (M = 2.77) (M = 1.19). Pour les niveaux de complexité, la surestimation était moins importante en condition « difficile » (M = 1.90) que dans les conditions « facile » (M = 2.72) et « moyenne » (M = 2.82) (voir aussi figure 35).

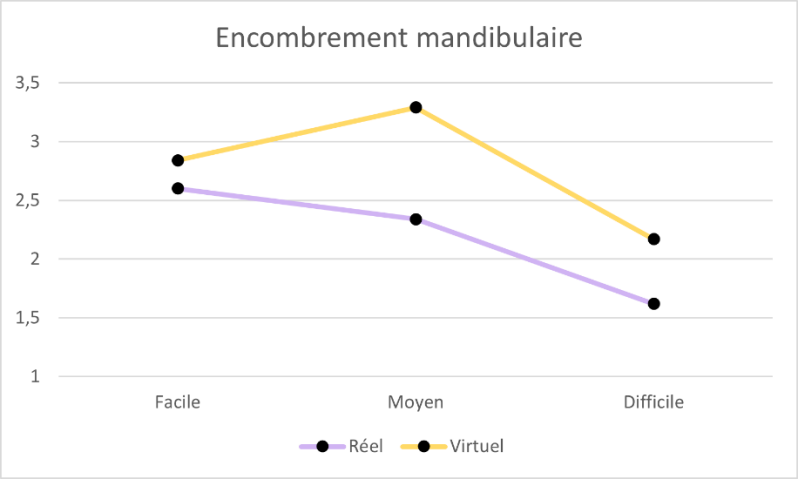


Figure 35 : Représentation graphique de l'encombrement mandibulaire.

Les résultats indiquent que la performance dans l'estimation de l'encombrement mandibulaire est significativement plus élevée sur les modèles réels que sur les modèles virtuels. Ils indiquent également que la performance augmente avec la difficulté des cas, et ce quel que soit le support utilisé.

5-3-2-2 Distance intercanine

L'analyse de variances a montré un effet principal du type de présentation (réel – vs. virtuel) [F (1, 49) = 6.053, p < .05], un effet principal du niveau de complexité [F (2, 98) = 4.078, p < .05], ainsi qu'une interaction entre le type de présentation et le niveau de complexité, [F (2, 98) = 3.154, p < .05].

Les statistiques descriptives (Tableau 7) ont montré que la distance intercanine était surestimée tant en mode réel qu'en mode virtuel, mais que la surestimation était plus importante en mode virtuel (M = 3.63) (M = 2.56 en mode réel). Pour les niveaux de complexité, la surestimation est la plus importante en condition « facile » (M = 4.03), puis diminue en condition « moyenne » (M = 2.76) et atteint son niveau minimal en condition « difficile » (M = 2.36). Enfin, l'interaction était due au fait que la différence entre mode virtuel et mode réel (surestimation plus importante en mode virtuel) était présente en condition « facile » et « moyenne », mais pas en condition « difficile ».

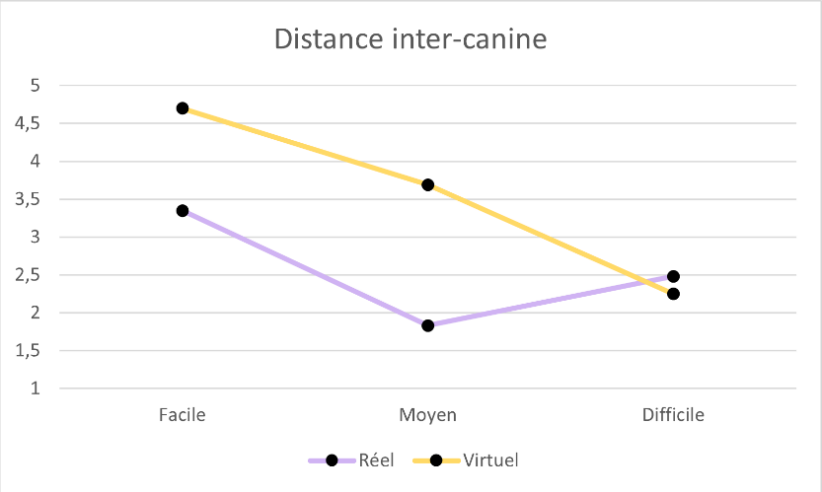


Figure 36 : Représentation graphique de la distance intercanine.

Les résultats indiquent que la distance intercanine était surestimée tant en mode réel qu'en mode virtuel, mais que la surestimation était plus importante en virtuel à l'exception des cas difficiles, où cette surestimation était légèrement supérieure en mode réel

5-3-2-3 Distance intermolaire

L'analyse de variance n'a pas montré d'effet principal du type de présentation (réel –vs. virtuel) [F (1, 49) = 0.441, p > .05], de niveau de complexité [F (2, 98) = 1.740, p > .05], et pas d'interaction entre le type de présentation et le niveau de complexité [F (2, 98) = 1.660, p > .05].

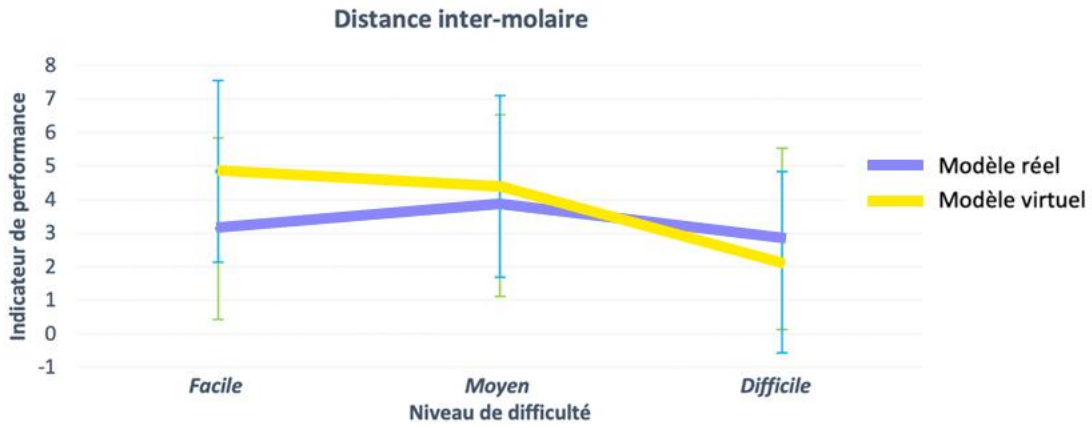


Figure 37 : Représentation graphique de la performance lors de l'estimation de la distance intermolaire.

	Encombrement mandibulaire						Distance inter-canine						Distance inter-molaire					
	Facile		Moyen		Difficile		Facile		Moyen		Difficile		Facile		Moyen		Difficile	
	M	σ	M	σ	M	σ	M	σ	M	σ	M	σ	M	σ	M	σ	M	σ
Réel	2.61	2.18	2.34	2.76	1.62	4.16	3.35	8.75	1.83	7.93	2.48	8.76	3.13	9.79	3.82	9.03	2.83	11.12
Virtuel	2.84	2.54	3.29	2.95	2.17	4.21	4.71	9.08	3.69	8.88	2.25	8.27	4.84	10.67	4.39	8.10	2.13	9.32

Tableau 7 : Statistiques descriptives de l'indicateur de l'encombrement, de la distance intercanine et de la distance intermolaire.

5-4 Discussion

Les analyses de variances réalisées sur les moyennes des trois estimations ont montré que les distances étaient surestimées tant en mode réel qu'en mode virtuel. Ces résultats concordent avec ceux obtenus par d'autres chercheurs. Plusieurs études ont confirmé une surestimation de la distance de l'encombrement mandibulaire lorsque l'évaluation avait lieu au moyen de la visualisation de modèles réels par des orthodontistes (Johal et Battagel, 1997 ; Naish et al., 2016 ; Wallis et al., 2014). D'autre part, une expérience menée par Makaremi et N'Kaoua (2021) au sein d'un groupe de 31 orthodontistes a montré une surestimation des distances au niveau de l'encombrement mandibulaire sur les modèles virtuels.

De plus, les résultats que nous avons obtenus montrent que les distances étaient moins bien perçues avec les modèles virtuels qu'avec les modèles en plâtre. En effet, les distances étaient surestimées dans la condition de RV comparée à la condition réelle.

Une explication possible de ces résultats réside dans l'effet des indices de profondeur non picturaux sur l'estimation de la distance. Ces indices correspondent, entre autres, aux informations provenant du système oculomoteur binoculaire humain permettant de percevoir la profondeur. Ainsi, la disparité binoculaire, l'accommodation et la vergence ont une grande importance dans la perception et dans l'estimation de la distance notamment dans l'espace personnel s'étendant sur 2 mètres de l'observateur (Cutting et Vishton, 1995). Les modèles en plâtre permettent un accès naturel aux différents indices de profondeur non picturaux tandis que, dans le cas des modèles

virtuels, présentés ici à l'aide d'un écran d'ordinateur, ces indices peuvent être absents ou contradictoires. La disparité binoculaire renvoie à la différence des images perçues par l'œil droit et par l'œil gauche (Cutting et Vishton, 1995). Wainman et al. (2018) ont montré, dans une tâche nécessitant une compréhension spatiale de l'anatomie du pelvis féminin, l'importance de cet indice dans la différence de performance entre modèle physique et modèle virtuel. Dans cette étude, les participants ont eu l'œil non-dominant caché (éliminant la vision binoculaire) ou non lors de l'estimation de la distance sur des modèles virtuels ou physiques. Lorsque la vision binoculaire était disponible, le modèle physique a permis une meilleure performance que le modèle virtuel. Cependant, sans vision binoculaire, les performances sur le modèle physique ont diminué, tandis qu'elles n'ont pas changé pour le modèle virtuel (pas de différence significative entre modèle physique et modèle virtuel dans la condition sans vision binoculaire). Ces résultats indiquent que la performance dans l'analyse spatiale est diminuée par de mauvaise informations fournies par la vision binoculaire sur des modèles virtuels par rapport à des modèles physiques. De plus, en situation virtuelle, l'accommodation s'effectue sur le plan constitué par l'écran de l'ordinateur plutôt que sur les objets virtuels eux-mêmes ce qui diminuerait l'acuité visuel en profondeur (Huckauf, 2005). A contrario, le point de convergence dépend de la parallaxe des objets présentés (Bruder et al., 2016 ; Renner et al., 2013). Cela peut donc donner lieu à un conflit entre l'accommodation et le point de convergence entraînant de la fatigue oculaire ou des maux de tête sur le long terme (Huckauf, 2005 ; Renner et al., 2013). Il a été reporté que lorsque ce conflit a lieu, c'est l'accommodation qui est priorisée en tant qu'indice de profondeur (Swenson, 1932). Par conséquent, la perception des distances en environnement virtuel serait mal orientée par l'accommodation (Bruder et al., 2016 ; Drascic et Milgram, 1996 ; Renner et al., 2013).

D'autres facteurs, comme le support de l'apprentissage utilisé par les participants (modèle virtuel ou modèle réel) ou leurs capacités visuospatiales, peuvent aussi influer sur la performance. En effet, il a été montré que de faibles capacités visuospatiales entrainent une baisse de la performance lors de l'étude de modèles virtuels, tandis que les capacités visuospatiales, fortes ou faibles, n'impactent pas la performance lors de l'étude de modèles physiques (Wainman et al., 2020). Par ailleurs, Wainman et al. (2018) ont analysé l'effet du support de l'apprentissage dans le contexte d'une tâche nécessitant une analyse spatiale du pelvis féminin sur un écran d'ordinateur. Ces auteurs ont alors montré que l'utilisation d'un modèle physique menait à de meilleures performances que l'utilisation d'un modèle virtuel. L'effet des supports utilisés pendant la

formation des praticiens (modèles réels, modèles virtuels ou les deux) et des capacités visuospatiales de ces derniers mériterait aussi de faire l'objet d'études.

Enfin, une dernière explication possible est que les tâches réalisées sur un modèle virtuel entrainent une charge mentale supérieure à celle observée lors des tâches réalisées sur un modèle réel. En effet, une hausse de la charge mentale est associée à une diminution de la performance (Jaquess et al., 2017 ; Chanquoy et al., 2007). Or, selon les résultats de notre étude, la charge mentale était plus élevée lorsque les estimations étaient réalisées sur un écran d'ordinateur que lorsqu'elles étaient réalisées sur des modèles en plâtre. Ce résultat est en accord avec des études ayant montré une hausse de la charge mentale dans des tâches réalisées en milieu virtuel (Koumaditis et al., 2020 ; Stevens et al., 2016). En effet, plusieurs auteurs affirment que l'interaction avec un environnement virtuel, du fait de sa nature complexe, pourrait augmenter la charge cognitive (Fraser et al., 2012 ; Huang et al., 2020). Koumaditis et al. (2020) ont par ailleurs observé que la hausse de la charge mentale était associée à une baisse de la performance dans une tâche de résolution de puzzle en 3D réalisée en RV.

Dans le cadre de notre étude, il est donc fort probable que, pour compenser les indices visuels non-picturaux conflictuels dans la condition avec modèles virtuels, les praticiens aient eu à s'appuyer sur leur compétences visuospatiales et/ou d'autres types de traitement cognitif pour appréhender le modèle virtuel et estimer les distances, entrainant une hausse de la charge mentale et une diminution de la performance. Il s'agit là aussi d'une hypothèse qu'il conviendrait de tester dans des travaux ultérieurs.

Concernant l'effet de la complexité des modèles, nos résultats indiquent que plus le niveau de complexité augmente, plus la charge mentale augmente, que ce soit pour les modèles en plâtre ou pour les modèles virtuels. Ces résultats sont en accord avec la littérature montrant qu'une charge mentale élevée est associée à une complexité élevée de la tâche (Chao et al., 2017 ; Das et al., 2020 ; Jacko et Ward, 1996 ; Li et Kim, 2021). Par exemple, Haji et al. (2015) ont montré que la charge mentale était plus élevée dans une tâche complexe que dans une tâche simple lors d'une simulation physique de nœud pour fixer un drain thoracique. Di Stasi et al. (2016) ont, quant à eux, trouvé que la charge mentale augmentait avec le niveau de complexité de la tâche dans différents exercices de laparoscopie réalisés en RV. Notre étude montre que c'est aussi le cas dans des tâches d'estimation de la distance réalisées en mode réel ou virtuel, où l'augmentation de la complexité des modèles augmente la charge mentale.

Dans notre étude, la complexité des modèles a également eu un impact sur l'estimation des distances. En effet, nos résultats montrent que l'augmentation de la complexité des modèles (virtuels et réels) était associée à de meilleures estimations des distances pour l'encombrement mandibulaire et la distance intercanine tandis qu'aucune différence d'estimation selon la complexité de la tâche n'a été observée pour la distance intermolaire. Un tel résultat pourrait alors s'expliquer par les enjeux de l'estimation des distances réalisée dans notre étude. En effet, l'encombrement mandibulaire et la distance intercanine jouent un rôle important dans le diagnostic et, par conséquent, dans le traitement qui sera mis en place, tandis que la distance intermolaire est une condition qui ne se présente pas dans la pratique professionnelle. Des auteurs ont démontré que les objectifs de la tâche médient la relation entre performance et complexité de la tâche. Des objectifs importants dans des tâches complexes peuvent améliorer la performance grâce à des processus motivationnels et cognitifs contribuant au développement de stratégies efficaces (Campbell, 1988 ; Wood et al., 1987). Ainsi, dans notre étude, l'estimation des distances pour des cas complexes peut être associée à des enjeux d'intervention chirurgicale plus graves ou pouvant avoir plus d'impact sur la prise en charge ultérieure. Ces enjeux plus importants lors de l'estimation de modèles complexes pourraient être à l'origine d'une attention accrue pendant l'évaluation des distances sur les modèles réels et virtuels, améliorant ainsi la performance. Bien entendu, une telle hypothèse reste à tester dans des études ultérieures, notamment en termes d'impact de l'enjeu thérapeutique sur l'exactitude des réponses quant aux évaluations en milieu réel et virtuel.

5-5 Conséquences cliniques des évaluations

Cliniquement, le paramètre présentant le plus d'intérêt lors de son évaluation est l'encombrement. En effet, Saghafi et al. (2017) (90) et Baumrind et al. (1996) (91) ont démontré que l'encombrement est le facteur principal de prise de décision d'extraction. Ils ont également défini les cas d'extractions « limites » par un encombrement entre 4 et 8 mm.

Selon les résultats de notre étude (Tableau 5), l'encombrement est surestimé en moyenne de 2,25 mm, soit 2 mm sur le modèle réel et 2,51 mm sur le modèle numérique. Cette surestimation est plus élevée pour les cas « faciles », indépendamment du support utilisé.

Les résultats que nous avons obtenus dans ce cadre peuvent être mis en parallèle avec les résultats d'autres études ayant trait à l'estimation de l'encombrement et à son influence sur la prise de décision.

Dans l'étude menée par Wallis et al. (2014), il a été demandé à 62 orthodontistes d'évaluer la valeur de l'encombrement de 8 modèles mandibulaires réels (sur des typodonts[8]). Ces chercheurs ont observé que les orthodontistes réalisaient des estimations au ½ millimètre près (comme demandé dans notre protocole). Ils se sont rendu compte que les orthodontistes ont renoncé à utiliser des outils de mesure, se contentant de la visualisation du modèle. Une surestimation de l'encombrement a été observée pour 7 modèles sur 8. De plus, les praticiens ont été questionnés sur les choix thérapeutiques à réaliser sur chacun des modèles (sans extraction, cas limite, extraction). De manière logique, plus la valeur de l'encombrement était élevée, plus on avait recours au processus d'extraction était proposé. Au-delà de 4 mm d'encombrement, la solution majoritairement choisie est l'extraction, et aucun orthodontiste ne choisit la solution sans extraction pour les modèles où l'on observe un encombrement supérieur à 5 mm. La valeur de l'encombrement est donc un facteur majeur qui influence la décision d'extraction, qui semble être prise au millimètre près.

Une étude plus ancienne a également été menée par Rheude et al. (2005). Elle comparait déjà l'influence du support de présentation des modèles (numérique ou réel) sur le diagnostic et la prise de décision. Ces chercheurs n'ont pas obtenu de résultats significatifs quant à la prise de décision d'extraction ou de chirurgie. Ils ont toutefois constaté que les procédures biomécaniques de traitement (quantité de stripping prévue, ajout d'un système d'ancrage maxillaire, etc.) variaient de manière significative selon le support de visualisation utilisé (modèle numérique ou modèle en plâtre). Cependant, le changement statistiquement significatif constaté entre modèles en plâtre et modèles numériques ne semble pas être cliniquement pertinent, sauf dans les cas les plus complexes (décompensation pré-chirurgicale, extractions asymétriques etc.), où les modèles en plâtre sont considérés comme plus fiables et plus précis, et donc indispensables.

Naish et al. (2016) se sont demandé si la connaissance de la vraie valeur de l'encombrement exerçait une quelconque influence sur la prise de décision comparée à une simple estimation par

[8] En dentisterie, un typodont est un modèle de la cavité buccale, y compris les dents, la gencive et le palais.

visualisation de modèles en plâtre. L'estimation simple conduit, selon eux, à une surestimation systématique et significative de la valeur de l'encombrement. La prise de décision d'extraction est modifiée de manière significative dès lors que les praticiens connaissent la vraie valeur de l'encombrement, en revanche, la prise de décision de stripping et d'expansion n'est pas modifiée, ce qui s'expliquerait par le fait que ces moyens de traitement sont mis en place dans les cas les plus « simples ». Nous constatons que ces résultats sont en accord avec ceux de notre étude, excepté pour la précision en fonction de la complexité des cas : dans l'étude que nous venons de citer, plus les cas sont complexes et plus la valeur de l'encombrement est élevée, plus l'évaluation manque de précision. Ces chercheurs concluent donc qu'une analyse plus précise de l'encombrement peut s'avérer nécessaire à une meilleure prise de décision.

La connaissance de la vraie valeur de l'encombrement est donc primordiale pour un diagnostic et une prise de décision de qualité, principalement dans les cas « limites » où l'extraction est envisagée. Les différentes études menées dans ce domaine révèlent globalement une surestimation significative de la valeur de l'encombrement. Une meilleure prise de conscience par les praticiens quant à leurs capacités d'estimation s'avère donc nécessaire.

5-6 Conclusion

L'évolution technologique dans le domaine de l'orthodontie contribue au remplacement progressif des modèles réels par des modèles virtuels. Les praticiens n'ayant plus la possibilité de manier les modèles de leurs mains, cette évolution conduit à une perte de l'apport haptique. Cette perte est l'une des conséquences du changement de paradigme généré par la transition numérique et digitale en orthopédie dento-faciale.

Les études présentées dans ce chapitre ont mis en lumière certaines conséquences pour le praticien :

- Une différence de précision lors de l'estimation des distances sur les modèles dentaires selon l'environnement considéré (réel ou virtuel) ;
- Une précision plus élevée de la part des praticiens lorsque l'estimation est effectuée sur des modèles réels ;

L'utilisation des outils numériques à des fins de diagnostic ou de traitement sera amenée à se développer dans les années à venir. Il est donc indispensable de mieux comprendre l'impact des

différentes variables (entre autres : expériences professionnelles, objectifs, matériel, type de modèle) sur les tâches nécessitant l'estimation de distances.

En effet, une perception des données paracliniques du patient sur un support numérique par le praticien au plus près de la réalité clinique est indispensable. Le changement de paradigme dû à l'apport des nouvelles technologies offre aux praticiens des opportunités nouvelles de perception.
L'enjeu de notre travail d'analyse et de réflexion sur la transition numérique est de proposer des pistes vers une stratégie d'augmentation, à travers laquelle les praticiens deviendront plus performants avec l'utilisation de l'outil numérique.

5-7 Essentiel à retenir

La charge mentale de l'estimation des distances sur des modèles orthodontiques est significativement plus élevée lorsque l'opération est réalisée sur des modèles virtuels que sur des modèles réels. De plus, la charge mentale augmente avec la difficulté des cas estimés, et ce quel que soit le type de support utilisé.
Les distances sont surestimées sur les modèles réels et virtuels, mais cette surestimation est plus importante de façon statistiquement significative sur les modèles virtuels. La performance de l'estimation des distances est moins bonne sur les modèles virtuels.
Par ailleurs, plus la complexité des cas augmente, moins cette surestimation est importante, ce qui signifie que la performance augmente avec la complexité de la tâche.
Il existe une différence significative en fonction du type de présentation (réel ou virtuel) : l'évaluation de l'encombrement et de la distance intercanine est moins satisfaisante dans le cas des modèles virtuels, l'évaluation de la distance intermolaire s'avérant, quant à elle, beaucoup moins dépendante du support (réel ou virtuel).
Une des sources d'explication possibles de ces résultats repose sur l'effet des différents indices de profondeur non-picturaux sur l'estimation des distances avec une perception des distances des modèles virtuels qui serait mal orientée par l'accommodation. D'autres facteurs, comme le type d'apprentissage réalisé par les participants, leurs capacités visuospatiales, ou une plus grande sollicitation de charge mentale, peuvent aussi être impliqués dans la baisse de la performance observée dans le cas des modèles virtuels.

La surestimation de l'encombrement des arcades à partir de modèles virtuels a de véritables conséquences cliniques, notamment lorsque le praticien est amené à décider de la pertinence de l'avulsion des prémolaires dans le cadre de traitements orthodontiques.

6 : Apport de l'interprétabilité des réseaux de neurones convolutionnels (CNN) à une meilleure perception des dysmorphoses par l'orthodontiste

6-1 Vision par ordinateur et réseaux neuronaux convolutionnels (CNN)

L'intelligence artificielle, et particulièrement la vision par ordinateur, fonctionne comme un troisième œil pour le praticien (Fourcade et al., 2019), permettant d'élargir et d'améliorer sa perception des données paracliniques. Mais la question se pose de savoir ce que voit précisément ce troisième œil.

L'interprétabilité des réseaux de neurones convolutifs permet d'apporter une réponse à cette question. En effet, les zones d'une image où le réseau de neurones prête attention pour la discriminer peuvent constituer une source d'information précieuse.

Ce chapitre après avoir exposé le fonctionnement des CNN et leur utilisation en médecine, explore la manière dont l'interprétabilité des réseaux de neurones peut contribuer à étendre le champ de perception de l'orthodontiste.

6-1-1 Définition

La vision par ordinateur est une discipline visant à élaborer des algorithmes capables d'acquérir une compréhension de haut niveau à partir d'images (Sonka et al., 2014).

Ce domaine a émergé dans les années 1960, lorsque les informaticiens ont commencé à intégrer dans l'algorithme des outils de détection de bords, de lignes, de courbes et d'autres structures remarquables pour aider l'algorithme à extraire et à identifier l'information souhaitée.

La vision par ordinateur a connu un essor remarquable ces dernières années. A travers les réseaux neuronaux convolutionnels (CNN), elle dépasse désormais les capacités humaines dans les tâches de décryptage de l'information visuelle. Cette capacité de la vision par ordinateur repose sur l'extraction de caractéristiques locales qui sont combinées pour former des caractéristiques de niveau supérieur. Cette technologie permet des applications importantes en matière de classification en imagerie médicale. Les CNN se sont avérés plus performants que les praticiens pour ce qui est de la classification des pathologies (Liu et al., 2019). La performance des CNN frôle désormais le taux de Bayes (Lin et al., 2014), un seuil absolu qui limite les capacités d'apprentissage tant des humains que des machines, en matière d'extraction d'information. Dès lors, les CNN

deviennent un outil précieux pour l'obtention de nouvelles informations sur différentes pathologies.

Dans un même temps, des techniques dites d'interprétabilité ont été développées. Elles consistent à fournir une information représentant à la fois le mécanisme interne de l'algorithme de Machine Learning et la représentation interne des données (Gilpin et al., 2018) En d'autres termes, l'interprétabilité consiste à transformer le flux digital interne du réseau neuronal en un format intelligible pour l'Homme. En imagerie, un exemple d'interprétabilité des réseaux neuronaux appliqué aux CNN sont les cartes de saillance qui permettent de restituer visuellement les caractéristiques locales détectées par les CNN, et qui ont permis au CNN de prendre sa décision de classification (par exemple en image physiologique vs. Pathologique).

6-1-2 Fonctionnement

6-1-2-1 Les neurones artificiels

Les neurones artificiels, aussi appelé neurones formels, sont des objets mathématiques calquant les propriétés des neurones biologiques. Ce sont des unités qui effectuent des calculs élémentaires (Guresen et Kayakutlu, 2004) (Figure 38) en mimant la diffusion de stimuli nerveux dans les neurones biologiques : ces derniers captent des influx afférents sur leur dendrites, puis effectuent leur sommation au niveau du soma. Si ce résultat dépasse une valeur seuil « θ » du neurone biologique, alors ce dernier va décharger et ainsi diffuser le résultat aux autres neurones biologiques via les synapses.

Le neurone artificiel reçoit des entrées notées « x » qui vont être traitées selon une fonction d'agrégation préalablement choisit, avant d'être soumis à une fonction d'activation : Heaviside, Tangeante, Sigmoïde. Dans le cas d'une fonction d'activation Heaviside, si la somme dépasse le seuil θ une sortie est émise. Pour les fonctions tangentes et sigmoïdes, c'est l'application des formules mathématiques sur le résultat de la fonction d'agrégation qui fournit un résultat.

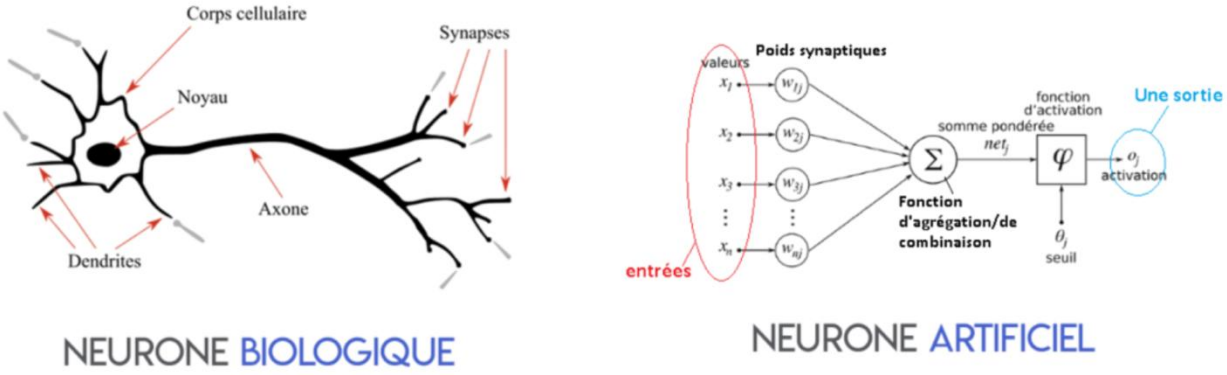


Figure 38 : Schéma comparatif : neurone biologique et neurone artificiel, d'après Maurice (2018).

Dans les systèmes de vision par ordinateur, les neurones artificiels sont organisés en réseaux neuronaux convolutifs ou réseaux neuronaux à convolution (en anglais CNN ou ConvNet pour Convolutional Neural Networks), un type de réseau de neurones artificiels acycliques (feed-forward) dans lequel le motif de connexion entre les neurones est inspiré par le cortex visuel des primates. Les neurones de cette région du cerveau sont arrangés de sorte à correspondre à des régions qui se chevauchent lors du pavage du champ visuel. Le fonctionnement des CNN est inspiré par les processus biologiques : ils consistent en un empilage multicouche, dont le but est de prétraiter de petites quantités d'informations. Ainsi un réseau de neurones convolutif est construit en empilant les couches : la sortie d'une couche correspond à l'entrée de la suivante. Les poids des couches sont déterminés par rétropropagation du gradient[9]. Dans le cas des CNN, ces paramètres désignent en particulier les caractéristiques (features) des images.

[9] La technique de rétropropagation du gradient (*Backpropagation* en anglais) est une méthode qui permet de calculer le gradient de l'erreur pour chaque neurone du réseau, de la dernière couche vers la première et permet un apprentissage.

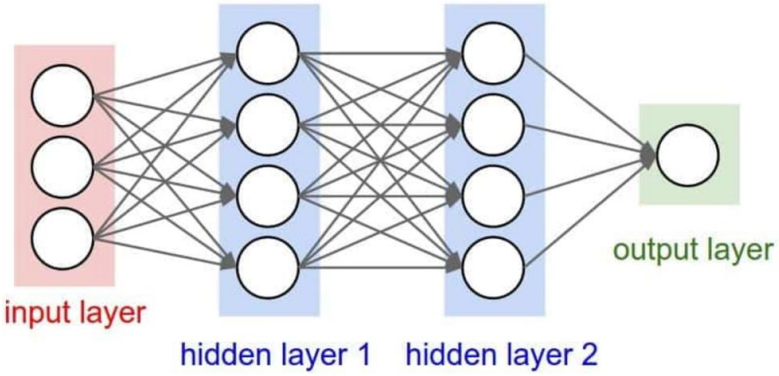


Figure 39 : Schéma d'une organisation convolutive.

Chaque rond blanc représente un neurone artificiel. La première couche en rouge correspond à l'image d'entrée. Elle aura autant de neurones que les pixels présents dans l'image. En bleu sont représentés deux couches convolutives. Ces couches sont dites « cachées ». Ce sont des couches où l'information est constituée d'un flux numérique abstrait. Plus le nombre de couche cachées est important, plus un réseau de neurones est dit profond. C'est de là que vient l'appellation « Deep learning » pour « réseaux neuronaux profonds ».

6-1-2.2 Apprentissage et classification

Les réseaux de neurones apprennent à partir de grands volumes de données. Ils fonctionnent en deux temps.

Une première phase est celle de l'apprentissage ou « training ». Elle nécessite une base de données contenant des exemples (images, par exemple). Et chaque image est labellisée du nom de la classe à laquelle elle appartient. Tous ces exemples vont être appris par le réseau. Le réseau va donc, d'un côté, voir une image, et de l'autre côté, savoir à quelle classe ladite image appartient. Le CNN extrait les informations visuelles caractéristiques lui permettant d'identifier correctement la classe et se « forge » une représentation latente de ces caractéristiques dans son réseau. Une fois l'apprentissage considéré satisfaisant l'apprentissage s'arrête et l'architecture du réseau avec ses poids mis à jour est sauvegardée. Le réseau peut être considéré alors comme « verrouillé », c'est-à-dire qu'il a acquis la connaissance préalable qui lui permettra par la suite, de détecter ces éléments

caractéristiques sur de nouvelles images et réaliser la bonne prédiction. C'est le principe de l'apprentissage automatique : le modèle se construit lui-même à partir d'exemples, sans avoir d'indication sur ce qu'il doit extraire des données.
La seconde phase est celle du test, qui consiste à utiliser le réseau pour prédire les classes. Il suffit désormais de présenter une image. A l'aide des caractéristiques apprises, le réseau peut identifier l'image et fournir une prédiction sur la classe à laquelle appartient l'image. C'est la tâche de classification.

6-1-2 .3 CNN et représentation interne

Pour extraire les caractéristiques de l'image, un CNN alterne majoritairement deux types de couches :

- Couches avec filtre convolutionnel : ces couches réalisent un produit entre un filtre (un tableau à deux dimensions dont les poids seront ajustés lors de l'apprentissage) et l'entrée de la couche. Les dimensions de ce filtre sont généralement 3×3 ou 5×5. La multiplication de ces couches au sein du réseau va permettre d'extraire des caractéristiques de plus en plus complexes qui permettront in fine de prédire une classe d'appartenance pour l'item présent dans l'image. C'est pour cela que l'on parle « d'apprentissage profond ».
- Couches avec pooling : ces couches permettent un sous-échantillonnage qui va compresser la dimension de l'image et réduire le coût computationnel des couches suivantes.

Lors de la phase d'apprentissage, les poids des différents filtres convolutionnels s'ajustent de sorte que le réseau converge vers la détection d'éléments au sein de l'image, qui permettront de prédire in fine la bonne catégorie (LeNet-5). Par exemple, la figure 40 ci-dessous montre le processus d'extraction de caractéristiques locales et la formation de caractéristiques de niveau supérieur pour la reconnaissance des visages. La première couche n'est capable de détecter que des formes générales, comme les arcs de cercles et les lignes. La deuxième couche utilise ce que détecte la couche précédente pour reconnaitre des formes plus précises, comme un nez, un œil ou une bouche. Enfin, la dernière couche combine les sorties précédentes pour reconnaitre des visages.

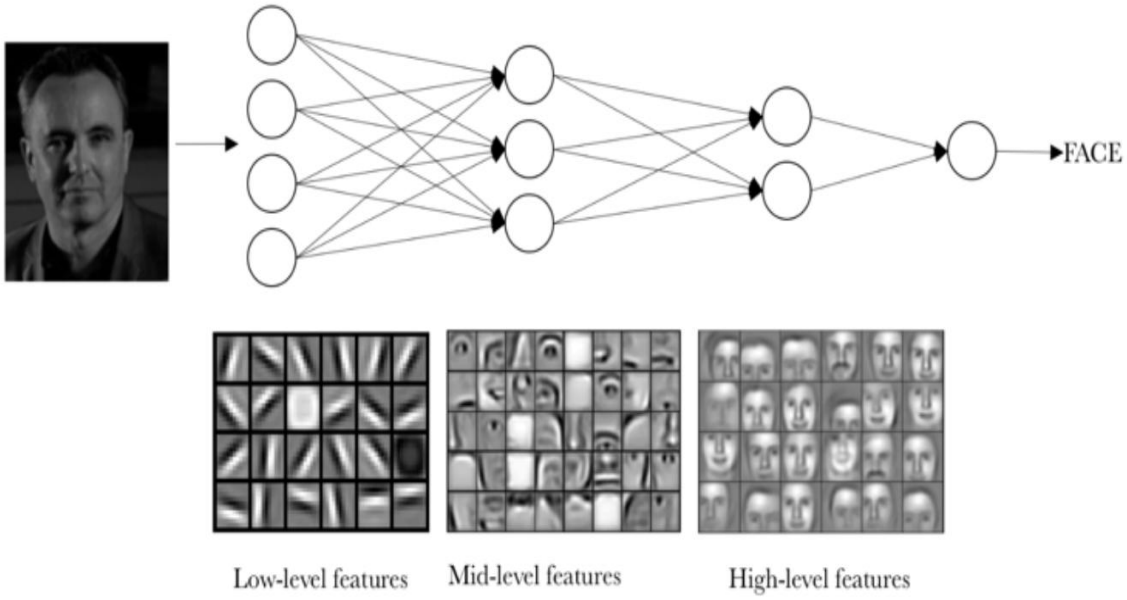


Figure 40 : Processus d'extraction des caractéristiques locales et supérieures d'après Jones (2014).

6-1-2.4 «Black Box Problem » et interprétabilité

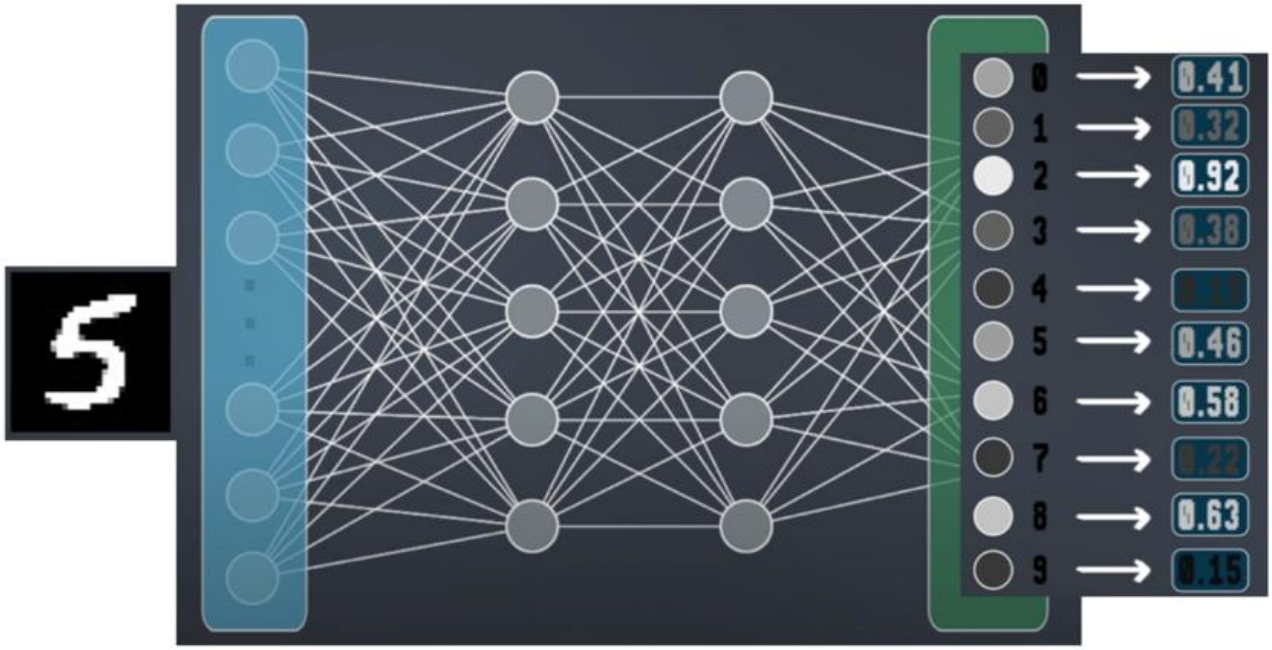


Figure 41 : Schéma d'un réseau de neurones de reconnaissance de chiffres, inspiré de Hoang (2018).

Le « Black Box Problem » de Sudher et Jain (2004) est la formalisation du problème de transparence des raisons qui ont conduit le réseau à prendre telle ou telle décision. En effet, le flux interne de

données numériques d'un réseau de neurones étant dépourvus de sens pour l'humain, cela rend la compréhension du cheminement et de l'analyse des caractéristiques utilisé pour la prise de décision, inaccessible pour un humain non expert du CNN.
La figure 41 présente le cas d'un réseau formé pour la reconnaissance des chiffres. Une image contenant un chiffre est introduite par la couche d'entrée, marquée ici en bleu. Puis, l'information est transmise de couche en couche jusqu'à la couche de sortie, indiquée ici en vert. Au niveau de cette couche de sortie, le réseau génére, pour chaque classe possible, la probabilité que l'image appartiennent à la classe considérée. La classe présentant la probabilité la plus élevée sera alors sélectionnée.
Le réseau reçoit donc en entrée une image et indique en sortie la classe qu'il identifie.
Il ne fournit aucune explication sur les éléments de l'image l'ayant amené à cette conclusion, mais se limite à une fonction de prédiction à cet égard.
C'est ce manque de transparence qui se situe au cœur de ce qu'on nomme le « Black Box Problem » (Watson et al., 2019).

Dès lors, des outils dits d'interprétabilités ont été élaborés. L'interprétabilité consiste à fournir des informations sur le cheminement interne du réseau dans un format intelligible par l'Homme, Zeiler et Fergus (2013) ont démontré que l'avant dernière couche N-1 (Pre-softmax) du CNN agit comme un champ récepteur (figure 42).

C'est-à-dire, qu'à sa sortie se trouve la décision de classe prédite par le CNN, et à son entrée le champ récepteur contenant les pixels ayants induits cette prise de décision.
Ainsi, Zhou et al. (2016) ont donc modifié cette couche (en la remplaçant par une « Class Activation Mapping » ou (CAM), afin de permettre d'identifier la contribution de chaque pixel à l'identification de la classe.
C'est le point de départ d'une technique d'interprétabilité nommée « carte de saillance » ou « carte d'activation ». Ces cartes permettent d'identifier, sur l'image d'origine, les zones d'intérêt (ROI) détectées par le CNN pour identifier la classe à laquelle appartient l'image.

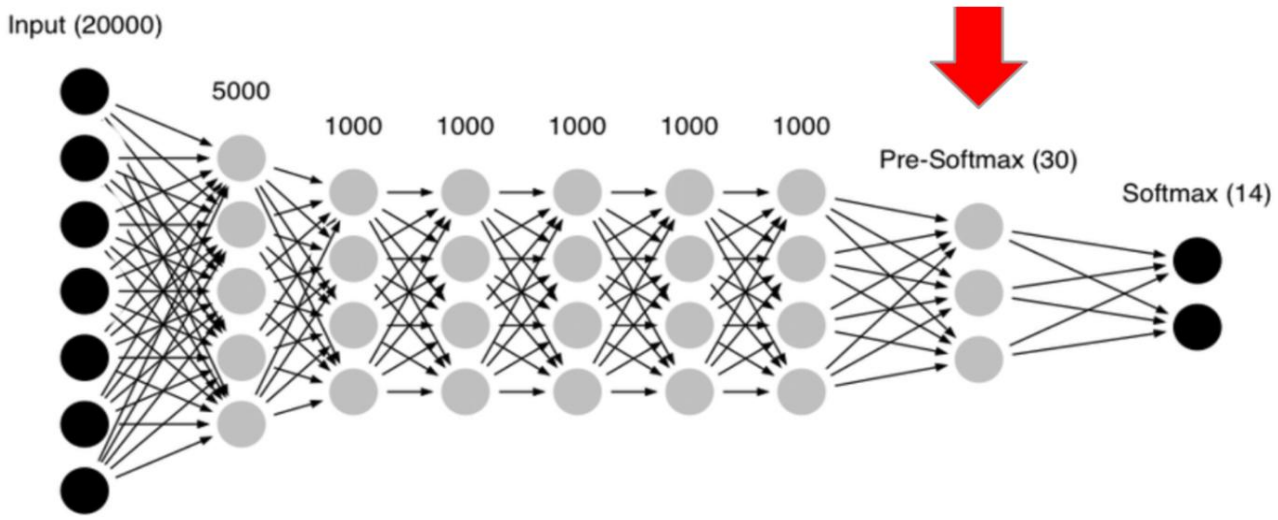


Figure 42 : Couche Pré-Softmax d'un réseau convolutif d'après Suhermi et al. (2018).

Les « carte de saillance » sont une technique d'interprétabilité qui permet d'expliciter sur l'image d'origine via une représentation thermique des pixels, ceux qui ont contribué le plus à la prédiction. La figure 43 ci-dessous présente la carte de saillance des pixels ayant permis d'identifier des activités comme le brossage des dents ou l'élagage d'arbre (Zhou et al. 2016). Les pixels ayant contribué le plus à l'identification dans ces cas précis sont indiqués par des couleurs chaudes (rouge par exemple).

Sur la base de cette CAM ont été élaborées de nombreuses déclinaisons de techniques visant à rendre compte aussi précisément que possible des régions d'intérêt d'une image (ROI).

Notamment, ces déclinaisons utilisent le gradient afin d'exploiter la CAM et remonter aux pixels. La technique la plus répandue en imagerie médicale est Grad-CAM et sa déclinaison Grad-CAM++ (Selvaraju et al., 2016 ; Chattopadhay et al., 2018). C'est à l'aide de ces méthodes que la grande majorité des études a été réalisée.

Bien que ces techniques soient populaires, récemment, (Wang et al., 2019) ont mis en évidence l'instabilité et la non-reproductibilité du gradient. Démontrant que les méthodes basées sur le gradient ne sont pas adéquates pour l'imagerie médicale, Haofan et al. (2020) ont proposé Score-CAM. Cette technique s'affranchit du gradient et permet de générer des cartes de saillances stables et robustes pour une utilisation médicale.

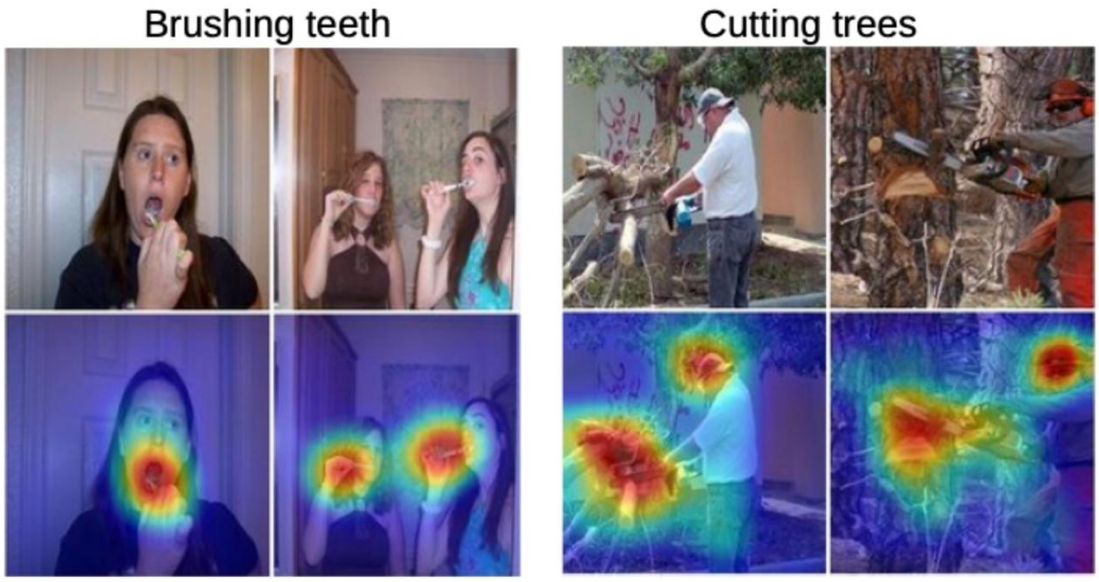


Figure 43 : Exemple de carte de saillance d'après Zhou et al. (2016).

Score-CAM

Pour déterminer les pixels d'importance, la méthode Score-CAM (figure 43) repose sur un principe innovant. La première phase consiste à laisser le CNN visualiser l'image et extraire les caractéristiques de l'image afin de donner sa prédiction.

La seconde phase récupère dans la CAM les caractéristiques détectées par le CNN. Les unes après les autres, ces caractéristiques sont utilisées comme masque sur l'image d'entrée.

L'idée est que si l'on masque les pixels d'une caractéristique importante, la confiance du CNN dans sa prédiction diminue fortement, affectant en conséquence la probabilité d'identification de la classe. De ce fait, l'utilisation du masque permet de considérer que lorsque la confiance dans la prédiction diminue c'est que la zone masquée était importante pour la prise de décision. C'est ainsi que la méthode Score-CAM permet de générer la carte de saillance et par conséquent de rendre compte des caractéristiques utilisées par le réseau dans sa prise de décision.

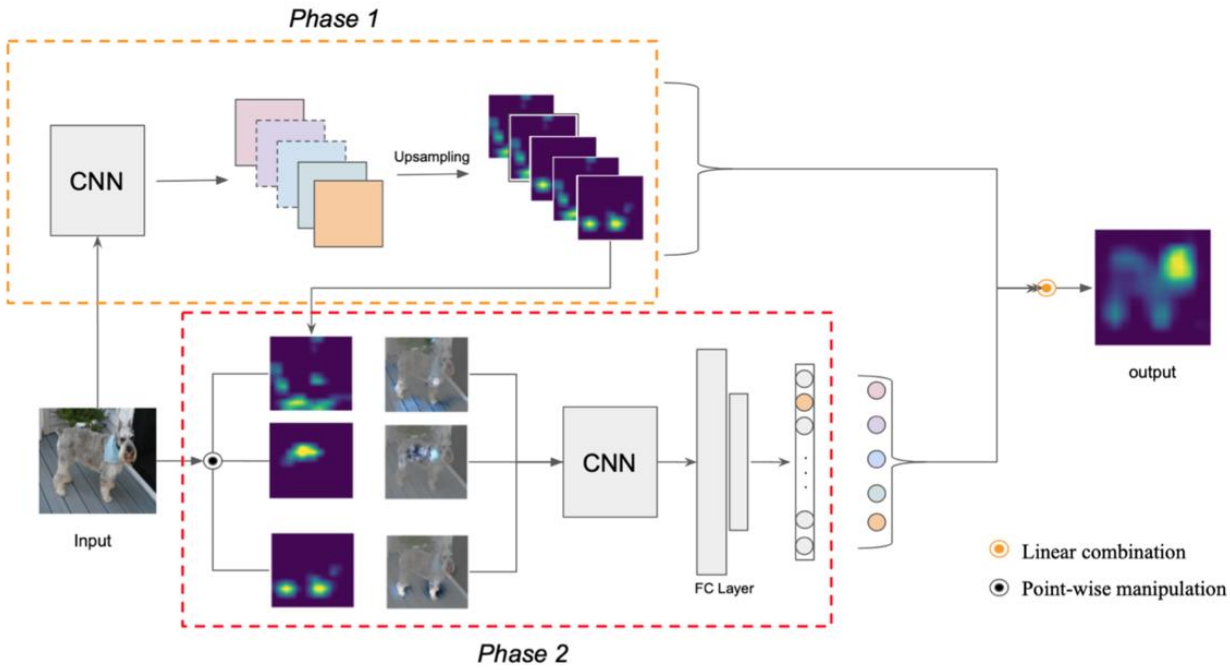


Figure 44 : Schéma du processus de Score-CAM d'après Wang et *al.* (2020).

6-2 Applications en imagerie médicale

6-2-1 Principales applications de la vision par ordinateur en médecine

L'utilisation des procédés de vision par ordinateur basés sur des réseaux de neurones convolutionels constitue la principale utilisation de l'intelligence artificielle en médecine (Jiang et Al, 2017). Plus particulièrement, les CNN connaissent un grand essor dans le système d'aide à la décision clinique (Antoniadi et al., 2021). Dans leur méta-analyse, (Liu et al., 2019) ont constaté une performance des CNN en diagnostique équivalente à celle des professionnels de la santé. Mais Aggarwal et al. (2021) contrastent ces propos et mettent en avant une grande hétérogénéité des performances en fonction des domaines. A l'aide des CNN, des systèmes experts ont été mis au point (Badea, 2016) pour assurer la détection et la classification des images.

Détection : La détection consiste à former les réseaux à reconnaitre certaines caractéristiques sur les images d'apprentissage. Le réseau peut par la suite identifier les images dans lesquelles les caractéristiques sont présentes. En imagerie médicale, les CNN de détection d'objets sont

principalement utilisés pour prioriser les images présentant un intérêt pour le praticien. En orthodontie, la détection d'objets est principalement utilisée pour la détection de repères en céphalométrie (Qian et al., 2019).

Classification : C'est la principale fonction des CNN en imagerie médicale (Nahid et Kong, 2017). Au cours des années 2017 et 2018, les CNN ont connu de multiples cas d'application, pour lesquelles on a pu observer qu'ils égalaient l'expert humain quant à la classification. Plus récemment, les CNN ont permis non seulement d'égaler mais aussi de dépasser la capacité de l'expert humain et frôlent désormais le taux de bayes, autrement dit l'erreur irréductible (Tumer et Ghosh, 1996). Les CNN ont connu leur premier succès dans le domaine de la classification des maladies de la peau (Shanthi et al., 2020) (Ismaël et al., 2021). Les principaux emplois en imagerie par rayon X portent sur les infections pulmonaires (Sirazitdinov, 2019) ou le cancer (Takemiya, 2019).

L'utilisation des CNN à des fins diagnostiques et de localisation sur imagerie par rayon-X a été particulièrement prolifique dans la détection du cancer du sein (Al-Antari et al., 2020), des cancers du poumon (Gu et al., 2021 ; Mastouri et al., 2020) et des infections respiratoires (Teixeira et al., 2021). En outre, elle a fait l'objet de nombreuses études pourtant sur la COVID19 : plus de 200 depuis 2020 (Alyasseri et al., 2021).
Des applications pour les affections osseuses via imagerie par rayon X sont également effectives, comme la détection d'arthrose du genoux (Bany et Yeasin, 2021), la détermination de l'âge (Karargyris et al., 2019) ou encore la détection de certaines fractures.
Concernant le crâne, de multiples applications pour la détection et la localisation de tumeurs de la mâchoire sur radiographie panoramique ont été réalisées (Poedjiastoeti et Suebnukarn, 2018). Les CNN s'avèrent également utiles en matière de détection de l'ostéoporose (Lee et al., 2020).

6-2-2 Apport de l'interprétabilité des réseaux neuronaux convolutionnels dans le domaine médical

L'interprétabilité des réseaux de neurones convolutionnels a été décrite comme une véritable opportunité d'ouvrir de nouveaux horizons en imagerie médicale (Reyes et al., 2020).
L'IA interprétable permet en effet de décliner des sous-types pathologiques et donc contribue de manière décisive à un diagnostic, un pronostic ou encore un plan de traitement plus adaptés (Guchet, 2014). Les réseaux neuronaux permettent de tenir compte de la singularité biologique de

chaque individu. Cette meilleure visualisation rend possible une meilleure compréhension des différentes pathologies. Le défi est donc de développer les outils d'interpétabilité les plus efficients pour leur adoption par les praticiens. Reyes et al. (2020) montrent que les cartes de saillances sont des outils bien admis par les praticiens. En parallèle, Pirovano et al. (2021) ont montré qu'une clé majeure d'adoption de l'intelligence artificielle par les praticiens n'est pas seulement la confiance envers l'outil mais également l'utilité concrète de celui-ci. Kermany et al. (2018) exposent comment l'interprétabilité (notamment les cartes de saillance) permettent d'étendre l'expertise du praticien en augmentant l'information sur l'image en mettant en valeur des éléments d'intérêt. Ils soulignent que l'interprétabilité soit efficace elle doit être en accord avec les processus cognitifs de reconnaissance des formes pour que le praticien intègre de manière pertinente ces informations (figure 45).

Dans des études ayant trait à l'imagerie par rayon X du bras, England et al. (2018) cherchaient à identifier les éléments sur lesquels se basent le CNN pour détecter des fractures osseuses du coude. Le coude étant naturellement pourvu de soudure, il est très difficile de détecter les fissures. La carte de saillance a permis aux chercheurs de voir que, dans le cas de petites fractures, les effusions sanguines dans les tissus mous étaient plus informatives que la recherche directe de la fissure sur l'os. Dans un autre registre, Baek et al. (2019) ont travaillé sur un modèle pour la détection du cancer du poumon à petites cellules. Grace à la carte de saillance, ils ont constaté que le CNN détectait de riches ensembles d'informations structurales ayant une valeur de pronostic remarquable. Ces auteurs ont pu par la suite relier ces éléments à l'activité d'angiogenèse préalable à l'émergence des tumeurs.

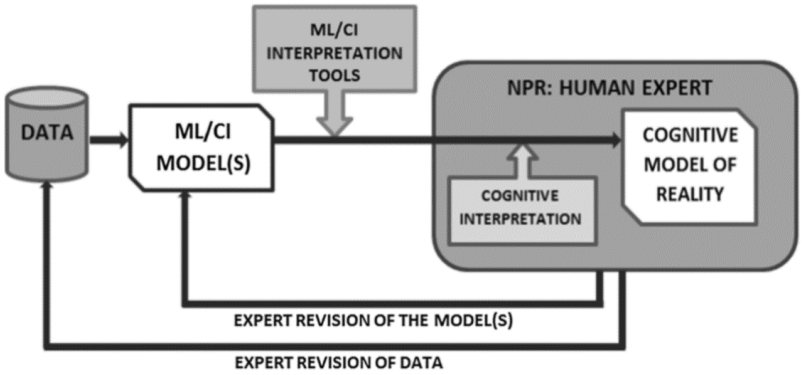


Figure 45 : Interprétabilité et cognition d'après Kermany et al. (2018).

L’emploi des CNN interprétables en imagerie médicales est particulièrement avancé et est courant dans les trois domaines suivants :

- L’histologie, où les CNN ont été rapidement développés et adoptés par les praticiens comme aide au diagnostic, vu le volume considérable de données disponibles (Kuntz et al., 2021 ; Ker et al., 2019). A l’aide des cartes de saillances, les CNN sont utilisés pour mettre en évidence les zones d’intérêt (ROI) et sélectionner les planches d’intérêt ;
- La dermatologie, où les CNN ont démontré une capacité supérieure à celles des dermatologues pour la détection de mélanomes et autres types de cancer de la peau (Thomas et al., 2021 ; Gareau et al., 2020 ; Jaworek-Korjakowska et al., 2021).
- L’analyse de fond de l’œil pour les rétinopathies, où le CNN permet la détection, la segmentation et la classification des lésions dans les images du fond d’œil pour déceler plus précocement l’émergence de maladies (Asiriet al., 2019).

6-3 Application des CNN en orthopédie dento-faciale (ODF)

En orthopédie dento-faciale (ODF), divers canaux d’acquisition d’images sont utilisés : les radiographies bidimensionnelles (panoramiques, téléradiographie de profil), la tomographie tridimensionnelle à faisceau conique (CBCT) et les scanners laser intra-oraux.

Les principales applications en ODF portent sur l'identification des points de repère, la classification squelettique, la prédiction de l'âge osseux et la segmentation des dents (Liu et al., 2021). Mohammad-Rahimi et al. (2021) montrent que l'IA a obtenu des résultats de pointe dans diverses applications orthodontiques, notamment la détection automatisée de points de repères sur les céphalogrammes latéraux et les images photographiques, la détermination du degré de maturation des vertèbres cervicales (Makaremi et al., 2019), la classification squelettique, les décisions d'extraction dentaire orthodontique, la prévision du besoin d'un traitement orthodontique ou de chirurgie orthognathique. La Płotka et al. (2021) retrace également les différentes applications des CNN en orthodontie.

Certaines de ces applications font d’ores et déjà l’objet d’une commercialisation (Tableau 8). En France, la société Kitview ainsi que la société EzOrtho développent des solutions pour les tracés céphalométriques. Le logiciel Planmeca Romexis® Cephalometric Analysis propose une aide à la planification chirurgicale. Enfin, Anatomage permet une analyse céphalométrique. Le tableau 8 présente une synthèse des applications existantes.

Compagny	Product	Function	Website	Origin
Denti.AI	Denti.AI	Dental disease detection	https://www.denti.ai/	Toronto,Canada
VidealHealth	VideaHealth	Dental disease detection	https://www.videa.ai/	Boston,USA
Pearl	Second Opinion	Dental disease detection	https://www.hellopearl.com/	West Hollywood,USA
Dentem	Dx Vision	Dental disease detection	https://www.dentem.co/	Toronto,Canada
Medihome	Medihome	Dental disease detection	https://medihome.jp/	Tokyo,Japan
DeepCare	DeepCare Dental Landscape	Dental disease detection	https://www.deepcare.com/	Beijing,China
ORCA Dental AI	Cephx	Cephalometric analysis Teeth segmentation Airways	https://www.orca-ai.com/	Las Vegas,USA
CellmatiQ GmbH	DentaliQ ortho DentaliQ opg	Cephalometric analysis	https://cellmatiq.com/	Hamburg,Germany
Uceph	Uceph	Cephalometric analysis	https://www.uceph.com/	Chengdu,China
Zhibeicloud	Zhibeicloud	Cephalometric analysis Bone age analysis Facial analysis	https:// www.aortho360.com/	Chengdu,China

Tableau 8 : Solution de logiciel d'IA pour l'orthodontie/dentisterie.

Téléradiographie de profil

La localisation des repères anatomiques (landmarks) est une étape cruciale pour l'analyse céphalométrique. Elle permet à l'orthodontiste de caractériser le crâne à l'aide des relations

spatiales entre ces points. Cette étape étant longue et fastidieuse, son automatisation présente un réel intérêt. En 2014, lors de l'International Symposium on Biomedical Imaging (ISBI) de l'IEEE, Wang et al. ont introduit le défi « Automatic Cephalometric X-Ray Landmark Detection Challenge 2014 **».** Ce défi propose comme tâche la localisation de 19 repères céphalométriques avec une marge d'erreur de 2 mm (qui est la marge d'erreur cliniquement acceptable).
Tant pendant l'édition de 2014 que pendant l'édition de 2015 de l'ISBI**,** ce sont les modèles non convolutifs qui ont prévalu. Le vainqueur de l'édition de 2015 obtint une erreur moyenne de 1,6 mm et un taux de détection réussi entre 75 et 85 % pour une plage de précision de 2 mm (Lindner et Cootes, 2015).

En parallèle, l'amélioration des performances des CNN et des bases d'apprentissage ont accru l'intérêt porté aux CNN en matière de détection de repères. La localisation de points de repères céphalométrique nécessite non seulement de relever la présence de l'objet remarquable mais également de générer une sortie qui le localise précisément. Cette tâche de localisation est généralement abordée comme une question relative à la régression (qui consiste à prédire une quantité) a l'inverse des tâches de classification (qui elles consistent à prédire une étiquette). Les CNN de localisation de repères présentent une architecture similaire aux CNN de classification, mais avec ceci de différent qu'ils intègrent un régreresseur sur leur dernière couche. Cette dernière couche se retrouve donc divisée en deux branches : un classifieur et un régresseur (Park et Kim, 2017).
Arik, Ibragimove et Xing (2017) ont implémenté un réseau qui classifie le profil osseux en se basant sur la position de ces repères. Parmi les modèles qui ont suivi, le plus remarquable fut Faster R-CNN (Ren, Girshick et Sun, 2016) et ses diverses déclinaisons. Il fournit une performance de précision en deçà de la marge de 2 mm cliniquement acceptable. Toutefois, il présentait le défaut de détecter des repères superflus. Pour limiter cet excès de détection, la méthode CephaNet (Qian et al., 2019) a été développée, en introduisant un facteur de « confiance maximale » qui permet d'éliminer les repères superflus.

Parmi les autres méthodes qui ont vu le jour, l'approche multi-échelle (Kwon et al., 2021) est fondée sur le constat que, contrairement au regard humain qui focalise son attention sur une zone de l'image, les CNN considèrent simultanément l'ensemble de l'image. En s'inspirant du processus

biologique, l'approche multi-échelle forme le CNN, en premier lieu, à porter un regard global sur l'ensemble des repères afin de les disposer avec parcimonie les uns par rapport aux autres.

Dans un second temps, cette approche se focalise sur chaque repère pour le positionner précisément en se basant sur des repères individuels. Cette méthode a mené à des résultats qui surpassent ceux obtenus à l'aide des méthodes conventionnelles. Suivant ce même principe, l'approche basée sur la « pyramide d'images » extrait des aperçus étroits de l'image d'entrée qu'elle affine de manière itérative pour accomplir des tâches de régression. Elle permet une performance de 86,72 % par rapport aux points positionnés à moins de 2 mm (Gilmour et Ray, 2020). Plus récemment, la méthode C-CNN (Cascade Convolutional Neural Network) (Kim et al., 2021) a montré une erreur moyenne de 1,36 mm s'approchant de la marge de 1,31 mm d'erreur humaine. Yu et al. (2020) ont développé un schéma multimodal d'apprentissage par transfert qui a montré une performance supérieure à 90 % à moins de 2 mm. Pour vérifier que le réseau se concentrait bien sur les zones pertinentes du crâne (sur des vues sagittales et transversales), ces chercheurs ont utilisé la technique des cartes de saillance.

En parallèle, des modèles de détection de repères sur téléradiographie de face ont été développés (Takeda et al., 2021). Par ailleurs, l'émergence de l'utilisation de CBCT (technique d'imagerie par tomodensitométrie) procure désormais des images en volume 3D. Des modèles de localisation de repères sur volume 3D ont donc été développés et réussissent une précision de l'ordre de 0,89 mm sur 18 repères (Chen et al., 2021).

La capacité de classification que possèdent les CNN rend également possible leur application à la téléradiographie de profil. Cette utilisation des CNN est plus complexe car elle nécessite une meilleure coordination entre les cliniciens en charge de la structuration et de la labélisation des données et les scientifiques (data scientists) en charge de l'élaboration du réseau de neurone. Mais cette utilisation présente un réel intérêt clinique, car elle apporte au praticien une information qualitative, précieuse pour son diagnostic et pour sa prise de décision thérapeutique. Ces apports sont illustrés par l'étude ayant trait au stade de maturation vertébrale sur téléradiographies de profil grâce à l'apprentissage d'un réseau neuronal multicouche, que nous avons mené et que nous présentons dans le chapitre suivant. Le point de départ de ce travail était le constat d'une difficulté pour le praticien de repérer systématiquement la maturation osseuse des vertèbres cervicales sur les téléradiographies de profil, alors qu'il s'agit d'une information précieuse dans la matrice de prise

de décision thérapeutique. Cette difficulté est probablement due au fait que sur l'imagerie, la zone des vertèbres cervicales est en dehors du champ visuel lorsque le praticien analyse la dysmorphose maxillo-faciale. Le praticien ne perçoit donc pas le secteur vertébral lorsque son attention est portée sur la dysmorphose maxillo-faciale. (Berbaum et al., 1990). Cette démarche est une illustration concrète de la stratégie d'augmentation, qui implique de s'appuyer sur les pratiques réelles du professionnel de santé pour envisager la manière dont ce travail pourrait être approfondi, accru et non pas réduit, par une utilisation élaborée de l'IA. Dans cette optique, nous allons présenter deux articles. Le premier est une application des CNN dans la reconnaissance des formes. Mais surtout, le second est une utilisation de l'interprétabilité des réseaux de neurones dans la compréhension d'une dysmorphose faciale.

6-4 Détermination du stade de maturation vertébrale sur téléradiographie de profil par l'apprentissage d'un réseau neuronal multicouche

Référence bibliographique de l'article présentant ces travaux et joint en annexe :

Masrour Makaremi, Camille Lacaule, Ali Mohammad-Djafari. (2019). Deep Learning and Artificial Intelligence for the Determination of the Cervical Vertebra Maturation Degree from Lateral radiography, Entropy(33),30 ; Doi: 10.3390

Le manuscrit de cet article est intégré aux annexes (Annexe 3).

6-4-1 Introduction

La quantité et la direction de la croissance maxillo-faciale peuvent être modifiées dans une certaine mesure par des facteurs extrinsèques. Cette possibilité constitue un des principaux leviers thérapeutiques en orthopédie dento-faciale. Cependant, le choix du moment où le traitement intervient est crucial. Un taux de croissance élevé peut conduire à corriger un déséquilibre des mâchoires, sinon une correction chirurgicale du décalage des mâchoires sera envisagée. Pour cette raison, le succès d'un traitement d'orthopédie dento-faciale dépend de la détermination temporelle de l'intervention considérée.

La dysmorphose dento-faciale la plus répandue est la classe II, qui correspond à une mandibule courte dont la croissance peut être stimulée par des traitements orthopédiques. De nombreuses

investigations radiographiques ont établi que la croissance condylienne / mandibulaire (figure 46) suit une courbe de croissance similaire à celle de la croissance staturale (Patcas et al., 2016). Ainsi, pour estimer son potentiel de croissance mandibulaire, le patient doit être localisé sur sa courbe de croissance staturale.

D'autres travaux (Jaqueira et al., 2010 ; Elhaddaoui et al., 2014) ont établi qu'il était possible de situer le patient par rapport à son pic de croissance pubertaire en se basant sur une classification des stades de maturation vertébrale sur une téléradiographie de profil (figure 46). Cependant, une expertise spécifique est nécessaire pour évaluer les étapes de la courbe de croissance vertébrale segmentée de CS1 a CS6 (Uysal et al., 2006). De plus, la zone des vertèbres cervicales est dissociée du champ visuel de la dysmorphose maxillo-faciale, ce qui complique le travail d'analyse mené par le praticien.

Les progrès de la vision par ordinateur grâce à l'apprentissage profond nous ont amenés à tenter d'identifier une méthode de détermination automatique du stade de maturation vertébrale afin d'en faire un outil d'aide à la prise de décision du praticien, concernant le choix du traitement à appliquer.

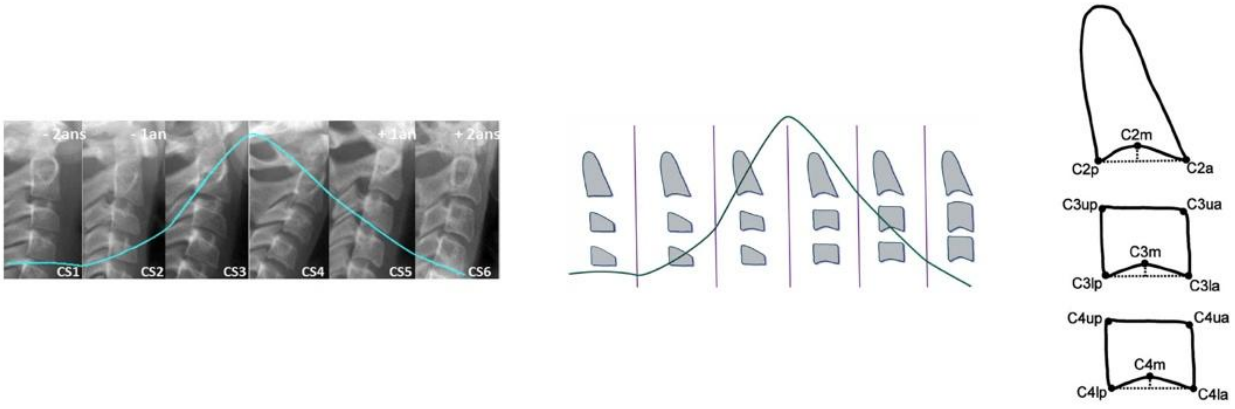


Figure 46 : Stades CS (cervical stage) radiologiques et morphologiques superposés à la courbe de Croissance staturale de Björk (Bjork et al., 1967).

6-4-2 Matériel et méthode

Afin de mener ce travail, nous avons puisé dans une base de données comportant 1870 téléradiographies de profil avec une répartition homogène des classes d'âges orthodontiques. Ces radios ont été extraites des dossiers de patients et anonymisées, et les patients ont donné leur

consentement éclairé à l'utilisation des données considérées. Le protocole a reçu la validation du comité d'éthique.

Chacune des 1870 radios a ensuite été labélisée par un expert qui a déterminé le stade de maturation vertébrale (de CS1 à CS6). Ces données labélisées ont alors été divisées en trois ensembles : l'entraînement, la validation et le test.

Seule une partie de la téléradiographie de profil nous est utile (figure 47). Nous avons effectué différents prétraitements, ce qui permet de réduire la taille de l'image de 2012 x 2012 à 256 x 256, ainsi que de réaliser l'indispensable économie de la puissance de calcul.

Dans cette même optique, nous avons également utilisé les propriétés du filtre entropique pour prétraiter les images, car nous nous intéressons davantage aux contours qu'aux niveaux de gris des images. Sur cette base, nous avons pu définir un filtre entropique capable de détecter les variations subtiles de la distribution des niveaux de gris localement.

La structure du modèle d'apprentissage profond est composée d'une couche convolutionnelle en entrée, de trois autres couches de réseaux de convolution (CNN) et d'une couche entièrement connectée de (32 x 32) à 6 classes.

Nous avons principalement utilisé deux algorithmes standard : Stochastic Gradient Descent (SGD) et Adaptive Moment Estimation (Adam).

Figure 47 : Partie isolée (vertèbres cervicales) d'une téléradiographie de profil.

6-4-3 Résultats

Le modèle avait pour objectif de classer les images selon leur niveau de maturation vertébrale (CS1, CS2, CS3, CS4, CS5, CS6). Pour former le modèle proposé, dans un premier temps, nous avons utilisé 360 images étiquetées disponibles avec ces données. La précision des images testées atteignait les 90% dans l'estimation du stade de maturation vertébrale (inférieur à 70%). Dans un deuxième temps, nous avons augmenté le nombre de données définies à 600, puis à 900 et enfin aux 1870 images disponibles. Quand on passe de 360 à 600 puis à 900 et enfin à 1870 images, les performances sont améliorées. Cependant, le progrès est moins important lors du passage de 900 à 1870, peut-être parce qu'un palier est atteint, mais également peut-être parce que le nombre d'images dans chacune des 6 classes n'est plus équitablement réparties lors du passage de 900 à 1870 radios.

La précision d'estimation du stade de maturation vertébrale obtenue avec l'ensemble de l'échantillon est tout à fait satisfaisante, légèrement meilleure que la précision mesurée chez des experts (orthodontistes entrainés) (Uysal et al., 2006). La performance optimum est obtenue avec les stades de maturation CS1 et CS6 (0.99) qui ont des formes anatomiques bien identifiables et la précision la plus faible concerne le stade de maturation CS3 (0.953) qui présente la plus grande variabilité anatomique (Tableau 9).

Ces résultats ont été obtenus à partir d'un réseau de neurones simple à six couches. Il faut noter (figure 48) que la performance baisse lorsque l'on passe de 6 couches à 7 couches.

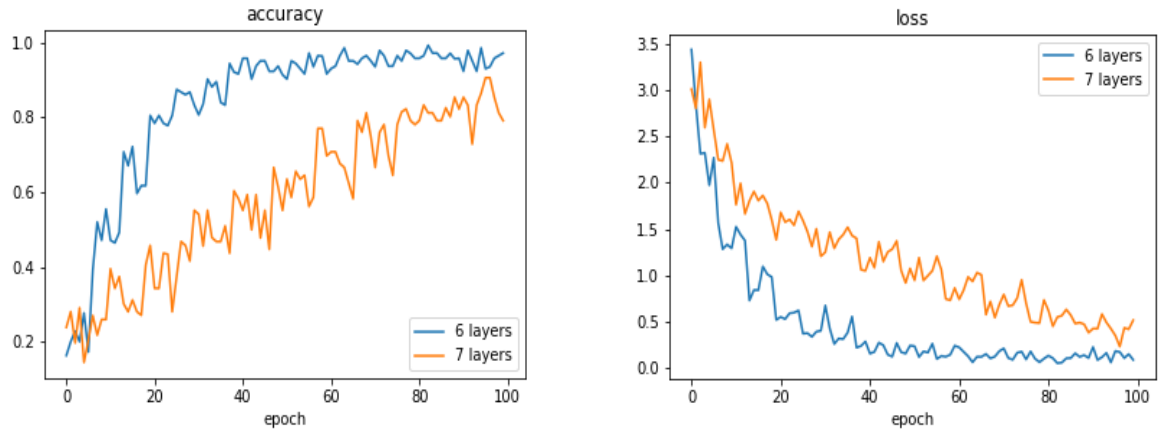


Figure 48 : Evolution de la fonction de perte et de la précision pour les 1870 images examinées. En bleu lévolution d'un réseau à 6 couches et en orange celle d'un réseau à 7 couches.

	CS1	CS2	CS3	CS4	CS5	CS6
Performance du réseau	0.990	0.970	0.953	0.980	0.977	0.990

Tableau 9 : Précision de la détermination du stade de maturation vertébrale. L'algorithme d'optimisation utilisé dans le cadre de la formation était ADAM.

6-4-4 Discussion

Les performances du réseau neuronal quant à l'évaluation sont satisfaisantes pour une utilisation clinique, avec un taux de précision dans la reconnaissance des différents stades de maturation vertébrale aussi fiable que d'autres utilisations de CNN dans le domaine médical (Antoniadi et al., 2021) et la performance d'estimation du stade de maturation vertebral par les experts (Uysal et al., 2006). Le filtre entropique capable de détecter les variations subtiles de la distribution des niveaux de gris a été un facteur clé dans la performance du réseau de neurones. Il est toutefois à noter que les images constituant cet échantillon proviennent toutes de la même unité de soins, sans oublier le manque de variabilité des appareils radiologiques utilisés afin de réaliser les téléradiographies de profil. Par conséquent, le volume des données doit être augmenté et, ce qui est encore plus important, leurs sources doivent être diversifiées, pour permettre une utilisation clinique performante.

Cet outil ne doit surtout pas être perçu comme se substituant à la compétence du praticien, mais comme un outil aidant ce dernier à se concentrer sur l'essentiel, c'est-à-dire la prise de décision. Cette démarche qui vise à intégrer l'efficacité du CNN dans une stratégie d'augmentation, c'est à dire permettant le meilleur rapport possible entre le praticien et l'IA, peut être renforcée grâce aux travaux sur l'interprétabilité des réseaux de neurones convolutionnels.

6-5 Utilisation de l'interprétabilité du Deep Learning pour l'extraction d'informations morphologiques en imagerie médicale : cas d'application pour la rétrognathie mandibulaire.

6-5-1-Introduction

Comme le souligne (Fourcade et al., 2019) l'intelligence artificielle, et particulièrement la vision par ordinateur, fonctionne comme un troisième œil pour le praticien, permettant d'élargir et d'améliorer sa perception des données paracliniques. Mais la question se pose de savoir ce que voit précisément ce troisième œil.

L'interprétabilité des réseaux de neurones convolutifs permet d'apporter une réponse à cette question. En effet, les zones d'une image à laquelle le réseau de neurones prête attention pour la discriminer peuvent constituer une source d'information précieuse.

Nous défendons l'hypothèse que cette information contribue à visualiser les corrélations entre les dysmorphoses maxillo-faciales et l'ensemble de l'architecture cranio- faciale.

Cette approche est illustrée dans ce paragraphe, où nous étudions l'apport de l'interprérabilité des réseaux de neurones convolutionnels (CNN, pour Convolutional Neural Netwoks) à la recherche de corrélations entre la rétrognathie mandibulaire et l'architecture cranio faciale. En effet depuis la différenciation du genre Homo, une cinétique concomitante de rétrusion mandibulaire a été établie (Bastir et al., 2005) dont les mécanismes phylogénétiques et ontogénétiques sont encore inconnus. Cette diminution de la longueur mandibulaire s'accompagne d'une augmentation du volume cérébral, et la relation entre le degré de rétrécissement mandibulaire et son impact sur l'architecture craniofaciale reste à clarifier. Néanmoins, le passage du prognathisme à un rétrognathisme de plus en plus accentué s'est accompagné d'une série de modifications craniofaciales au cœur de l'évolution humaine (Puech et al., 2021).

Ces transitions évolutives se sont accompagnées de changements alimentaires et comportementaux tels que la bipédie et l'émancipation du milieu arboricole qui ont modifié les contraintes sélectives. La révolution industrielle a entraîné une nouvelle mutation des habitudes alimentaires avec l'émergence de produits transformés, majoritairement mous, et d'un mode de vie plus sédentaire. Ces modifications entraînent un déséquilibre des forces subies par le crâne au cours de son développement, ce qui se traduit par une augmentation des cas de classe II par

rétrognathie mandibulaire (Gazit et al., 1984). La classe II par rétrognathie mandibulaire (C2Rm) est une pathologie affectant la mandibule inférieure. Il en résulte une mandibule sous-développée qui perd son alignement avec la mâchoire supérieure. La conséquence est une perte de stabilité qui entraîne une nouvelle répartition des contraintes biomécaniques communiquées à l'ensemble du crâne (Liberman et al., 2009).

Une meilleure compréhension des points d'impact du C2Rm sur l'ensemble du crâne est donc un enjeu majeur dans le diagnostic, le traitement et la prise en charge de cette dysmorphie, mais aussi dans une réflexion sur les modifications craniospinales survenues au cours de l'évolution humaine. La relation sagittale du maxillaire à la mandibule est un critère diagnostique important dans le traitement orthodontique.

Bien que l'observation clinique aide à déterminer la relation sagittale dans une certaine mesure, une évaluation précise n'est possible qu'avec l'aide de la céphalométrie. Divers paramètres angulaires ont été formulés pour évaluer les écarts de mâchoire dans le plan sagittal. Un modèle paramétrique classique avec trois angles céphalométriques standard est largement utilisé pour évaluer la projection de la face vis-à-vis de la base du crâne (Katz, 1992). L'ouverture de l'angle Maxillaire-Nasion-Mandibulaire (ANB) permet d'estimer le décalage entre la mandibule et le maxillaire. Il permet de classer la classe squelettique (classe I [-1°-6°], classe II [>6°], classe III [<-1°]). Les projections antéropostérieures du maxillaire et de la mandibule sont mesurés respectivement par les angles Sella-Nasion-Maxillaire (SNA) et Sella-Nasion-Madibulaire (SNB). Ils permettent de classer la position de la mandibule est de déterminer la situation physiologique vs Rétrognathe/prognathe (Rétrognathie [SNB<76°], Physiologique [76°-80°], Prognathie [SNB>80°]) et le maxillaire (Rétrognathie [SNB<78°], Physiologique [78°-82°], Prognathie [SNB>82°]) (Yüksel et Esenlik, 2020).

Des décennies de recherche basée sur les paramètres céphalométriques étudient les altérations cranio-faciales corrélées à la C2Rm. Des corrélations statistiques entre rétrognathie mandibulaire et vertèbres cervicales (Gomes, 2014), voies respiratoires supérieures (Shokri, 2018 ; Chuan, 2021), sinus frontal (Tehranchi, 2017), base crânienne (Almeida, 2017) ont été mise en évidence par de nombreux auteurs. Si cette technique est suffisamment efficace pour dépister la dysmorphie, elle ne rend pas compte de toutes les structures du crâne atteintes. Deux limites de cette méthode

paramétrique ont été mises en évidence : l'analyse étudie individuellement la corrélation entre C2Rm et l'altération d'une structure, mais une vision globale de l'architecture n'est pas possible. La seconde limite est la difficulté de montrer statistiquement la corrélation en fonction de la sévérité de la pathologie du fait de la forte variabilité interindividuelle de l'architecture craniospinale. Ces limites ont été en partie contournées par la Morphométrie Géométrique (GM) qui a proposé une approche appelée « analyse de Procuste » (Ross, 2004). La distance de Procuste, utilisée pour caractériser l'impact de la C2Rm sur la forme du crâne est une mesure de la différence de forme entre deux structures. Mais les limites de l'approche GM ont également été largement soulignées (Adams et al., 2012). Globalement, ces méthodes nécessitent d'analyser les structures indépendamment et donc de les prédéfinir, et ne permettent pas d'isoler ni de définir les unités morphogénétiques réelles qui interagissent. Elles ne permettent pas non plus d'identifier ni de quantifier leurs interactions directes et indirectes ni de comprendre les processus par lesquels elles interagissent (Liberman, 2000).

À notre connaissance, aucune étude n'a pu rendre compte de l'ensemble des structures impactées par la C2RM. Cet objectif se heurte aux limites des outils méthodologiques classiques de l'anthropométrie ce qui amène Daniel Liberman (2000) à déclarer que *le défi de la morphométrie est de trouver un outil capable de capturer la co-variation*. Les progrès récents des réseaux de neurones convolutifs (CNN) apparaissent comme un outil pertinent pour relever ce défi.

Ces dernières années, en imagerie médicale, les réseaux de neurones profonds (DNNs) et en particulier les réseaux de neurones convolutifs (CNNs) ainsi que les méthodes de class activation mapping (CAM) et leurs variant Score-CAM (Wang et al., 2019) sont devenus particulièrement performant pour localiser les régions discriminantes (biomarqueurs) dans les images anormales. Une telle approche a été appliquée à divers problèmes d'imagerie médicale (Kim, Rajaraman et Antani, 2019) tels que les images du fond d'œil rétinien (Daanouni et al., 2021), la détection de cancer (Masud et al., 2020), la détection de la pneumonie (Panwar et al., 2020) ou le diagnostic de la maladie d'Alzheimer (Irie et al., 2020).

L'objectif de ce travail est d'utiliser pour la première fois des méthodes de deep learning (DL) appliquées aux images médicales de rétrognathie mandibulaire afin de :

a) préciser les modifications crâniennes impliquées dans cette dysmorphie. Ces informations seront cruciales à la fois en termes de diagnostic mais aussi pour la mise en place d'approches thérapeutiques innovantes ;
b) préciser (également pour la première fois) les structures impactées en fonction de la sévérité de cette dysmorphie (la sévérité sera appréhendée à l'aide de l'angle ANB) ;
c) d'utiliser ces données pour discuter des liens entre la diminution de la longueur de la mandibule et les modifications craniofaciales au cours de l'évolution humaine.

6-5-2 Méthodes

6-5-2.1 Aperçu des méthodes

L'étude est résumée dans la figure 50. Un modèle binaire de classification de crâne a été entraîné et validé avec un jeu de données de 1317 images de radiographies latérales de classe I physiologique et 1377 images de radiographies latérales de classe II pathologique. La classification a été réalisée par deux orthodontistes (figure 49) avec un logiciel de mesure basé sur la classification céphalométrique de Tweed (Kula et Ghoneima, 2020). L'ensemble des images a ensuite été utilisé pour entraîner un modèle de classification permettant de diagnostiquer la CL II.

La méthode Score-CAM a été utilisée pour générer des cartes de saillance. Cette méthode identifie la zone de l'image céphalométrique utilisée pour la classification. Le résultat a été visualisé en deux étapes. La première consiste à identifier la carte d'activation moyenne de la classe II. La deuxième étape consiste à sous-échantillonner la distribution de la classe II selon l'angle ANB dans une fenêtre de 5 grades : [4,5°-6°[, [6°-7°[, [7°-8°[, [8°-9°[, [>9°] et à générer une carte de saillance pour chaque grade.

6-5-2.2 Ensemble de données

23479 téléradiographies de profil recueillies entre janvier 2012 et décembre 2021 dans 4 cliniques d'orthodontie, de patients français se plaignant d'une dysmorphose dento-faciale et/ou d'une malocclusion, ont été utilisées pour concevoir une base de 2694 téléradiographies d'entraînement et de test d'un modèle d'apprentissage profond (1/3 d'hommes et 2/3 de femmes avec un âge

moyen de 12,63 ± 5,06 ans). Ces images céphalométriques ont été obtenues à l'aide d'un appareil de radiographie CARESTREAM CS 9000-C.

Pour réussir à obtenir ces 2694 téléradiographies, plusieurs étapes de sélection ont été nécessaires. Partant des 23479 téléradiographies, deux orthodontistes ont sélectionné le premier pool d'images sur la base de leur qualité et de leur haute reproductibilité (crâne en position standard - inclus dans un céphalostat avec le plan de Frankfort parallèle au sol (Silva et al., 2003) et radiographies plein champ du crâne entier). 11 193 images ont ainsi été sélectionnées. Ces images ont ensuite été classées en deux catégories, physiologique (2700) et pathologique (4400), en utilisant la mesure des angles ANB [0°-5°], SNB [<77°] et SNA [79°-85°].

Dans l'étape suivante, les cas se situant à la limite entre physiologique et pathologique (ANB [(-1°) - 1°]) ont été exclus pour améliorer la discrimination. Un ensemble d'entraînement final comprenant 1317 images physiologiques de classe I et 1377 images de classe II a ainsi été obtenu. La figure 49 présente ce processus de labélisation.

Compte tenu de l'importance du prétraitement de la redimensionnalisation (Makaremi et al., 2019) toutes les images radiographiques ont été redimensionnées à 256 x 256 pixels comme présenté au point 6-6-2-4 Prétraitement des images.

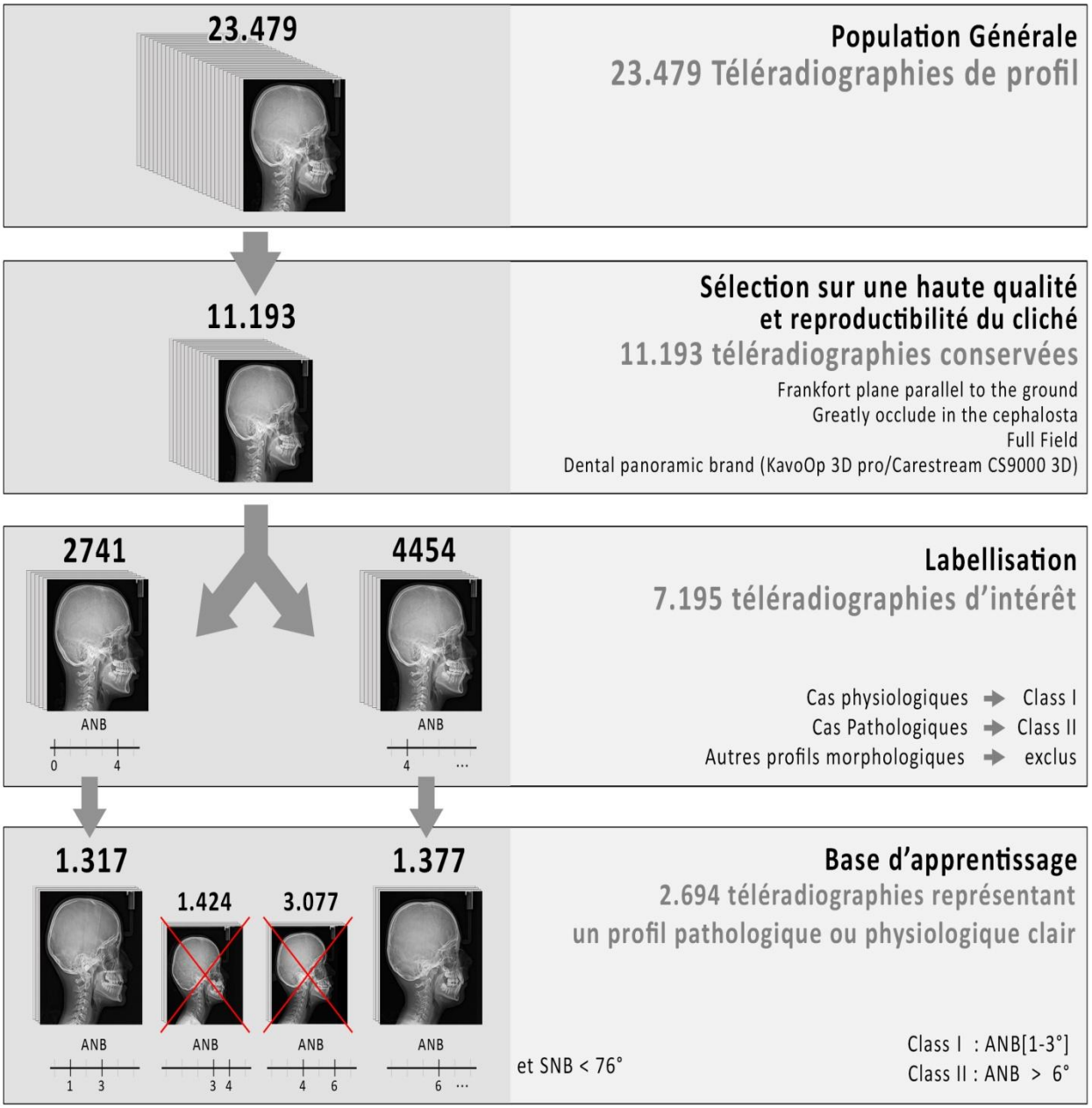


Figure 49 : Processus de sélection et de labellisation des téléradiographies pour l'élaboration du jeu de données.

6-5-2.3 Implémentation du réseau de classification

Un modèle de classification binaire a été réalisé afin d'extraire les caractéristiques de la forme des os à partir des radiographies et de les utiliser pour discriminer les radiographies de classe I et les radiographies de classe II.

Pour concevoir l'architecture du modèle, il faut un réseau suffisamment profond comportant assez de paramètres pour stocker l'ensemble des données et extraire des caractéristiques de haut niveau (Bengio et Dellaleau, 2011 ; Eldan et Shamir, 2016).
Tout d'abord, pour ce qui est de la taille du modèle, un modèle à 7 couches avec 867 178 paramètres a montré la meilleure précision avec 95% d'identification correcte de la classe (physiologique ou pathologique). Chacune de ces couches a été couplée à une normalisation par groupe et à une fonction d'activation ReLU (He et al., 2018).

Une fonction d'activation softmax a été utilisée sur la dernière couche entièrement connectée pour réaliser la prédiction de la classe. Le modèle a été affiné à l'aide de l'optimiseur ADAM (Kingma et Ba,2014) avec un taux d'apprentissage décroissant commençant à 10^{-3} et diminuant de 0,95 par 100 époques. Un total de 1000 époques a été atteint avec une taille de lot de 100 dans une GPU P100 Nvidia, en utilisant 70 Gbs de mémoire.

Dans ces méthodes de classification d'apprentissage profond, comme le nombre de paramètres est élevé, la théorie suggère qu'une certaine forme de régularisation est nécessaire pour assurer que l'erreur de généralisation reste minime et éviter le surajustement (Zhang et al., 2016).
Dans ce contexte, la théorie de l'apprentissage statistique dispose d'un certain nombre d'outils différents capables de contrôler et d'éviter le surajustement. Dans notre étude, des méthodes d'augmentation des données et de dropout ont été utilisés pour contrôler l'erreur de généralisation (Zhang et al., 2016).

- Augmentation des données : Pendant l'apprentissage, les radiographies prétraitées ont été augmentées avec les opérations suivantes : translations horizontales et verticales aléatoires ; rotation par pas de 10 degrés, décalage de la largeur et de la hauteur et zoom dans la plage de 3%.

- Dropout : une couche de dropout avec un taux de 0,5 a été utilisée.

6-5-2.4 Prétraitement des images

L'étape de prétraitement permet de normaliser les données entrantes pour le CNN et d'optimiser l'unité de traitement graphique (GPU) en termes de consommation de mémoire. En raison de la nature de la fonction sigmoïde (ou softmax) utilisée dans l'algorithme de backpropagation, les valeurs de l'intensité des pixels doivent être normalisées entre 0 et 1 avant d'être transmises au réseau neuronal (Dawson et Wilby, 1998). En outre, pour améliorer la détection des bords, un processus de Sobel a été appliqué (Gao et al., 2010) pour supprimer les régions non-variables et mettre en évidence les forts changements comme les bords.

6-5-2.5 Précision du modèle

Pour déterminer la précision du modèle, une validation croisée en 5 volets a été utilisée et a montré une précision de 95%.

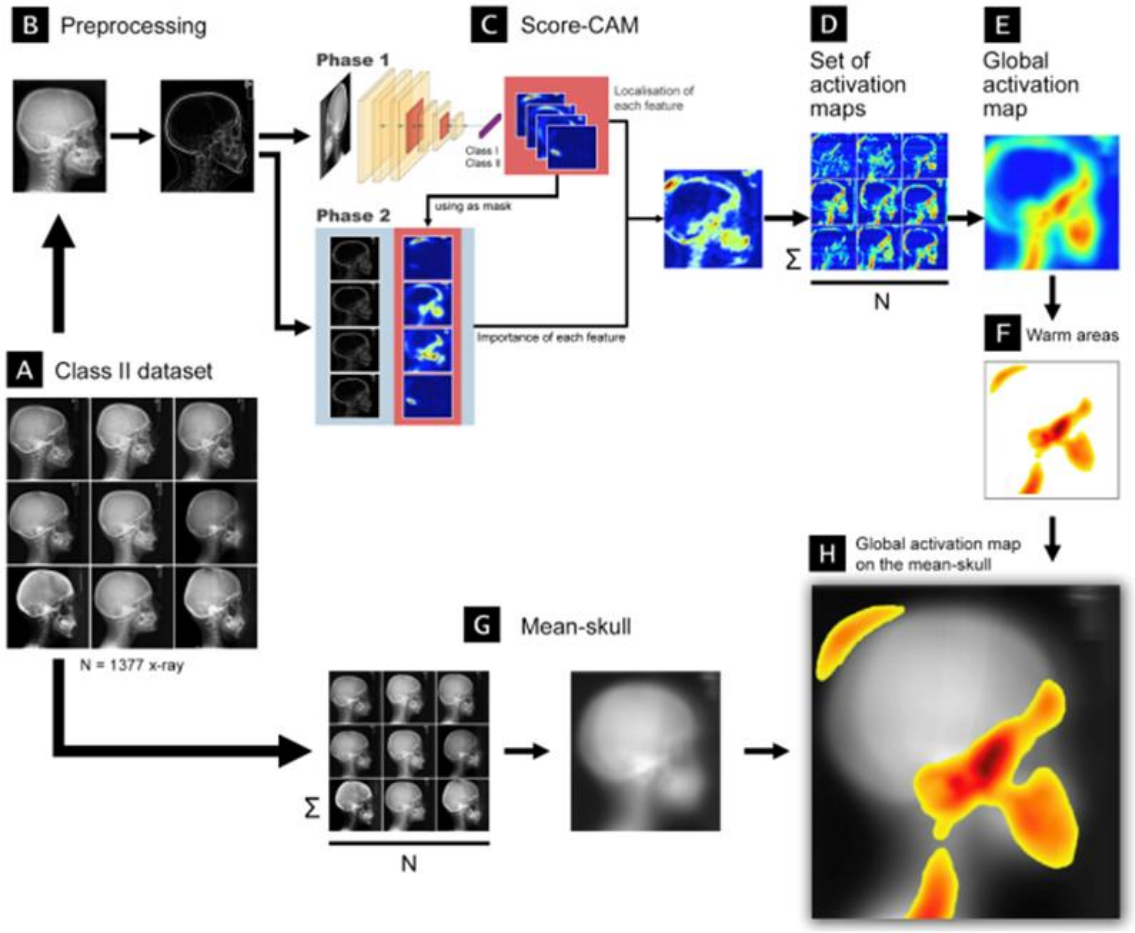


Figure 50 : A- Jeu de données : le jeu de données est composé de 1 377 radiographies latérales de classe II. B- Pour chaque radiographie, un prétraitement Sobel a été effectué. C- Les radiographies prétraitées sont utilisées dans un processus de classification DL-CNN d'entraînement-évaluation-test, puis des cartes d'activation sont générées. D- Les cartes d'activation générées de toutes les radiographies sont collectées. E- Une carte d'activation moyenne est générée à partir des 1377 cartes d'activation. F- La zone chaude de la carte d'activation moyenne est extraite. G- Un crâne moyen est généré à partir des 1377 cartes d'activation. H- Enfin, la température chaude de la carte d'activation globale est extraite et superposée au crâne moyen.

6-5-2.6 Carte d'activation : Technique de superposition

Pour chacune des images, une carte d'activation a été générée à l'aide de Score-CAM (Wang et al., 2019). La position constante du crâne sur l'image (*cf.* : 6-6-2-2 Ensemble de données) permet de superposer les cartes d'activations individuelles.

6-5-2.7 Comparaison par zones

Cinq zones crânio-faciales distinctes ont été identifiées sur la base de critères anatomiques (Vion, 1997) : le sinus frontal, la base du crâne, la partie corticale de l'os occipital, l'os pariétal et les vertèbres cervicales.

Ces zones ont été identifiées sur la base d'une revue portant sur les caractéristiques crânio-faciales des sujets atteints de rétrognathie mandibulaire. La recherche a été menée sur PubMed, Embase et la Cochrane Library. D'autres articles pertinents ont été obtenus par des recherches manuelles et des recherches sur Google Scholar. Les critères d'inclusion étaient le recrutement de patients diagnostiqués avec une classe squelettique II basée sur l'angle ANB et SNB, la disponibilité de la téléradiographie de profil pour chaque participant et la présence d'un coefficient de corrélation entre une variable céphalométrique et la rétrognathie mandibulaire.
Les zones identifiées ont ensuite été colorisées sur une téléradiographie de profil ci-dessous figure 51.

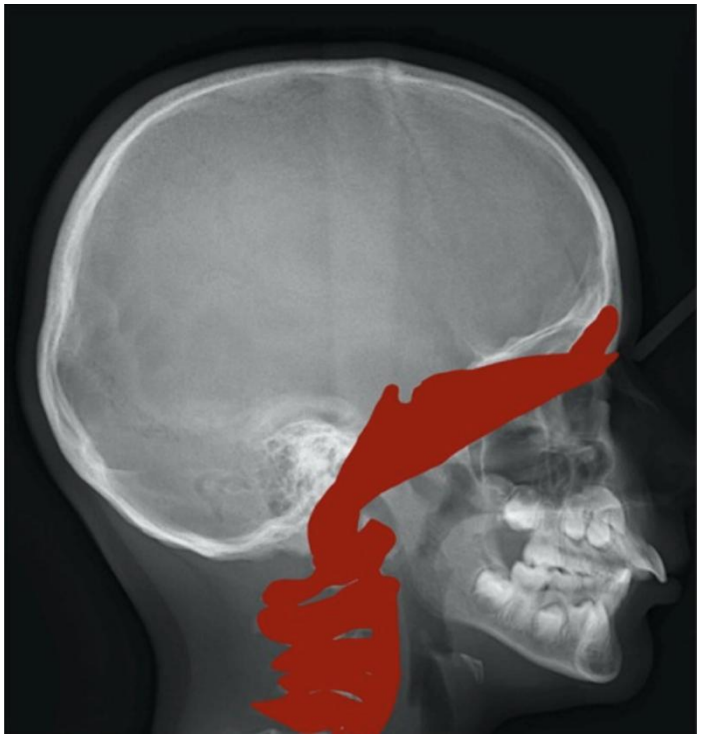


Figure 51 : Schéma colorisé sur la base des zones identifiées comme significativement corrélées a la classe II par rétrognathie mandibulaire.

6-5-3 Résultats

Les résultats sont présentés en 2 parties : l'analyse descriptive et l'analyse statistique.

Chacune de ces parties présentes :

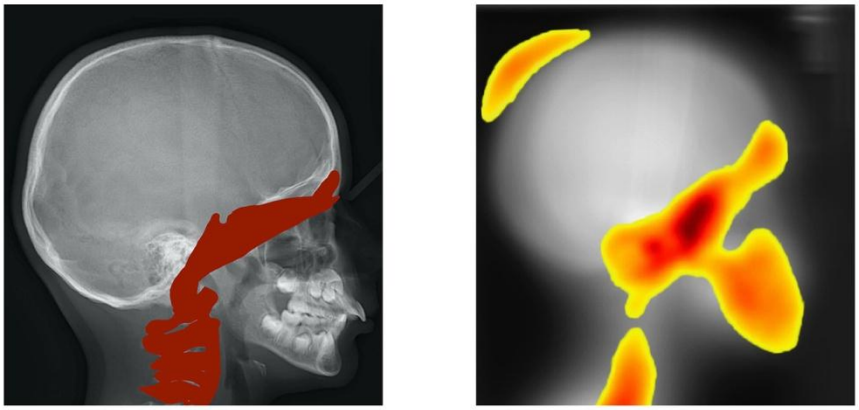
1) Tout d'abord la carte de saillance globale (classisfication selon l'état physiologique vs. Parthologique)
2) Les cartes de saillances selon la gravité de la rétrognatie (gravité formalisée à l'aide l'angle ANB).

6-5-3.1 Analyse descriptive

Carte de saillance globale

La figure 52 montre en B la carte d'activation établie sur toutes les classes II classées par le CNN. La carte d'activation montre les zones de l'image utilisées par le réseau pour fournir sa prédiction. Cette carte d'activation est établie en faisant la moyenne de l'ensemble de l'échantillon des classes II. Elle représente par le biais d'un code couleur l'importance des pixels de l'image qui participe à la prédiction de la classe. Le jaune indique les pixels de faible importance, l'orange, les pixels d'importance moyenne, et le rouge, les pixels très importants.

En A est affichée (pour comparaison) la carte des structures significativement impactées par la classe II, identifiée dans la littérature. Le processus détaillé d'élaboration de cette carte est décrit dans la section (6-6-2-7 comparaison par zone) de la partie Méthodes.



A - Carte des structures corrélées à l'angle SNB

B - Carte de saillance

Figure 52 : A structures identifiées dans la littérature et B Résultat de notre carte de saillance.

Nos résultats montrent la présence de différentes zones chaudes. Dans la partie supérieure gauche, une première zone légèrement incurvée représentée en jaune est mise en évidence. Dans la partie inférieure droite, un halo d'intensité jaune-rouge émerge. Enfin, une troisième zone plus mince se déploie en diagonale, partant du quart inférieur gauche et aboutissant au quart supérieur droit. Une inflexion dans la partie centrale divise cet axe en deux segments, tous deux d'intensité jaune-rouge. Le point d'inflexion est également marqué par un halo d'intensité rouge.
Ces zones semblent chevaucher des structures osseuses. La première zone est superposée à l'os pariétal. La deuxième zone est centrée sur les mâchoires. La troisième zone est divisée en deux segments. Le premier segment couvre les vertèbres cervicales. Il commence à C6 et se poursuit jusqu'à l'atlas. L'activation suit un plan de gravité le long de cet axe vertébral. Le point d'inflexion apparaît centré sur l'os sphénoïde. Le deuxième segment part du sphénoïde et suit la base du crâne jusqu'au sinus frontal.

Ces résultats montrent qu'il existe un morphotype spécifique à la population extrait par le CNN. De plus, ce morphotype se superpose aux structures identifiées dans la littérature (vertèbres cervicales, sphénoïde, base du crâne au sinus frontal). Mais, notre carte d'activation montre également une activation de l'os pariétal, activation qui n'avait jamais été identifiée auparavant.

Carte de saillance selon le degré de séverité de la rétrognathie

Pour approfondir, nous examinons l'évolution de la carte thermique en fonction de la sévérité de la classe II (figure 53). Le degré de sévérité est appréhendé par l'amplitude de l'angle ANB. A cet effet, nous présentons sur la ligne du bas les cartes de saillance pour les angles ANB sous-échantillonnés à partir de [4.5°-6°] et augmentant progressivement jusqu'à [9°<] par pas de 1°. Sur la ligne supérieure est affichée, à titre d'exemple, la radiographie d'un patient pour chaque intervalle correspondant.

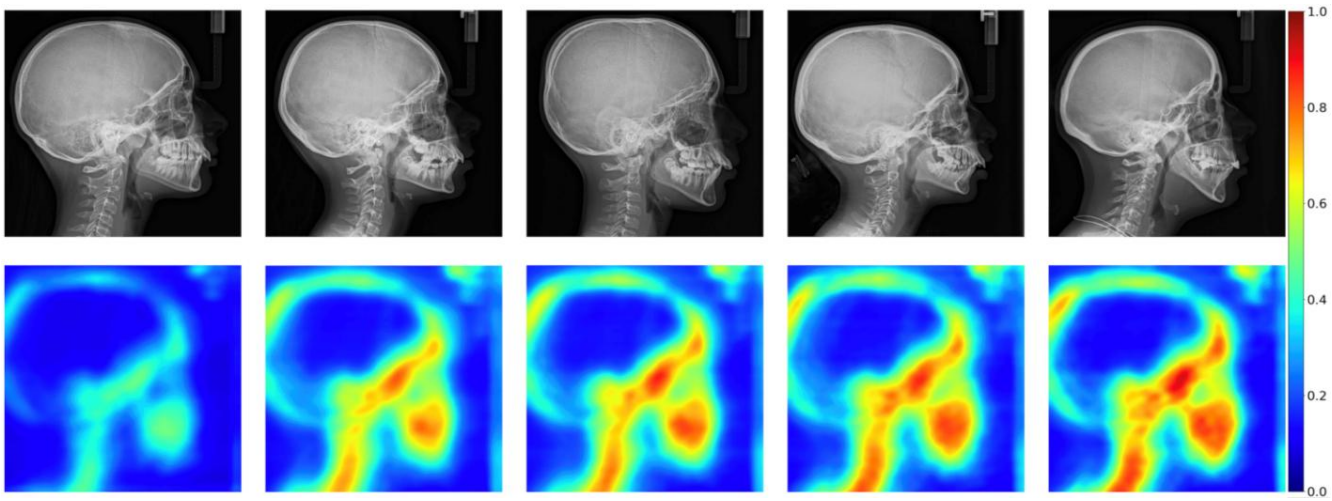


Figure 53 : La rangée du haut montre, pour chaque intervalle de sévérité, un exemple de crânes. La rangée du bas montre les cartes de saillance pour chaque intervalle par ordre croissant de sévérité. La barre de couleur indique la quantification de la gravité.

La première carte [4.5°-6°] montre une faible activation globale. Son intensité maximale est distribuée le long des structures osseuses (zone verte). La deuxième carte montre une évolution du même type, les couleurs chaudes apparaissent sur certaines zones qui apparaissaient auparavant en vert. Ces couleurs chaudes sont également plus diffuses. Plus précisément, la partie inférieure du segment cervical, le segment supérieur jusqu'au sinus et le halo mandibulaire apparaissent en intensité orange-rouge. Le reste des vertèbres et la base crânienne, ainsi que l'os pariétal, sont représentés en jaune. La troisième carte montre une amplification de l'intensité des zones, qui s'intensifie plus fortement vers le rouge. Les segments cervicaux et basilaires sont maintenant

complètement rouges. Le diamètre du halo mandibulaire reste le même, mais le rouge s'intensifie. La quatrième carte montre un épaississement des segments cervico-basilaires et une activation du point d'inflexion (sphénoïde). Le halo mandibulaire continue de s'étendre vers la partie supérieure du maxillaire. L'os pariétal passe du jaune à l'orange. La cinquième colonne est caractérisée par l'activation de l'os pariétal (orange à rouge) et l'épaississement de toutes les autres zones rouges.
Dans l'ensemble, sur toute l'amplitude de l'angle ANB, la cinétique d'intensité croissante des pixels et l'expansion de la surface activée deviennent évidentes.

6-5-3.2 Analyse statistique

Afin d'évaluer l'impact de l'angle ANB (sévérité de la retrognatie) sur le score de classe et la surface de pixels d'importance pour les quatre structures anatomiques, nous appliquons une analyse de variance (ANOVA) à deux facteurs (Angle ANB et structures anatomiques). L'analyse effectuée montre un effet significatif du facteur angle à la fois sur le score de classe $F(4, 13992) = p<0.001$ et sur la surface chaude $F(4, 13992) = p<0.001$ avec une augmentation du score de classe le long de l'angle ($\bar{x}$([4.5°-6°[)=0. 464, $\bar{x}$ ([6°-7°[)=0.474, $\bar{x}$([7°- 8°[)=0.487, $\bar{x}$([8°-9°[)=0.503, $\bar{x}$([>9°])=0.500) et aussi une augmentation de la surface chaude ($\bar{x}$([4. 5°-6°[)=0,416, $\bar{x}$([6°-7°[)=0,433, $\bar{x}$([7°-8°[)=0,451, $\bar{x}$([8°-9°[)=0,466, $\bar{x}$([>9°])=0,471).
L'analyse montre également un effet significatif de la structure anatomique sur le score de classe $F(3, 13992) = p<0,001$. La base crânienne et les vertèbres cervicales ($\bar{x}$(Base crânienne)=0,588, $\bar{x}$(Vertèbres)=0,557) présentent un score de classification plus élevé que l'os pariétal et le sinus frontal ($\bar{x}$(Sinus frontal)=0,353, $\bar{x}$(Pariétal)=0,387). De même, l'analyse montre un effet significatif de la variable aire sur surface chaude $F(3,13992) = p<0,001$. La base crânienne et les vertèbres cervicales ($\bar{x}$(Base crânienne) =0,566, $\bar{x}$(Vertèbres)=0,534) ont une surface chaude plus grande que le sinus frontal et l'os pariétal ($\bar{x}$(Sinus frontal)=0,305, $\bar{x}$(Pariétal)=0,304).
Aucun effet d'interaction n'est observé pour le score de classification $F(12, 13992) = p=0,379$, ni pour la surface $F(12 , 13992)= p=0,530$.
Le manque de significativité de l'interaction est dû au fait que l'augmentation est la même pour chacune des structures.

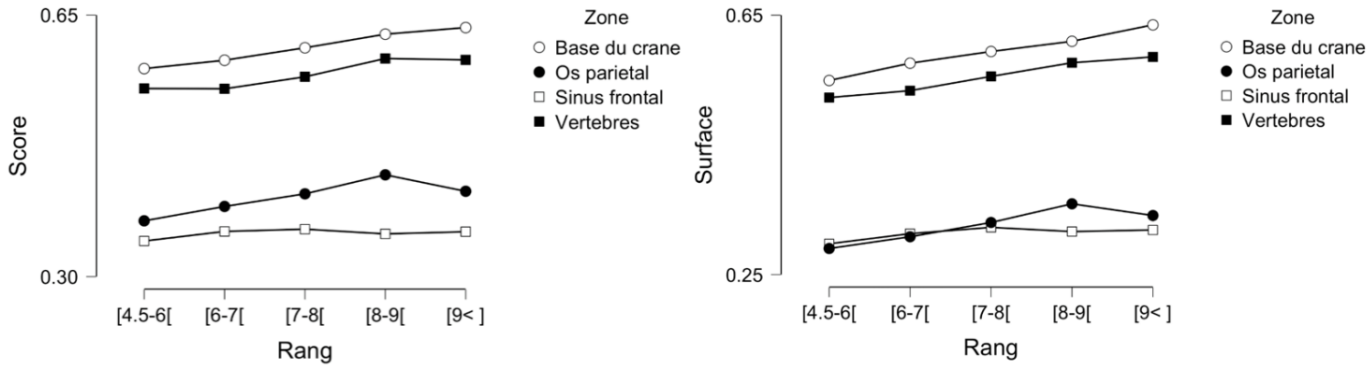


Figure 54 : Pour chacune des quatre structures squelettiques (base crânienne, sinus frontal, os pariétal, vertèbres cervicales), le graphique de gauche montre l'évolution du score de classe suivant l'angle ANB (rang). Le graphique de droite montre l'évolution de la surface des pixels chauds suivant l'angle ANB (rang).

6-5-4 Discussion

6-5-4-1 Extraction d'informations morphologiques

A notre connaissance, notre étude est la première à placer directement comme objectif la génération d'une carte de saillance globale afin d'identifier les zones d'intérêt detectées par le réseau. Cette inversion constitue un renversement du paradigme de l'interprétabilité qui utilise ordinairement les cartes de saillance pour vérifier que le réseau a correctement effectué la classification (Herman, 2017). Cet outil a permis de synthétiser le contenu informationnel de tous les échantillons pathologiques en générant des cartes thermiques globales de la pathologie. De cette manière, nous avons éliminé la variabilité interindividuelle et ainsi extrait le pattern (biomarqueur) de la pathologie. Ce biomarqueur confirme que le morphotype pathologique est étendu sur les zones occlusales. Notre technique a permis de se passer des mesures céphalométriques habituelles mais de tout de même retrouver (comme éléments clés de la rétrognathie mandibulaire) les structures établies dans la bibliographie (base crânienne, vertèbres et sinus frontal). L'un d'eux, le sinus frontal, était particulièrement controversé. Notre carte de saillance indique que cette zone est discriminante. Ce résultat plaide donc pleinement en faveur d'un impact du C2Rm sur le sinus frontal. Par ailleurs, notre carte de saillance a montré que l'os pariétal est également discriminant. Cette zone n'avait jamais été identifiée dans la literature.

Dans notre seconde analyse, La représentation offerte par Score-CAM a permis une visualisation inédite de la cinétique d'apparition et d'expansion des dysmorphyses sur ces structures périphériques en fonction du degré de sévérité de la pathologie tant sur le score de classe que sur la surface chaude avec une augmentation du score de classe le long de l'angle. Les cartes de saillance par niveau de sévérité de la pathologie ont permis de visualiser le comportement global de certaines structures. Ce comportement se traduit par un "allumage progressif" du score de classe ($\bar{x}$([4.5°-6°[)=0.464, $\bar{x}$([6°-7°[)=0.474, $\bar{x}$([7°- 8°[)=0.487, $\bar{x}$([8°-9°[)=0.503, $\bar{x}$([>9°])=0. 500) et une extension de la surface chaude ($\bar{x}$([4,5°-6°[)=0,416, $\bar{x}$([6°-7°[)=0,433, $\bar{x}$([7°-8°[)=0,451, $\bar{x}$([8°-9°[)=0,466, $\bar{x}$([>9°])=0,471) sur un axe dépendant du degré de gravité. Cet axe comprend les vertèbres cervicales, la base du crâne et le sinus frontal. Jusqu'à présent, ces structures ont été étudiées séparément. Une telle visualisation globale n'avait jamais été réalisée auparavant. Nos résultats montrent à travers l'angle ANB une cinétique croissante du score de classe et de la surface de pixels discriminants du crâne. De plus, ces cinétiques sont homogènes pour chacune de ces structures (class score $F(12, 13992)$= $p=0.379$; surface chaude $F(12 , 13992)$= $p=0.530$).

Notre outil s'est donc révélé plus performant que les méthodes céphalométriques et procustéenne utilisées jusqu'à présent pour comparer les profils crâniens. Liberman (1999) a déclaré que le défi de l'étude craniofaciale intégrée était de trouver des outils capables d'évaluer l'interaction entre les structures associées. Pour cet auteur, "il est difficile de définir ou d'évaluer des hypothèses quantitatives d'intégration craniofaciale car nous en savons si peu sur l'étendue et la nature des nombreuses interactions qui se produisent probablement entre les régions crâniennes". Notre outil apparaît comme une réponse pour capturer les covariations des structures craniofaciales. Plus largement, les cartes de saillance apparaissent comme un nouvel outil d'évaluation morphométrique pour toutes les applications utilisant l'imagerie médicale.

6-5-4.2 Impacte de la rétrognathie mandibulaire sur l'architecture craniofaciale

Comme indiqué ci-dessus, notre méthodologie permet de retrouver les résultats classiques de la littérature pour certaines structures (base crânienne, vertèbres cervicales). Elle permet également d'identifier le sinus frontal (qui a donné lieu à des résultats contradictoires) ainsi que l'os pariétal qui n'avait pas été mis en évidence dans les études précédentes.

Le sinus frontal

La zone frontale reflète la transmission des forces occlusales. Par l'intermédiaire des piliers canines et molaires les forces masticatrices sont transmises à la partie supérieure (frontale) de la face ou est logé le sinus frontal. La classe II par rétrognathie mandibulaire altère la répartition des forces masticatrices et est donc susceptible d'altérer la formation du sinus frontal.

De précédentes études constatent effectivement un degré de corrélation entre la classe II et l'air du sinus frontal (Yassaei, Emani et MirbeigiI, 2018). Mais d'autres études ne retrouvent pas cet effet (Prashar et al., 2012 ; Kapasiawala, Raval et Patil, 2016). Ces dernières études constatent, au niveau de l'air du sinus frontal, un effet significatif de la class III comparé à la Class I ou à la Class II mais ne trouve pas de différence significative entre la class II et la class I.

Une explication possible de ces résultats contradictoires est la faible taille des échantillons vis-à-vis de la grande variabilité anthropologique (respectivement 80, 60 et 116 individus subdivisés en trois classes squelettiques). Notre étude porte sur un échantillon de 2694 individus, diminuant ainsi largement l'effet de la variabilité anthropologique. De plus les études précédemment citées mesuraient la surface du sinus par analyse procucéenne. En explorant l'entièreté de l'espace de pixels notre méthodologie est moins restrictive et ne préjuge pas de la dimension à explorer (surface, diamétres ect). Si l'altération du sinus frontal engendrée par la classe II ne concerne pas la surface mais une autre dimension, les analyses Procuste sont susceptibles de ne pas les saisir contrairement à la méthodologie CNN que nous avons utilisé.

L'os pariétal

L'os pariétal n'avait encore jamais été mis en évidence ni même suspecté comme faisant partie du schéma architectural pathologique de la classe II. Les méthodes céphalométriques et procucéenne nécessitant de définir à l'avance les structures d'intérêt à mesurer n'ont donc pas permis d'identifier cette structure (cette structure n'était pas jugée a priori comme pertinente). Comme nous l'avons indiqué, le CNN analyse simultanément l'ensemble de l'espace de pixels de l'image, et ne nécessite donc pas de prédéfinir les structures d'importances. C'est certainement la raison pour laquelle cette méthodologie a permis d'identifier cette nouvelle structure en lien avec la class 2.

De plus cette structure mise en évidence par le réseau est pertinente dans la mesure où elle s'inscrit dans une logique biomécanique du crane.

Comme indiqué précédemment la class 2 par rétrognathie mandibulaire a pour conséquence une modification des forces masticatrices. Deux muscles relient la mâchoire à l'os pariétal :

- Premièrement les muscles temporaux connectent la mandibule à la voute du crane, plus précisément à l'os pariétal. Donc une altération de la mandibule (telle que la classe II par rétrognathie mandibulaire) modifie les forces communiquées a l'os pariétal et est donc susceptible d'altérer sa formation (Thilloy, 1982 ; De Brondeau, 2008).

- Selon le même principe que la voie du muscle temporal, une seconde voie plus indirecte passerait, elle, par les muscles ptérygoïdes qui transmettraient des forces masticatrices de la mandibule au sphénoïde. Le sphénoïde est la structure centrale sur laquelle s'établi l'ensemble de la voute crânienne (dont fait parti l'os pariétal). Donc, une altération de la mandibule (telle que la classe II par rétrognathie mandibulaire) modifie les forces masticatoire communiquées au sphénoïde et est donc susceptible d'altéré sa forme et, d'affecter ainsi l'édification de la voute. De plus, nos résultats en fonction de la sévérité de la rétrognatie montrent que le sphénoïde est dysmorphique dès les cas les moins sévères, alors que l'impact sur le pariétal est plus tardif. Cela renforce l'hypothèse d'une implication du sphénoïde en amont du pariétal dans la mise en place du schéma pathologique.

Cette implication du sphénoïde permet d'appuyer le model développemental « sphenoide shortning » proposé par Lieberman (1998). Notamment, le schéma mis en évidence par notre carte de saillance recouvre le schéma proposé par Lieberman (1998) qui place le sphénoïde comme élément central du réarrangement de l'architecture craniofaciale. Cette théorie met en jeu l'élément mandicateur avec la base du crane et prédit un impacte sur le sinus frontal (Figure.55). Comme nous l'avons vu, les résultats contradictoires de l'état de l'art ne permettaient pas de parvenir à un consensus sur une dysmorphie ou non du sinus frontal. Or, Notre carte de saillance montre une dysmorphie concomitante de l'ensemble de ces trois zones (sinus frontal, base du crane et élément mandicateur), et confirme un impact sur le sinus frontal, ce qui apporte un argument important en faveur du modèle développemental proposé par lieberman

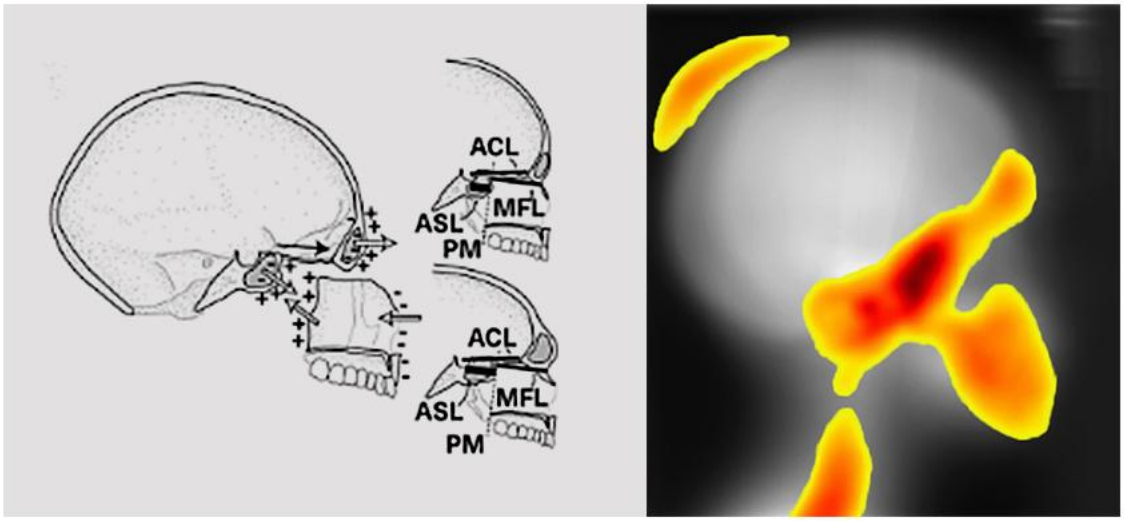


Figure 55 : “Sphenoid shortening” et évolution de la forme de la base du crane des Humains modernes Lieberman (1998).

De plus, notre analyse de cinétique en fonction de la sévérité de la rétrognatie (figure 53) montre une direction préférentielle d’amplification de la dysmorphose crânienne, avec un comportement d’ensemble qui suggère une interaction entre ces trois acteurs (élément mandicateur, base du crane comprenant le sphénoïde, et sinus frontal).

Cette interaction est également abordée par Lieberman (1998) qui postule une influence a double sens : à la fois de la base du crane sur le développement maxillo-mandibulaire au travers du sphénoïde, et réciproquement de la zone maxillo-mandibulaire sur la base du crane mais dans une moindre mesure, le tout réorganisant l’ensemble de l’architecture craniofaciale. Notre cas d’application sur la classe II par rétrognathie mandibulaire permet de mettre en évidence des bases mécanistiques sous-jacentes en visualisant la manière dont une modification des contraintes occlusales peut impacter le schéma spino-cranio-facial et établir un nouveau morphotype.

6-5-4.3 Implication anthropologiques

Les changements évolutifs sont sous tendu par des changements microévolutif (Von cramon-tautadel, 2014) qui constituent des modifications ayant lieu au sein d’une population donnée sur quelques générations. Leur compréhension permet de mettre en évidence les forces évolutives et d’éclairer les mécanismes en œuvre dans la mise en place de nouveaux morphotypes. Dans ce

cadre, de nombreux auteurs ont tenté d'expliquer pourquoi, alors que les gènes des os crâniens sont relativement stables, la capacité crânienne augmente considérablement au fur et à mesure de l'émergence des nouveaux homininés.

Les travaux de Bjork et Enlow placent la croissance cérébrale comme élément moteur de la réorganisation de l'architecture craniofaciale au cours du développement des hominides. Selon ces auteurs, l'augmentation du volume cérébrale engendrerait une pression dans le neurocrâne conduisant à un plissement de la base du crâne et à une globularisation de la voûte du crâne (via la dure-mère). Et c'est cette flexion de la base du crâne qui serait également à l'origine de la diminution de la projection de la face.

La « masticatory-functional hypothesis » nuance cette théorie en introduisant la mastication comme mécanisme modulateur de l'édification de la partie inférieure de la face. Dans cette théorie la croissance cérébrale expliquerait la croissance de la boite crânienne (comme dans les travaux de Bjork et Enlow), mais la mastication influencerait le développement de la face inférieure et dans une moindre mesure le développement de la base du crâne antérieure. La mastication aurait donc une influence locale sur la partie antérieure du crâne.

Dans ce cadre, nos résultats montrent que la classe II par rétrognathie mandibulaire, qui est une pathologie des mâchoires, a un impact sur la face inférieure (mâchoires et le massif facial) mais également sur la voûte du crâne au niveau de l'os pariétal. Nos travaux permettent d'apporter un éclairage nouveau et complémentaire sur le débat concernant le développement de la forme du crane au cours de l'évolution des hominidés. Ils permettent effectivement de montrer que la mastication peut avoir un impact bien plus général que ne le prévoyait la « masticatory-functional hypothesis » et que cet impact concerne à la fois la face inférieure (comme l'indique cette théorie) mais également dans une certaine mesure, le neurocrâne.

Cette influence masticatrice soumet donc le phénotype craniofaciale a des facteurs locaux de modulation due au terroir et aux comportements alimentaires. Cette hypothèse pourrait rendre compte de certaines limites explicatives des modèles classiques basés sur la morphologie ou sur la génétique en reconsidérant le rôle des facteurs locaux dans l'établissement des modèles phylogéniques.

A titre d'exemple, le cas récent du passage du mésolithique au néolithique a vu émerger une gracilisation du morphotype craniofaciale en l'absence de modification génétiques. En revanche ce passage s'est caractérisé par une transition d'un mode de vie de chasseur cueilleur a agropastoral. Un changement de régime alimentaire en a résulté avec une sophistication de la préparation des aliments (broyage cuisson trempage lixiviation) et une consistance nettement plus molle que celle des chasseurs cueilleurs. Un morphotype gracile a émergé se caractérisant notamment par un aplanissement de la face et une augmentation de la capacité crânienne. La concomitance du morphotype gracile et du changement alimentaire avait déjà amené certains auteurs à suggérer l'hypothèse de l'influence masticatrice, notamment sur la projection de la face mais cette hypothèse n'expliquait pas les changements de capacité crânienne. Nos résultats apportent une voie mécanistiques sous-jacente plus complète (incluant la voute du crane) à ce changement micro évolutif et soulève l'importance des facteurs locaux dans l'expression d'une morphologie locale.

Sur des échelles de temps plus anciennes, l'implication du pariétal et du sphénoïde que nous avons constaté s'inscrit dans l'évolution macroévolutive des hominides. En effet le passage de l'australopithèque a homo habilis puis erectus et sapiens s'est traduit conjointement par une diminution du bourrelet suborbitaire adjacent au sinus frontal), par un dégagement pariétal et une rétrusion de la face. A notre échelle nous avons mis en évidence cette même cinétique aux niveaux micro-évolutifs induit par l'activité masticatoire. Il se trouve que ces nouveaux hominides ont émergés suite à des modifications environnementales qui ont profondément modifiés le biotope et donc l'alimentation de ces individus. L'étendu de la modulation que nous avons démontrée confère une plasticité d'adaptation de l'ensemble du crane a ces changements alimentaires et pourrait être un vecteur de l'émergence de nouveau chainon.

6-5-5 Conclusion

A partir d'un cas d'application (classe II par rétrognathie mandibulaire), notre étude introduit une nouvelle méthodologie d'analyse morphologique à base d'imagerie médicale reposant sur les réseaux de neurones artificiels. La classe II par rétrognathie mandibulaire se traduit par une face rétrusive et une altération des contraintes occlusales notamment masticatrices. Or, ces éléments sont au centre de l'évolution paléoanthropologique, ce qui fait de cette pathologie un model pertinent d'étude de l'évolution des hominides. Enfin, cette pathologie est la plus fréquente en

orthodonthie ce qui facilite la collecte d'un nombre extrêmement important de radiographies destinées à constituer la base d'apprentissage du CNN.

Tout d'abord, notre réseau de neurones convolutifs (CNN) s'est avéré capable de classifier les radiologies des patients atteint ou non de la pathologie avec une performance de 95%. Cette performance élevée a permis de générer une carte de saillance via la méthode de visualisation Score-CAM en vue de restituer les caractéristiques morphologiques utilisées par le réseau pour discriminer les cas pathologiques. Une carte de saillance globale de la pathologie a ainsi pu être générée et le patron (bio-marqueur) de la pathologie a pu être extrait après élimination de la variabilité inter-individuelle. Cette carte offre une représentation inédite des structures osseuses impactées par la malocclusion. D'une part, notre approche confirme l'atteinte de la base du crane et des vertèbres cervicales déjà attestées dans la littérature scientifique. D'autre part, nos résultats indiquent également une atteinte du sinus frontal (structure donnant lieu a des résultats contradictoires) ainsi que de l'os pariétal (impact non identifié par les études précédentes).

En anthropologie, les études intraspécifiques visent à déterminer si différentes régions sont plus ou moins susceptibles d'évoluer de manière coordonnée ou indépendante. Notre résultat concernant la variation modulaire de différentes structures (sinus frontal, os pariétal, vertébres cervicales, base du crane) en fonction de la sévérité de la rétrognatie mandibulaire fournit une base de discussion à l'hypothèse de l'influence de la mastication dans l'émergence de caractères dérivés locaux et la résurgence de caractère. Plus exactement, le sinus frontal (via le torus supra orbitaire qui le juxtapose) et l'os pariétal sont tous deux des marqueurs de rusticité paléoanthropologique. La mise en évidence de leur dysmorphie suite à une malocclusion permet de reconsidérer la théorie masticatoire avec une influence sur l'architecture craniofaciale plus vaste que celle habituellement admise, plaidant ainsi en faveur de l'hypothèse d'un rôle des modifications des habitudes alimentaires dans la gracilisation de l'architecture craniofaciale et l'élaboration de morphotypes régionaux.

Il faut noter que notre étude porte sur des vues sagittales médiales et n'offre donc pas une vue complète du crâne. Notamment les ailes du sphénoïde (particulièrement impliquées dans les modifications de l'architecture craniofaciale) n'apparaissaient pas sur téléradiographie de profil et l'os pariétal se déploie majoritairement en vue transversale. Il serait donc particulièrement intéressant de compléter notre étude par une analyse d'images du crâne en vues tranversales. Quoi

qu'il en soit, l'approche méthodologique novatrice que nous proposons apparait comme pertinente pour extraire et visualiser les systèmes morphologiques complexes et identifier, à l'avenir, les biomarqueurs de nombreuses situations pathologiques à partir de bases de données d'imagéries médicales.

7 Discussion générale

En réponse au défi soulevé par Michelle Serres de « permettre à nos cerveaux de travailler efficacement avec l'outil numérique », cette thèse aborde, sous l'angle des sciences cognitives, la question de l'optimisation de l'interface entre les nouvelles technologies et le praticien, en s'appuyant sur l'orthopédie dento-faciale comme domaine d'application. Elle interroge notamment l'impact de la transition numérique et digitale sur la perception des données paracliniques par l'orthodontiste et, par conséquent, sur le processus de prise de décision.

Notre travail s'est articulé en deux axes : le premier étudie l'impact de la transition numérique sur la perception et l'utilisation par l'orthodontiste de modèles virtuels en 3D des arcades dentaires. Le second explore la manière dont l'interprétabilité des réseaux de neurones convolutifs peut contribuer à une meilleure compréhension des images médicales et peut apporter aux praticiens des informations nouvelles en termes de diagnostic et de suivi thérapeutique. Ici le secteur d'application concerne les interactions entre les différentes régions anatomiques crânio-faciales impliquées dans la rétrognathie mandibulaire. Cette dernière application permet également de discuter de l'évolution de la forme du crâne au cours du développement des hominides.

7-1-1 Flux de travail numérique et sens clinique

Ce travail a fait l'objet d'un article épistémologique qui a introduit l'importance des sciences cognitives comme catalyseur entre les nouvelles technologies et l'orthodontiste. Il a permis de mieux connaître les biais intuitifs pour reconnecter l'intuition du praticien au flux de travail numérique.

En orthodontie, Andreiko (1994) introduit le Computer-aided Design (conception assistée par ordinateur) / Computeur-aided Manufacturing (fabrication assistée par ordinateur) dans le cadre du projet Elan de l'entreprise Ormco, qui voit émerger le développement des premiers appareils orthodontiques. Cet auteur souligne la grande variabilité interindividuelle de l'homme moderne et pose le constat que l'appareillage ne peut pas être conçu par une simple logique algorithmique mais doit être fait sur mesure. Cela pose le défi de la médecine personnalisée (Mathieu et al., 2013). Dans la continuité, Makaremi (2016) formalise le flux de travail de l'orthodontiste en y intégrant la sphère cognitive.

Contrairement à la machine, le praticien n'a pas besoin d'expliciter sa représentation des actions à réaliser pour établir son plan de traitement. Partant de l'observation des données cliniques et paracliniques, le praticien élabore intuitivement le plan de traitement à l'aide de son sens clinique. Ce sens clinique est intuitif et non paramétrique. Mais, du fait de la transition numérique, le praticien est aujourd'hui confronté à de nombreuses données surlesquels des solutions algorithmiques de traitement sont déjà proposées (Hennessy et Al-Awafhi, 2016). Mais, ce processus de planification automatisée proposé par les industriels diffère fondamentalement de l'approche intuitive du praticien. En effet, l'élaboration du plan de traitement par le praticien peut être vue comme matrice d'analyse qui pondère des éléments de natures différentes (potentiels de croissance, disposition présente, projection d'incidence d'acte comme l'extraction de prémolaire, etc.), en lien avec les attentes du patient. L'état de l'art montre qu'il est pour l'heure impossible d'apprendre à l'algorithme un «langage commun» permettant de pondérer ces différents éléments. Cela pose la question des limites des outils numériques mais également, en parallèle, de la nature de l'expertise humaine.

7-1-2 Base cognitive de l'expertise

Bilalic et al. (2010) ont décrit l'expertise comme un processus cognitif complexe non explicite, qui repose sur l'intuition. En orthodontie cette intuition est le sens clinique du praticien. La théorie des Chunks (Miller, 1956) formule que les unités d'informations ne sont pas mémorisées individuellement mais par regroupement au sein d'unités signifiantes. Peu importe le nombre d'éléments contenus dans le chunk, il sera retenu comme une unité (Ehrlich, 1972). Johnson (1970) propose une description des chunks dont l'un des traits «le processus de décodage» s'avère particulièrement impliqué dans l'expertise. Ce processus de décodage met en jeu la reconnaissance d'une configuration donnée, préalablement engrammée à la suite d'un processus de formation. L'activation de cette configuration permet d'établir plus vite la stratégie. Dans ce contexte la dimension perceptive, mais surtout les processus descendants (des connaissances vers la perception) apparaissent donc essentiels pour l'expression de l'expertise. Considéré sous le prisme de la psychologie, l'expertise est décrite par le modèle CHREST (Gobet, 1993) qui souligne les liens entre la perception et la mémoire à long terme. Dans son ensemble, ce travail bibliographique souligne l'importance de la perception (et de l'influence des connaissances sur la perception) dans l'expertise et a donc motivé la suite de nos travaux.

7-1-3 Étude de l’estimation des distances sur des modèles d’arcades dentaires en réalité virtuelle

L’estimation des distances est au cœur de la pratique orthodontique; elle constitue le point de départ de l’expression du sens clinique. Ces estimations permettent de saisir l’essence du problème à traiter et d’établir un plan de traitement et d’intervention thérapeutique. La transition numérique rend donc crucial de tester la fiabilité des estimations réalisées en environnement virtuel.

La perception des distances est généralement divisée en deux catégories. Egocentrique (entre l’observateur et l’objet) et exocentrique (entre deux objets). Peillard, Argelaguet et al., (2019) ont montré que la distance égocentrée est sous-estimée en environnement virtuel tandis qu'à l'inverse la distance exocentrique est surestimée. De plus, l'immersion a été mise en évidence comme influent l’estimation. Un haut sentiment d’immersion diminue la surestimation des distances exocentriques (Naceri et al., 2010 ; Waller, 1999). Les modèles orthodontiques 3D virtuels sont des environnements virtuels non immersifs, c'est-à-dire des environnements dans lesquels "l'utilisateur est confronté à un simple écran et à quelques dispositifs qui lui permettent d'interagir avec l’environnement virtuel” (Chellali, 2009).
Ils ne bénéficient donc pas de cette diminution de l’erreur d’estimation (liée à l’immersion).

De plus, Watenberg et Wilborg (2003) ont observé que le taux d'erreur augmente lorsque la distance à estimer devient plus grande dans les environnements virtuels, un effet amplifié par l'interaction avec un environnement non immersif. Ainsi, dans les environnements virtuels non-immersifs, l'estimation de la distance est erronée et cette erreur serait plus importante que dans des environnements immersifs.
Pour tester cela nous avons donc évalué l’estimation des distances réalisées sur des arcades dentaires virtuelles 3D à l'aide de 7 paramètres habituellement estimés par les orthodontistes. Une comparaison avec des modèles en plâtre a également été réalisée. Enfin, des modèles de 3 niveaux de complexités différents ont également été proposés aux participants.

Globalement, les distances étaient surestimées quelle que soit la situation. Mais, la surestimation était plus importante en mode virtuel (M = 3.37) qu'en mode réel (M = 2.67). Pour ce qui est des niveaux de complexité, la surestimation était plus importante dans la condition « facile » (M = 3.58), diminuait dans la condition « moyenne » (M = 3.23), et atteignait son niveau le plus bas dans la condition « difficile » (M = 2.44). De plus, l’interaction était due au fait que les différences entre

mode réel et mode virtuel (des surestimations plus importantes ont été observées avec des modèles virtuels) disparaissaient en situation « difficile ».

Les distances sont surestimées tant en mode réel qu'en mode virtuel. Ces résultats concordent avec ceux obtenus par d'autres chercheurs. Plusieurs études ont confirmé une surestimation de la distance de l'encombrement mandibulaire lorsque l'évaluation avait lieu au moyen de la visualisation de modèles réels par des orthodontistes (Johal et Battagel, 1997 ; Naish et al., 2016 ; Wallis et al., 2014). D'une part, une expérience menée par Makaremi (2021) et d'autre part au sein d'un groupe de 31 orthodontistes a montré une surestimation des distances au niveau de l'encombrement mandibulaire sur les modèles virtuels.

De plus, les résultats que nous avons obtenus montrent que les distances étaient moins bien perçues avec les modèles virtuels qu'avec les modèles en plâtre.

Nous avons alors proposé une explication reposant sur l'effet des indices de profondeur non picturaux sur l'estimation de la distance. Les modèles en plâtre permettent un accès naturel aux différents indices de profondeur non picturaux tandis que, dans le cas des modèles virtuels, présentés ici à l'aide d'un écran d'ordinateur, ces indices peuvent être absents ou contradictoires. Une telle hypothèse reste évidemment à confirmer par des études ultérieures. Nous avons également proposé de tester l'influence des capacités visuo-spatiales (qui affecterait les performances d'estimation sur des modèles virtuels mais non sur des modèles en plâtre (Wainman et al., 2020) sur l'estimation des distances.

Un autre résultat de notre étude était que la charge mentale était plus élevée lorsque les estimations étaient réalisées sur un écran d'ordinateur que lorsqu'elles étaient réalisées sur des modèles en plâtre. Ce résultat est en accord avec des études ayant montré une hausse de la charge mentale dans des tâches réalisées en milieu virtuel (Koumaditis et al., 2020 ; Stevens et al., 2016). En effet, plusieurs auteurs affirment que l'interaction avec un environnement virtuel, du fait de sa nature complexe, pourrait augmenter la charge cognitive (Fraser et al., 2012 ; Huang et al., 2020).
Nos résultats indiquent également que plus le niveau de complexité augmente, plus la charge mentale augmente, que ce soit pour les modèles en plâtre ou pour les modèles virtuels. Ces résultats sont également en accord avec la littérature montrant qu'une charge mentale élevée est

associée à une complexité élevée de la tâche (Chao et al., 2017 ; Das et al., 2020 ; Jacko et Ward, 1996 ; Li et Kim, 2021).

Enfin, nous montrons également que l'augmentation de la complexité des modèles (virtuels et réels) était associée à de meilleures estimations des distances pour l'encombrement mandibulaire et la distance intercanine tandis qu'aucune différence d'estimation selon la complexité de la tâche n'a été observée pour la distance intermolaire. Nous avons expliqué ce résultat en terme d'enjeux thérapeutiques de l'estimation des distances réalisée. En effet, l'encombrement mandibulaire et la distance intercanine jouent un rôle important dans le diagnostic et, par conséquent, dans le traitement qui sera mis en place, tandis que la distance intermolaire est une condition qui ne se présente pas dans la pratique professionnelle. Des auteurs ont démontré que les objectifs de la tâche médient la relation entre performance et complexité de la tâche. Des objectifs importants dans des tâches complexes peuvent améliorer la performance grâce à des processus motivationnels et cognitifs contribuant au développement de stratégies efficaces (Campbell, 1988 ; Wood et al., 1987). La encore, ce résultat, important en terme de pratique clinique reste à confirmer et à expliquer dans le cadre d'études ultérieures.

7-2-1 Vision par ordinateur et expertise médicale

Foucarde et al. (2019) ont introduit la vision par ordinateur comme un troisième œil pour le praticien, permettant d'élargir et d'améliorer sa perception des données paracliniques.
L'émergence de l'imagerie médicale engendre une augmentation sans précédent des données paracliniques. Ces données offrent une transparence et une richesse d'information sur le patient.

Ces données permettent d'obtenir plus d'information sur le patient mais le volume d'information à traiter devient un défi pour le praticien. De plus, le praticien se retrouve limité par ses capacités cognitives. Il apparaît donc nécessaire de développer des outils capables d'exploiter, de synthétiser ces informations et de les mettre à disposition du praticien pour optimiser sa prise de décision.

7-2-2 Elaboration d'un CNN de classification de maturation osseuse

Dans un premier temps, nous nous sommes intéressés à aider le praticien, par un système numérique (CNN), à déterminer le pic de croissance osseuse. En fait, la quantité et la direction de la croissance maxillo-faciale peuvent être modifiées dans une certaine mesure par l'orthodontiste.

Cette possibilité constitue un des principaux leviers thérapeutiques en orthopédie dento-faciale, notamment pour stimuler la croissance mandibulaire (dont le sous développement est la principale pathologie traitée en orthodontisterie (rétrognathie mandibulaire). Cependant, cette intervention (appareil dentaire permettant une activation de croissance) ne peut être appliquée que pour une durée de six mois (au-delà, des effets parasites surviennent). Pour maximiser l'effet, il est donc crucial que ces six mois de traitement soit appliqué dans une période de forte croissance osseuse. Si cette période est dépassée, la correction ne sera plus orthopédique mais chirurgicale. La détermination du pic de croissance osseuse est donc un enjeu primordial pour une bonne prise en charge orthopédique.

En cabinet orthodontiste la principale imagerie effectuée sur les patients est la téléradiographie de profil. Cette téléradiographie englobe la face, le crâne et les vertèbres cervicales. Plusieurs études (Jaqueira et al., 2010 ; Elhaddaoui et al., 2014) ont établi qu'il était possible de situer le patient par rapport à son pic de croissance pubertaire en se basant sur la morphologie des vertèbres cervicales. Mais cette reconnaissance des différents stades de croissance cervicale nécessite une expertise spécifique (Uysal et al., 2006).

Les récents succès des réseaux de neurones en traitement d'imagerie au rayon-x ((Bany et Yeasin, 2021 ; Karargyris et al., 2019) rendent leur emploi pertinent pour former un automate à la réalisation de cette tâche comme aide au diagnostic. Nos résultats indiquent que notre réseau est parvenu à une performance d'identification du stade de croissance osseuse satisfaisant pour une utilisation clinique avec un taux de précision équivalent à l'état de l'art d'autres cas d'application médical (Antoniadi et al., 2021) et la performance d'estimation du stade de maturation vertébrale par les experts (Uysal et al., 2006). Ce résultat, extrêmement prometteur, confirme clairement que des outils numériques type CNN peuvent constituer des aides précieuses au diagnostic à destination des cliniciens.

En revanche, le réseau a été formé sur un échantillon provenant d'une même unité de soin. Cette provenance commune engendre un sur apprentissage spécifique à cette clinique. Afin de rendre le réseau valide pour une application plus générale il sera nécessaire de fournir au réseau des exemples d'apprentissages provenant de multiples sources.

De plus, cet outil ne doit pas être perçu comme se substituant à la compétence du praticien, mais comme un outil d'aide au diagnostic pour que ce dernier puisse se concentrer sur l'essentiel, c'est-à-dire la prise de décision. Cette démarche vise à intégrer l'efficacité du CNN dans une stratégie

d’augmentation, c'est-à-dire une stratégie permettant le meilleur rapport possible entre le praticien et l’IA pour parvenir à l'optimum de soin.

7-2-3 Réseau de neurones convolutionnels et expertise médicale

L’utilisation des procédés de vision par ordinateur basés sur des réseaux de neurones convolutionnels constitue la principale utilisation de l’intelligence artificielle en médecine (Jiang et Al, 2017). Plus particulièrement, les CNN connaissent un grand essor dans le système d’aide à la décision clinique (Antoniadi et al., 2021). Dans leur méta-analyse, Liu et al. (2019) ont constaté que les CNN se sont avérés plus performants que les praticiens pour ce qui est de la classification des pathologies (Liu et al., 2019). La performance des CNN frôle désormais le taux de Bayes (Lin et al., 2014), un seuil absolu qui limite les capacités d’apprentissages tant des humains que des machines, en matière d’extraction d’information. Ces performances supérieures à l’humain pour la classification des pathologies signifient donc que les réseaux de neurones tirent des bases d’imagerie, un savoir supérieur aux experts humains. Ce savoir supplémentaire pourrait enrichir la compréhension des pathologies et la pratique clinique. Plus exactement, appliqué sur une image, les réseaux de neurones conventionnels agissent comme des extracteurs de caractéristiques. Au fil des couches du réseau, les caractéristiques extraites sont de plus en plus haut niveau. La dernière couche condense l’ensemble des caractéristiques de l’image extraites par les réseaux. Et ces caractéristiques lui permettent de prédire la classe de l’image. Tout l'enjeu est donc d’accéder à ces caractéristiques et à les rendre compréhensibles et accessibles au praticien. Quand le réseau classifie, il fournit uniquement sa prédiction de classe sans expliciter les éléments sur lesquels il se base pour réaliser sa classification. Ce problème a été explicité sous le nom de « Black Box problem » (problème de la boite noire) (Watson et al., 2019). Pour y remédier, en parallèle de l’essor des réseaux de neurones convolutionnels, des techniques d’interprétabilités ont été développées pour restituer le raisonnement interne du réseau.

En imagerie, Zeiler et Fergus (2013) ont montré qu’il est possible, en partant de la dernière couche du réseau (où sont stockées les caractéristiques extraites), de remonter aux pixels de l’image qui portent cette caractéristique. Dès lors, Zhou et al., (2016) ont introduit la « classe activation mapping » qui permet de pondérer l’importance de chaque pixel dans la prédiction de la classe sous forme de carte de saillance. C’est le point de départ de multiples méthodes de carte de saillance. Ces cartes de saillances ont pour principales caractéristiques de montrer les zones

d’intérêt qui ont conduit le réseau à prendre sa décision de classification (acronyme ROI pour Région Of Interest).

Dans notre dernière étude, nous avons proposé une utilisation innovante des cartes de saillance en générant une carte de saillance unique représentative de l’ensemble des individus de l’échantillon pathologique que nous avons exploré. Cette superposition a été rendue possible par une sélection stricte des données où la position du crâne est dans une position standardisée. Cette approche est novatrice dans la mesure où habituellement les réseaux sont formés sur des bases de données les plus variables possible pour éviter le sur ajustement et maximiser la généralisation.

De façon standard, les cartes de saillances sont utilisées pour vérifier, sur un individu donné, que le réseau de neurone s’est concentré sur des zones définies à priori comme pertinentes (ex. le poumon pour détecter une infection respiratoire). Dans notre utilisation, la superposition des cartes de saillances de l’ensemble des individus (utilisation d’une position standard de la tête sur des téléradiographies pour permettre cette superposition) a permis d’identifier l’ensemble des structures impliquées dans la pathologie considérée (ici la rétrognathie mandibulaire). C’est cette utilisation, par superposition et par identification des structures pertinentes, qui constitue la dimension novatrice de notre approche.

La carte que nous avons générée permet donc d’isoler des structures osseuses impliquées dans le schéma pathologique de la rétrognathie mandibulaire. Ces structures rejoignent les structures déjà attestées dans la bibliographie (pour la base du crâne et les vertèbres cervicales). Nous avons de plus identifié le sinus frontal (structure fortement discutée dans la littérature). La carte a mis également en évidence une structure qui n’avait jamais été identifiée (l’os pariétal). De plus nous avons restitué de manière innovante les cartes de saillances pour différents degrés de sévérité de la pathologie. Cette représentation a mis en évidence une cinétique d’intensification du dysmorphisme de ces structures en fonction de la sévérité. Ce patern de structure mis en évidence permet d’appréhender l’étendu du schéma pathologique sur l’ensemble de la face, de la base du crâne et jusqu'à la voûte du crâne. Ces résultats nous ont permis de discuter les théories morpho développementales du crâne humain, notamment de proposer/confirmer l’influence de la mastication au-delà des standards admis par la théorie classique appelée « masticatory-functional hypothesis » (hypothèse de mastication fonctionnelle). En effet, la convexité et la hauteur de l’os pariétal et le sinus frontal sont des éléments centraux de l’évolution du crâne des Hominides. La

mise en évidence, dans notre étude, de l'influence de la zone occlusale (mâchoire) sur ces structures (os pariétal et sinus frontal) permet de considérer l'impact des changements masticatoires (induits par les changements alimentaires) comme un élément majeur de l'évolution morphologique des Hominides.

8 Conclusion :

L'intégration des outils numériques dans la pratique clinique de l'orthodontiste est un défi crucial induit par la digitalisation du flux de travail. Les sciences cognitives permettent de prendre en compte la cognition de l'expert confronté à ces nouvelles technologies numériques, dans le cadre d'une stratégie d'augmentation. Cette dernière ne peut réussir qu'à la condition que les traitements cognitifs mis en œuvre par l'expert restent connectés au flux du processus numérique. C'est dans cet enjeu que s'est inscrit cette thèse.

Si les modèles numériques 3D ont permis de s'affranchir des modèles physiques, cette transition questionne sur l'expertise mise en œuvre sur ces environnements virtuels. Cette thèse met en lumière les limites de ces environnements virtuels qui affectent les distances perçues et augmente la charge mentale. Ces limites doivent être considérées par le praticien pour garantir l'optimum thérapeutique. D'autre part, ces limites devraient être prises en compte par les industriels pour le développement des futurs outils en envisageant, par exemple, d'augmenter l'immersivité et la plurimodalité des environnements virtuels proposés afin de diminuer les biais perceptifs/cognitifs.

Les progrès des sciences numériques ouvrent également de nouvelles perspectives pour la perception et la compréhension des différentes pathologies. Nous avons, pour notre part, établi que l'interprétabilité des réseaux de neurones convolutifs permet une analyse novatrice des dysmorphoses crânio-faciales, par exemple, en établissant les interactions des différentes régions anatomiques crânio-faciales en fonction de la sévérité de la rétrognathie mandibulaire. Les perspectives de l'utilisation de l'interprétabilité des réseaux de neurones comme une approche complémentaire pour mieux appréhender tant l'imagerie maxillo-faciale que les dysmorphoses crânio-faciales est donc un domaine très prometteur.

Ainsi, les sciences cognitives permettent de prendre en compte la cognition de l'expert, à laquelle les nouvelles technologies numériques sont destinées, et ceci selon une stratégie d'augmentation. Nous estimons que cette stratégie d'augmentation ne peut réussir qu'à la condition que l'intuition de l'expert reste connectée au flux du processus numérique.

Liste des figures et tableaux

Figures :

Figure 1 : Conception actuelle de l'outil thérapeutique orthodontique selon le procédé CAD/CAM.

Figure 2 : Le praticien doit connecter le patient à l'outil thérapeutique issu des nouvelles technologies.

Figure 3 : Interactions entre les différentes disciplines constituant les sciences cognitives et l'orthodontie.

Figure 4 : Symbiose entre le praticien et l'intelligence artificielle (ci-après IA).

Figure 5 : Activité cérébrale différentielle lors d'une perception virtuelle (droite) et réelle (gauche) d'après Allami et al. (2008).

Figure 6 : C'est en créant une monnaie commune entre différents paramètres (esthétique, fonctions oro-faciales**,** croissance faciale) que le praticien parvient à l'optimum individuel pour le patient.

Figure 7 : Un workflow (flux de travail) optimisé permettant une symbiose entre le praticien et l'IA dans le procédé CAD/CAM.

Figure 8 : Le workflow numérique.

Figure 9 : Les quatre piliers de l'apprentissage d'après Dehaene (2018).

Figure 10 : Le flux du travail numérique renferme les quatre piliers de l'apprentissage.

Figure 11 : Fonctionnement du modèle d'architecture cognitive CHREST d'après Gobet (1993).

Figure 12 : Schéma des caractéristiques de la réalité virtuelle (ci-après RV).

Figure13 : Schéma des indices de profondeur picturaux d'après Cutting et Vishton (1995).

Figure 14 : Schéma des portées effectives des différents indices de profondeur inspiré de Cutting et Vishton (1995).

Figure 15 : Puzzle en 3D et son modèle virtuel utilisé dans l'étude de Tourancheau et al. (2012).

Figure 16 : Schéma des quatre groupes de facteurs influant sur la perception de la distance et les estimations des distances basées sur Renner et al. (2013).

Figure 17 : Schéma des différentes parallaxes des objets : (a) parallaxe négative, (b) parallaxe nulle et (c) parallaxe positive.

Figure 18 : Schéma du champ visuel d'affichage (CVA) et du champ visuel géométrique (CVG)

Figure 19 : Exemple d'indice de perspective améliorant l'estimation des distances d'après Delucia, 1991.

Figure 20 : Technique de perspective aérienne (Heinrich et al., 2021).

Figure 21 : Technique de profondeur pseudo-chromatique (Heinrich et al., 2021).

Figure 22 : Principes des techniques de renforcement de l'occlusion d'après Ropinski et al. (2006).

Figure 23 : Modèle d'étude orthodontique en plâtre et en 3D numérique.

Figure 24 : Identification de la population participant à l'étude.

Figure 25 : Les 5 cas cliniques utilisés pour l'estimation de la distance.

Figure 26 : Présentation des sept variables à estimer.

Figure 27 : Organigramme de la procédure.

Figure 28 : Processus de fabrication des modèles en plâtre à partir des empreintes numériques des arcades mandibulaires.

Figure 29 : Modèles mandibulaires inclus dans l'étude.

Figure 30 : Questionnaire de charge mentale – NASA-TLX.

Figure 31 : Fiche explicative – mesures.

Figure 32 : Présentation sommaire du protocole expérimental.

Figure 33 : Représentation graphique de la charge mentale selon le niveau de complexité de la tâche.

Figure 34 : Représentation graphique de la moyenne des estimations de l'encombrement mandibulaire, de la distance intercanine et de la distance intermolaire.

Figure 35 : Représentation graphique de l'encombrement mandibulaire.

Figure 36 : Représentation graphique de la distance intercanine.

Figure 37 : Représentation graphique de la performance lors de l'estimation de la distance intermolaire.

Figure 38 : Schéma comparatif : neurone biologique et neurone artificiel, d'après Maurice (2018).

Figure 39 : Schéma d'une organisation convolutive.

Figure 40 : Processus d'extraction des caractéristiques locales et supérieures d'après Jones (2014).

Figure 41 : Schéma d'un réseau de neurones de reconnaissance de chiffres, inspiré de Hoang (2018).

Figure 42 : Couche Pré-Softmax d'un réseau convolutif d'après Suhermi et al. (2018).

Figure 43 : Exemple de carte de saillance d'après Zhou et al. (2016).

Figure 44 : Schéma du processus de Score-CAM d'après Wang et al. (2020).

Figure 45 : Interprétabilité et cognition d'après Kermany et al. (2018).

Figure 46 : Stades CS (cervical stage) radiologiques et morphologiques superposés à la courbe de Croissance de Björk.

Figure 47 : Partie isolée (vertèbres cervicales) d'une téléradiographie de profil.

Figure 48 : Evolution de la fonction de perte et de la précision pour les 1870 images examinées.

Figure.49- Processus de sélection et de labellisation des téléradiographies pour l'élaboration du jeu de données.

Figure.50- Jeu de données : le jeu de données est composé de 1 377 radiographies latérales de classe II. B- Pour chaque radiographie, un prétraitement Sobel a été effectué. C- Les radiographies prétraitées sont utilisées dans un processus de classification DL-CNN d'entraînement-évaluation-test, puis des cartes d'activation sont générées. D- Les cartes d'activation générées de toutes les radiographies sont collectées. E- Une carte d'activation moyenne est générée à partir des 1377 cartes d'activation. F- La zone chaude de la carte d'activation moyenne est extraite. G- Un crâne moyen est généré à partir des 1377 cartes d'activation. H- Enfin, la température chaude de la carte d'activation globale est extraite et superposée au crâne moyen).

Figure 51 : Schéma colorisé sur la base des zones identifiées comme significativement corrélées a la classe II par rétrognathie mandibulaire.

Figure-52. A structures identifiées dans la littérature et B Résultat de notre carte de saillance

Figure-53. La rangée du haut montre, pour chaque intervalle de sévérité, un exemple de crânes. La rangée du bas montre les cartes de saillance pour chaque intervalle par ordre croissant de sévérité. La barre de couleur indique la quantification de la gravité.

Figure 54 : Pour chacune des quatre structures squelettiques (base crânienne, sinus frontal, os pariétal, vertèbres cervicales), le graphique de gauche montre l'évolution du score de classe suivant l'angle ANB (rang). Le graphique d ge droite montre l'évolution de la surface des pixels chauds suivant l'angle ANB (rang).

Figure 55 : “Sphenoid shortening” et évolution de la forme de la base du crane des Humains modernes Lieberman (1998).

Tableaux :

Tableau 1 : Utilisation de la RV en médecine.

Tableau 2 : Présentation des 5 cas cliniques utilisés dans notre étude.

Tableau 3 : Moyennes et écarts types pour chaque modèle. x1 représente la valeur obtenue via le logiciel de mesure, x2 la moyenne des mesures et σ l’écart type.

Tableau 4 : Résultats du test à un échantillon pour toutes les variables. Ce tableau montre la valeur moyenne, l'écart type, la valeur du test t et la signification pour chaque variable. Le tableau du haut présente les résultats avec 0 cm comme valeur de référence. Le tableau du bas présente les résultats avec 1 cm comme valeur de référence (seules les variables qui se sont avérées significatives avec la valeur 0 sont testées avec la valeur 1).

Tableau 5 : Tableau descriptif de la population.

Tableau 6 : Statistiques descriptives de la charge mentale

Tableau 7 : Statistiques descriptives de l’indicateur de l’encombrement, de la distance intercanine et de la distance intermolaire.

Tableau 8 : Solution de logiciel d’IA pour l’orthodontie/dentisterie.

Tableau 9 : Précision de la détermination du stade de maturation vertébrale. L'algorithme d'optimisation utilisé dans le cadre de la formation était ADAM.

Bibliographie

Adams, D. C., Rohlf, F. J., & Slice, D. E. (2004). Geometric morphometrics : Ten years of progress following the 'revolution'. *Italian Journal of Zoology*, *71*(1), 5-16. https://doi.org/10.1080/11250000409356545

Aggarwal, R., Sounderajah, V., Martin, G., Ting, D. S. W., Karthikesalingam, A., King, D., Ashrafian, H., & Darzi, A. (2021). Diagnostic accuracy of deep learning in medical imaging : A systematic review and meta-analysis. *Npj Digital Medicine*, *4*(1), 1-23. https://doi.org/10.1038/s41746-021-00438-z

Agus, M., Bettio, F., Gobbetti, E., & Pintore, G. (2007). Medical Visualization with New Generation Spatial 3D Displays. *Eurographics Italian Chapter Conference*, 6 pages. https://doi.org/10.2312/LOCALCHAPTEREVENTS/ITALCHAP/ITALIANCHAPCONF2007/161-166

Ahn, S., Kim, S., & Lee, S. (2021). Effects of Visual Cues on Distance Perception in Virtual Environments Based on Object Identification and Visually Guided Action. *International Journal of Human–Computer Interaction*, *37*(1), 36-46. https://doi.org/10.1080/10447318.2020.1805875

Ackerman, M. J. (1998). The Visible Human Project. *Proceedings of the IEEE*, *86*(3), 504-511. https://doi.org/10.1109/5.662875

Al-antari, M. A., Al-masni, M. A., & Kim, T.-S. (2020). Deep Learning Computer-Aided Diagnosis for Breast Lesion in Digital Mammogram. In G. Lee & H. Fujita (Éds.), *Deep Learning in Medical Image Analysis* (Vol. 1213, p. 59-72). Springer International Publishing. https://doi.org/10.1007/978-3-030-33128-3_4

Alexandre, L. (2011). *La mort de la mort : Comment la technomédecine va bouleverser l'humanité*. JC Lattès.

Alexandre, L. (2017). *La guerre des intelligences : Comment l'intelligence artificielle va révolutionner l'éducation*. J.C. Lattès.

Alexandrova, I. V., Teneva, P. T., de la Rosa, S., Kloos, U., Bülthoff, H. H., & Mohler, B. J. (2010). Egocentric distance judgments in a large screen display immersive virtual environment. *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization - APGV '10*, 57. https://doi.org/10.1145/1836248.1836258

Allami, N., Paulignan, Y., Brovelli, A., & Boussaoud, D. (2008). Visuo-motor learning with combination of different rates of motor imagery and physical practice. *Experimental Brain Research*, *184*(1), 105-113. https://doi.org/10.1007/s00221-007-1086-x

Al-Sheakli, I. I., Mohammed, S. A., & Taha, S. S. (2013). The frontal sinus dimensions in mouth and nasal breathers in Iraqi adult subjects. *Journal of Baghdad College of Dentistry*, *25*(2), 155-163.

Almeida, K. C. M. de, Raveli, T. B., Vieira, C. I. V., Santos-Pinto, A. dos, & Raveli, D. B. (2017). Influence of the cranial base flexion on Class I, II and III malocclusions : A systematic review. *Dental Press Journal of Orthodontics*, *22*(5), 56-66. https://doi.org/10.1590/2177-6709.22.5.056-066.oar

Alyasseri, Z. A. A., Al-Betar, M. A., Doush, I. A., Awadallah, M. A., Abasi, A. K., Makhadmeh, S. N., Alomari, O. A., Abdulkareem, K. H., Adam, A., Damasevicius, R., Mohammed, M. A., & Zitar, R. A. (2021). Review on COVID -19 diagnosis models based on machine learning and deep learning approaches. *Expert Systems*. https://doi.org/10.1111/exsy.12759

Amirthalingam, J., Paidi, G., Alshowaikh, K., Jayarathna, A. I., Salibindla, D. B. A. M. R., Karpinska-Leydier, K., & Ergin, H. E. (2021). Virtual Reality Intervention to Help Improve Motor Function in Patients Undergoing Rehabilitation for Cerebral Palsy, Parkinson's Disease, or Stroke : A Systematic Review of Randomized Controlled Trials. *Cureus*, *13*(7). https://doi.org/10.7759/cureus.16763

Andreiko, C. (1994). Craig Andreiko, DDS, MS, on the Elan and Orthos systems. Interview by Dr. Larry W. White. *Journal of Clinical Orthodontics: JCO*, *28*(8), 459-468.

Antoniadi, A. M., Du, Y., Guendouz, Y., Wei, L., Mazo, C., Becker, B. A., & Mooney, C. (2021). Current Challenges and Future Opportunities for XAI in Machine Learning-Based Clinical Decision Support Systems : A Systematic Review. *Applied Sciences*, *11*(11), 5088. https://doi.org/10.3390/app11115088

Arik, S. Ö., Ibragimov, B., & Xing, L. (2017). Fully automated quantitative cephalometry using convolutional neural networks. *Journal of Medical Imaging*, *4*(1), 014501. https://doi.org/10.1117/1.JMI.4.1.014501

Armbrüster, C., Wolter, M., Kuhlen, T., Spijkers, W., & Fimm, B. (2008). Depth Perception in Virtual Reality : Distance Estimations in Peri- and Extrapersonal Space. *CyberPsychology & Behavior*, *11*(1), 9-15. https://doi.org/10.1089/cpb.2007.9935

Asiri, N., Hussain, M., Al Adel, F., & Alzaidi, N. (2019). Deep learning based computer-aided diagnosis systems for diabetic retinopathy : A survey. *Artificial Intelligence in Medicine*, *99*, 101701. https://doi.org/10.1016/j.artmed.2019.07.009

Autor, D., Katz, L., & Krueger, A. (1998). Computing Inequality : Have Computers Changed the Labor Market? *The Quarterly Journal of Economics*, *113*(4), 1169-1213.

Ayoub, A. F., Xiao, Y., Khambay, B., Siebert, J. P., & Hadley, D. (2007). Towards building a photo-realistic virtual human face for craniomaxillofacial diagnosis and treatment planning. *International Journal of Oral and Maxillofacial Surgery*, *36*(5), 423-428. https://doi.org/10.1016/j.ijom.2007.02.003

Ayoub, A., & Pulijala, Y. (2019). The application of virtual reality and augmented reality in Oral & Maxillofacial Surgery. *BMC Oral Health*, *19*(1), 238. https://doi.org/10.1186/s12903-019-0937-8

Azer, S. A., & Azer, S. (2016). 3D Anatomy Models and Impact on Learning : A Review of the Quality of the Literature. *Health Professions Education*, *2*(2), 80-98. https://doi.org/10.1016/j.hpe.2016.05.002

Badea, M. S., Felea, I. I., Florea, L. M., & Vertan, C. (2016). The use of deep learning in image segmentation, classification and detection. *arXiv:1605.09612 [cs]*. http://arxiv.org/abs/1605.09612

Baek, S., He, Y., Allen, B. G., Buatti, J. M., Smith, B. J., Tong, L., Sun, Z., Wu, J., Diehn, M., Loo, B. W., Plichta, K. A., Seyedin, S. N., Gannon, M., Cabel, K. R., Kim, Y., & Wu, X. (2019). Deep segmentation networks predict survival of non-small cell lung cancer. *arXiv:1903.11593 [cs, eess]*. http://arxiv.org/abs/1903.11593

Baldominos, A., Saez, Y., & Pozo, C. G. del. (2015). An Approach to Physical Rehabilitation Using State-of-the-art Virtual Reality and Motion Tracking Technologies. *Procedia Computer Science*, *64*, 10-16. https://doi.org/10.1016/j.procs.2015.08.457

Banwari, A., Sengar, N., Dutta, M., & Travieso-González, C. (2016). Automated segmentation of colon gland using histology images. *2016 Ninth International Conference on Contemporary Computing (IC3)*. https://doi.org/10.1109/IC3.2016.7880223

Bany Muhammad, M., & Yeasin, M. (2021). Interpretable and parameter optimized ensemble model for knee osteoarthritis assessment using radiographs. *Scientific Reports*, *11*(1), 14348. https://doi.org/10.1038/s41598-021-93851-z

Barré, J., Michelet, D., Truchot, J., Jolivet, E., Recanzone, T., Stiti, S., Tesnière, A., & Pourcher, G. (2019). Virtual Reality Single-Port Sleeve Gastrectomy Training Decreases Physical and Mental Workload in Novice Surgeons : An Exploratory Study. *Obesity Surgery*, *29*(4), 1309-1316. https://doi.org/10.1007/s11695-018-03680-9

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577-660. https://doi.org/10.1017/S0140525X99002149

Bastir, M., Rosas, A., & Sheets, H. D. (2005). The Morphological Integration of the Hominoid Skull : A Partial Least Squares and PC Analysis with Implications for European Middle Pleistocene Mandibular Variation. In D. E. Slice (Éd.), *Modern Morphometrics in Physical Anthropology* (p. 265-284). Kluwer Academic Publishers-Plenum Publishers. https://doi.org/10.1007/0-387-27614-9_12

Bastir, M., Sobral, P. G., Kuroe, K., & Rosas, A. (2008). Human craniofacial sphericity : A simultaneous analysis of frontal and lateral cephalograms of a Japanese population using geometric morphometrics and partial least squares analysis. *Archives of Oral Biology*, *53*(4), 295-303. https://doi.org/10.1016/j.archoralbio.2007.09.011

Baumrind, S., Korn, E. L., Boyd, R. L., & Maxwell, R. (1996). The decision to extract : Part 1—Interclinician agreement. *American Journal of Orthodontics and Dentofacial Orthopedics*, *109*(3), 297-309. https://doi.org/10.1016/S0889-5406(96)70153-1

Bédard, J. (1991). Expertise and Its Relation to Audit Decision Quality. *Contemporary Accounting Research*, *8*(1), 198-222. https://doi.org/10.1111/j.1911-3846.1991.tb00842.x

Bengio, Y., & Delalleau, O. (2011). On the Expressive Power of Deep Architectures. *Discovery Science*. https://doi.org/10.1007/978-3-642-24477-3_1

Bellot-Arcís, C., Montiel-Company, J. M., Almerich-Silla, J. M., Paredes-Gallardo, V., & Gandía-Franco, J. L. (2012). The use of occlusal indices in high-impact literature. *Community Dental Health*, *29*(1), 45-48.

Benjelloun, L., El Harti, K., & El Wady, W. (2012). Place de la tomographie volumique à faisceau conique en imagerie dento-maxillofaciale. *Actualités Odonto-Stomatologiques*, *258*, 115-126. https://doi.org/10.1051/aos/2012204

Berbaum, K. S., Franken, E. A., Dorfman, D. D., Rooholamini, S. A., Kathol, M. H., Barloon, T. J., Behlke, F. M., Sato, Y., Lu, C. H., & el-Khoury, G. Y. (1990). Satisfaction of search in diagnostic radiology. *Investigative Radiology*, *25*(2), 133-140. https://doi.org/10.1097/00004424-199002000-00006

Berchon, M., & Luyt, B. (2014). *L'impression 3D* (2e éd). Eyrolles.

Berthoz, A., & Petit, J.-L. (2006). *Physiologie de l'action et phénoménologie*. O. Jacob.

Biederman, I. (1987). Recognition-by-components : A theory of human image understanding. *Psychological Review*, *94*(2), 115-147. https://doi.org/10.1037/0033-295X.94.2.115

Bieri, L. S., de Meneses, Y. L., & Jacot, J. (2003). Vision tridimensionnelle par lumière structurée appliquée au contrôle qualité en ligne (No. CONF).

Bilalić, M. (2017). *The neuroscience of expertise*. Cambridge University press.

Bilalić, M., Grottenthaler, T., Nägele, T., & Lindig, T. (2016). The faces in radiological images: Fusiform face area supports radiological expertise. *Cerebral Cortex*, *26*(3), 1004-1014. https://doi.org/10.1093/cercor/bhu272

Bilalić, M., Langner, R., Erb, M., & Grodd, W. (2010). Mechanisms and neural basis of object and pattern recognition: A study with chess experts. *Journal of Experimental Psychology: General*, *139*(4), 728-742. https://doi.org/10.1037/a0020756

Birbara, N. S., Sammut, C., & Pather, N. (2020). Virtual Reality in Anatomy : A Pilot Study Evaluating Different Delivery Modalities. *Anatomical Sciences Education*, *13*(4), 445-457. https://doi.org/10.1002/ase.1921

Bittner, A. C., Byers, J. C., Hill, S. G., Zaklad, A. L., & Christ, R. E. (1989). Generic Workload Ratings of a Mobile Air Defense System (LOS-F-H). *Proceedings of the Human Factors Society Annual Meeting*, *33*(20), 1476-1480. https://doi.org/10.1177/154193128903302026

Björk, A. (1963). Variations in the Growth Pattern of the Human Mandible : Longitudinal Radiographic Study by the Implant Method. *Journal of Dental Research*, *42*(1), 400-411. https://doi.org/10.1177/00220345630420014701

Björk, A., & Helm, S. (1967). Prediction of the age of maximum puberal growth in body height. *The Angle Orthodontist*, *37*(2), 134-143. https://doi.org/10.1043/0003-3219(1967)037<0134:POTAOM>2.0.CO;2

Bookstein, F. L., Gunz, P., Mitterœcker, P., Prossinger, H., Schæfer, K., & Seidler, H. (2003). Cranial integration in Homo : Singular warps analysis of the midsagittal plane in ontogeny and evolution. *Journal of Human Evolution*, *44*(2), 167-187. https://doi.org/10.1016/S0047-2484(02)00201-4

Bouletreau, P., Makaremi, M., Ibrahim, B., Louvrier, A., & Sigaux, N. (2019). Artificial Intelligence : Applications in orthognathic surgery. *Journal of Stomatology, Oral and Maxillofacial Surgery*, *120*(4), 347-354. https://doi.org/10.1016/j.jormas.2019.06.001

Brady, L., Wang, Y.-N., Rombokas, E., & Ledoux, W. R. (2021). Comparison of texture-based classification and deep learning for plantar soft tissue histology segmentation. *Computers in Biology and Medicine*, *134*, 104491. https://doi.org/10.1016/j.compbiomed.2021.104491

Brondeau, F. de. (2015). Les variations de la morphologie de la base du crâne au cours de la phylogenèse : Les différentes hypothèses. *Revue d'Orthopédie Dento-Faciale*, *49*(1), 59-70. https://doi.org/10.1051/odf/2014050

Brondeau, F. de. (2008). *La morphologie de la base du crâne et ses relations avec le maxillaire et la mandibule : Variations et auxologie* [These de doctorat, Bordeaux 1]. https://www.theses.fr/2008BOR13670

Bruder, G., Argelaguet, F., Olivier, A.-H., & Lécuyer, A. (2016). CAVE Size Matters : Effects of Screen Distance and Parallax on Distance Estimation in Large Immersive Display Setups. *Presence: Teleoperators and Virtual Environments*, *25*(1), 1-16. https://doi.org/10.1162/PRES_a_00241

Burdea, G., & Coiffet, P. (1993). *La réalité virtuelle*. Hermès.

Busselle, R. (2020). Perceived Realism. In The International Encyclopedia of Media Psychology, J. Bulck (Ed.). https://doi.org/10.1002/9781119011071.iemp0206

Bruzina, R. 1982 Art and Architecture, Ancient and Modem, *Research in Philosophy and Technology* 5: 163187.

Campbell, D. J. (1988). Task Complexity : A Review and Analysis. *Academy of Management Review*, *13*(1), 40-52. https://doi.org/10.5465/amr.1988.4306775

Campbell, S., & Goldstein, G. (2021). Angle's Classification–A Prosthodontic Consideration : Best Evidence Consensus Statement. *Journal of Prosthodontics*, *30*(S1), 67-71. https://doi.org/10.1111/jopr.13307

Carlson, D. S., & Van Gerven, D. P. (1977). Masticatory function and post-pleistocene evolution in Nubia. *American Journal of Physical Anthropology*, *46*(3), 495-506. https://doi.org/10.1002/ajpa.1330460316

Cederquist, R., & Dahlberg, A. A. (1979). Age changes in facial morphology of an Alaskan Eskimo population. *Ossa*, *6*, 39-67.

Cevidanes, L., Bailey, L., Tucker, G., Styner, M., Mol, A., Phillips, C., Proffit, W., & Turvey, T. (2005). Superimposition of 3D cone-beam CT models of orthognathic surgery patients. *Dentomaxillofacial Radiology*, *34*(6), 369-375. https://doi.org/10.1259/dmfr/17102411

Chan, L. K., & Cheng, M. M. W. (2011). An analysis of the educational value of low-fidelity anatomy models as external representations. *Anatomical Sciences Education*, *4*(5), 256-263. https://doi.org/10.1002/ase.239

Chao, C.-J., Wu, S.-Y., Yau, Y.-J., Feng, W.-Y., & Tseng, F.-Y. (2017). Effects of three-dimensional virtual reality and traditional training methods on mental workload and training performance. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *27*(4), 187-196. https://doi.org/10.1002/hfm.20702

Chanquoy, L., Tricot, A., & Sweller, J. (2007). *La charge cognitive : Théorie et applications (Collection U)*. ARMAND COLIN.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55-81. https://doi.org/10.1016/0010-0285(73)90004-2

Chattopadhay, A., Sarkar, A., Howlader, P., & Balasubramanian, V. N. (2018). Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks. *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 839-847. https://doi.org/10.1109/WACV.2018.00097

Chellali, A. (2009). *Étude des interactions homme-homme pour l'élaboration du référentiel commun dans les environnements virtuels collaboratifs.* [Phdthesis, Université de Nantes]. https://tel.archives-ouvertes.fr/tel-00467441

Chen, X., Lian, C., Deng, H. H., Kuang, T., Lin, H.-Y., Xiao, D., Gateno, J., Shen, D., Xia, J. J., & Yap, P.-T. (2021). Fast and Accurate Craniomaxillofacial Landmark Detection via 3D Faster R-CNN. *IEEE Transactions on Medical Imaging*, *40*(12), 3867-3878. https://doi.org/10.1109/TMI.2021.3099509

Chen, Z., Cao, Y., Zou, D., & Gu, Q. (2021). How Much Over-parameterization Is Sufficient to Learn Deep ReLU Networks? *arXiv:1911.12360 [cs, math, stat]*. http://arxiv.org/abs/1911.12360

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices*. *Cognitive Science*, *5*(2), 121-152. https://doi.org/10.1207/s15516709cog0502_2

Clark, H. M., & Robin, D. A. (1998). Generalized motor programme and parameterization accuracy in apraxia of speech and conduction aphasia. *Aphasiology*, *12*(7-8), 699-713. https://doi.org/10.1080/02687039808249567

Codd, A. M., & Choudhury, B. (2011). Virtual reality anatomy : Is it comparable with traditional methods in the teaching of human forearm musculoskeletal anatomy? *Anatomical Sciences Education*, *4*(3), 119-125. https://doi.org/10.1002/ase.214

Collins, K. L., Russell, H. G., Schumacher, P. J., Robinson-Freeman, K. E., O'Conor, E. C., Gibney, K. D., Yambem, O., Dykes, R. W., Waters, R. S., & Tsao, J. W. (2018). A review of current theories and treatments for phantom limb pain. *Journal of Clinical Investigation*, *128*(6), 2168-2176. https://doi.org/10.1172/JCI94003

Combe, T., Chardonnet, J.-R., Merienne, F., & Ovtcharova, J. (2021). CAVE vs. HMD in Distance Perception. *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 448-449. https://doi.org/10.1109/VRW52623.2021.00106

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances : The integration, relative potency, and contextual use of different information about depth. In *Perception of space and motion* (p. 69-117). Academic Press. https://doi.org/10.1016/B978-012240530-3/50005-5

Daanouni, O., Cherradi, B., & Tmiri, A. (2021). Automatic detection of diabetic retinopathy using custom cnn and grad-cam. In *Advances on Smart and Soft Computing* (pp. 15-26). Springer, Singapore.

Das, S., Maiti, J., & Krishna, O. B. (2020). Assessing mental workload in virtual reality based EOT crane operations : A multi-measure approach. *International Journal of Industrial Ergonomics*, *80*, 103017. https://doi.org/10.1016/j.ergon.2020.103017

Davenport, T. H., & Kirby, J. (2015, juin 1). Beyond Automation. *Harvard Business Review*. https://hbr.org/2015/06/beyond-automation

Dawson, C. W., & Wilby, R. (1998). An artificial neural network approach to rainfall-runoff modelling. *Hydrological Sciences Journal*, *43*(1), 47-66. https://doi.org/10.1080/02626669809492102

De Groot, A. D. (1965). Thought and Choice in Chess. 1978. *The Hague, Netherlands: Mouton Publishers*.

Decety, J., Jeannerod, M., & Prablanc, C. (1989). The timing of mentally represented actions. *Behavioural Brain Research*, *34*(1-2), 35-42. https://doi.org/10.1016/S0166-4328(89)80088-9

Dehaene, S. (2014). *Le code de la conscience*. Odile Jacob.

Dehaene, S. (2018). *Apprendre ! Les talents du cerveau, le défi des machines*. Odile Jacob.

Dehaene, S., & Changeux, J.-P. (2011). Experimental and Theoretical Approaches to Conscious Processing. *Neuron*, *70*(2), 200-227. https://doi.org/10.1016/j.neuron.2011.03.018

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., & Cohen, L. (2010). How Learning to Read Changes the Cortical Networks for Vision and Language. *Science*, *330*(6009), 1359-1364. https://doi.org/10.1126/science.1194140

Delaire, J. (2003). Le développement «adaptatif» de la base du crâne. Justification du traitement précoce des dysmorphoses de classe III. *Revue d'Orthopédie Dento-Faciale*, *37*(3), 243-265. https://doi.org/10.1051/odf/2003019

DeLucia, P. R. (1991). Pictorial and motion-based information for depth perception. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(3), 738-748. https://doi.org/10.1037/0096-1523.17.3.738

Dervin, F., & Abbas, Y. (2009). *Technologies numériques du soi et co-constructions identitaires*. l'Harmattan.

Deshayes, M. J. (1993). Vers un nouveau procede biometrique de modelisation des anomalies de la morphologie cranienne et faciale. *Cahiers d'Anthrop. et Biometrie humaine*, *0*(1), 135-151.

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, *18*(1), 193-222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Dhiman, I., Singla, A., Mahajan, V., Jaj, H. S., Seth, V., & Negi, P. (2015). Reliability of frontal sinus with that of maxillary sinus in assessment of different types of skeletal malocclusions. *Journal of Indian Orthodontic Society*, *49*(2), 96-103. https://doi.org/10.4103/0301-5742.162265

Di Stasi, L. L., Diaz-Piedra, C., Rieiro, H., Sánchez Carrión, J. M., Martin Berrido, M., Olivares, G., & Catena, A. (2016). Gaze entropy reflects surgical task load. *Surgical Endoscopy*, *30*(11), 5034-5043. https://doi.org/10.1007/s00464-016-4851-8

Díaz, J., Ropinski, T., Navazo, I., Gobbetti, E., & Vázquez, P.-P. (2017). An experimental study on the effects of shading in 3D perception of volumetric models. *The Visual Computer*, *33*(1), 47-61. https://doi.org/10.1007/s00371-015-1151-6

Drascic, D., & Milgram, P. (1996). *Perceptual issues in augmented reality* (M. T. Bolas, S. S. Fisher, & J. O. Merritt, Éds.; p. 123-134). https://doi.org/10.1117/12.237425

Duchaine, B., & Yovel, G. (2015). A Revised Neural Framework for Face Processing. *Annual Review of Vision Science*, *1*(1), 393-416. https://doi.org/10.1146/annurev-vision-082114-035518

Dunbar, A., Bearn, D., & McIntyre, G. (2014). The Influence of Using Digital Diagnostic Information on Orthodontic Treatment Planning—A Pilot Study. *Journal of Healthcare Engineering*, *5*(4), 411-428. https://doi.org/10.1260/2040-2295.5.4.411

Duncan, J. (2013). The structure of cognition : Attentional episodes in mind and brain. *Neuron*, *80*(1), 35-50. https://doi.org/10.1016/j.neuron.2013.09.015

Dupuy, J. P. (2013). *Aux origines des sciences cognitives*. La découverte.

Ebrahimi, E., Altenhoff, B. M., Pagano, C., & Babu, S. V. (2015). Carryover effects of calibration to visual and proprioceptive information on near field distance judgments in 3D user interaction. *2015 IEEE Symposium on 3D User Interfaces (3DUI)*. https://doi.org/10.1109/3DUI.2015.7131732

Ebrahimi, E., Hartman, L. S., Robb, A., Pagano, C. C., & Babu, S. V. (2018). Investigating the Effects of Anthropomorphic Fidelity of Self-Avatars on Near Field Depth Perception in Immersive Virtual Environments. *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 1-8. https://doi.org/10.1109/VR.2018.8446539

Ehrlich, S. (1972). *La capacité d'appréhension verbale*. FeniXX. https://www.sorbonne-librairie.fr/ebook/9782705911850-la-capacite-d-apprehension-verbale-stephane-ehrlich/

Eisová, S., Rangel de Lázaro, G., Píšová, H., Pereira-Pedro, S., & Bruner, E. (2016). Parietal Bone Thickness and Vascular Diameters in Adult Modern Humans : A Survey on Cranial Remains: A Survey on Cranial Remains. *The Anatomical Record*, *299*(7), 888-896. https://doi.org/10.1002/ar.23348

Eldan, R., & Shamir, O. (2016). The Power of Depth for Feedforward Neural Networks. *Conference on Learning Theory*, 907-940. https://proceedings.mlr.press/v49/eldan16.html

Elhaddaoui, R., Benyahia, H., Azaroual, F., & Zaoui, F. (2014). Intérêt de la méthode de maturation des vertèbres cervicales (CVM) en orthopédie dento-faciale : Mise au point. *Revue de Stomatologie, de Chirurgie Maxillo-faciale et de Chirurgie Orale*, *115*(5), 293-300. https://doi.org/10.1016/j.revsto.2014.09.003

Ellis, S. R. (1994). What are virtual environments? *IEEE Computer Graphics and Applications*, *14*(1), 17-22. https://doi.org/10.1109/38.250914

England, J. R., Gross, J. S., White, E. A., Patel, D. B., England, J. T., & Cheng, P. M. (2018). Detection of Traumatic Pediatric Elbow Joint Effusion Using a Deep Convolutional Neural Network. *American Journal of Roentgenology*, *211*(6), 1361-1368. https://doi.org/10.2214/AJR.18.19974

Ericsson, K. A., Charness, N., Feltovich, P. J., & Hoffman, R. R. (2006). *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge University Press.

Evans, J. S. B., & Frankish, K. E. (2009). *In two minds: Dual processes and beyond*. Oxford University Press.

Evans, M. (2007). Being driven mad: Towards understanding borderline and other disturbed states of mind through the use of the counter-transference. *Psychoanalytic Psychotherapy*, *21*(3), 216-232.

Eve, E. J., Koo, S., Alshihri, A. A., Cormier, J., Kozhenikov, M., Donoff, R. B., & Karimbux, N. Y. (2014). Performance of dental students versus prosthodontics residents on a 3D immersive haptic simulator. *Journal of Dental Education*, *78*(4), 630-637.

Ezhov, M., Gusarev, M., Golitsyna, M., Yates, J. M., Kushnerev, E., Tamimi, D., Aksoy, S., Shumilov, E., Sanders, A., & Orhan, K. (2021). Clinically applicable artificial intelligence system for dental diagnosis with CBCT. *Scientific Reports*, *11*(1), 15006. https://doi.org/10.1038/s41598-021-94093-9

Falk, E. B., Berkman, E. T., Mann, T., Harrison, B., & Lieberman, M. D. (2010). Predicting Persuasion-Induced Behavior Change from the Brain. *Journal of Neuroscience*, *30*(25), 8421-8424. https://doi.org/10.1523/JNEUROSCI.0063-10.2010

Ferréol, G., Ferréol, G., & Jucquois, G. (2003). *Dictionnaire de l'altérité et des relations interculturelles*. Armand Colin.

Fourcade, A., & Khonsari, R. H. (2019). Deep learning in medical image analysis: A third eye for doctors. *Journal of stomatology, oral and maxillofacial surgery*, *120*(4), 279-288.

Fraser, K., Ma, I., Teteris, E., Baxter, H., Wright, B., & McLaughlin, K. (2012). Emotion, cognitive load and learning outcomes during simulation training : Emotion and cognitive load during simulation. *Medical Education*, *46*(11), 1055-1062. https://doi.org/10.1111/j.1365-2923.2012.04355.x

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410-8415. https://doi.org/10.1073/pnas.1319030111

Frey, J., Hachet, M., & Lotte, F. (2017). EEG-based neuroergonomics for 3d user interfaces : Opportunities and challenges. *Le Travail Humain*, *80*(1), 73-92.

Fuchs, P., Moreau, G., Auvray M. et École nationale supérieure des mines de Paris (2006). *Le traité de la réalité virtuelle*. Paris : Presses de l'Ecole des mines.

Fukusima, S. S., Loomis, J. M., & Da Silva, J. A. (1997). Visual perception of egocentric distance as assessed by triangulation. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(1), 86-100. https://doi.org/10.1037/0096-1523.23.1.86

Gallina, J.-M. (2006). *Les représentations mentales*. Dunot.

Ganascia, J.-G. (2017). *Le mythe de la singularité : Faut-il craindre l'intelligence artificielle ?* Éditions Points.

Ganzer, N., Feldmann, I., Liv, P., & Bondemark, L. (2018). A novel method for superimposition and measurements on maxillary digital 3D models—Studies on validity and reliability. *European Journal of Orthodontics*, *40*(1), 45-51. https://doi.org/10.1093/ejo/cjx029

Gao, W., Zhang, X., Yang, L., & Liu, H. (2010). An improved Sobel edge detection. *2010 3rd International Conference on Computer Science and Information Technology*. https://doi.org/10.1109/ICCSIT.2010.5563693

Gareau, D. S., Browning, J., Correa Da Rosa, J., Suarez-Farinas, M., Lish, S., Zong, A. M., Firester, B., Vrattos, C., Renert-Yuval, Y., Gamboa, M., Vallone, M. G., Barragán-Estudillo, Z. F., Tamez-Peña, A. L., Montoya, J., Jesús-Silva, M. A., Carrera, C., Malvehy, J., Puig, S., Marghoob, A., … Krueger, J. G. (2020). Deep learning-level melanoma detection by interpretable machine learning and imaging biomarker cues. *Journal of Biomedical Optics*, *25*(11). https://doi.org/10.1117/1.JBO.25.11.112906

Gartner, L. P. (2003). Anatomical sciences in the allopathic medical school curriculum in the United States between 1967-2001. *Clinical Anatomy*, *16*(5), 434-439. https://doi.org/10.1002/ca.10159

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*(2), 191-197. https://doi.org/10.1038/72140

Gazit, E., Lieberman, M., Eini, R., Hirsch, N., Serfaty, V., Fuchs, C., & Lilos, P. (1984). Prevalence of mandibular dysfunction in 10?18 year old Israeli schoolchildren. *Journal of Oral Rehabilitation*, *11*(4), 307-317. https://doi.org/10.1111/j.1365-2842.1984.tb00581.x

Gentili, G., & Stoppato, C. (2021). Geometric function theory over quaternionic slice domains. *Journal of Mathematical Analysis and Applications*, *495*(2), 124780. https://doi.org/10.1016/j.jmaa.2020.124780

Gerig, N., Mayo, J., Baur, K., Wittmann, F., Riener, R., & Wolf, P. (2018). Missing depth cues in virtual reality limit performance and quality of three dimensional reaching movements. *PLOS ONE*, *13*(1), e0189275. https://doi.org/10.1371/journal.pone.0189275

Geuss, M. N., Stefanucci, J. K., Creem-Regehr, S. H., & Thompson, W. B. (2012). Effect of viewing plane on perceived distances in real and virtual environments. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1242-1253. https://doi.org/10.1037/a0027524

Ghinea, M., Frunză, D., Chardonnet, J.-R., Merienne, F., & Kemeny, A. (2018). Perception of Absolute Distances Within Different Visualization Systems : HMD and CAVE. In L. T. De Paolis & P. Bourdot (Éds.), *Augmented Reality, Virtual Reality, and Computer Graphics* (Vol. 10850, p. 148-161). Springer International Publishing. https://doi.org/10.1007/978-3-319-95270-3_10

Gilbert, S. B. (2016). Perceived Realism of Virtual Environments Depends on Authenticity. *Presence: Teleoperators and Virtual Environments*, *25*(4), 322-324. https://doi.org/10.1162/PRES_a_00276

Gilmour, L., & Ray, N. (2020). Locating Cephalometric X-Ray Landmarks with Foveated Pyramid Attention. *Proceedings of the Third Conference on Medical Imaging with Deep Learning*, 262-276. https://proceedings.mlr.press/v121/gilmour20a.html

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018). Explaining Explanations : An Overview of Interpretability of Machine Learning. *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, 80-89. https://doi.org/10.1109/DSAA.2018.00018

Gleicher, M., Correll, M., Nothelfer, C., & Franconeri, S. (2013). Perception of Average Value in Multiclass Scatterplots. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2316-2325. https://doi.org/10.1109/TVCG.2013.183

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *WIREs Cognitive Science*, *1*(4), 586-596. https://doi.org/10.1002/wcs.55Gobet, F. (1993). *A computer model of chess memory*. http://bura.brunel.ac.uk/handle/2438/2129

Gobet, F. (2011). *Psychologie du talent et de l'expertise*. De Boeck Supérieur. https://doi.org/10.3917/dbu.gobet.2011.01

Gobet, F., Lane, P., Croker, S., Cheng, P., Jones, G., Oliver, I., & Pine, J. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, *5*(6), 236-243. https://doi.org/10.1016/S1364-6613(00)01662-4

Gomes, A., & de Lima, E. (2006). Mandibular growth during adolescence. *The Angle Orthodontist*, *76*, 786-790. https://doi.org/10.2319/080405-263

Gomes, L. d. C. R., Horta, K. O. C., Goncalves, J. R., & Santos-Pinto, A. d. (2014). Systematic Review : Craniocervical posture and craniofacial morphology. *The European Journal of Orthodontics*, *36*(1), 55-66. https://doi.org/10.1093/ejo/cjt004

Gómez-Gómez, E., Carrasco-Valiente, J., Valero-Rosa, J., Campos-Hernández, J. P., Anglada-Curado, F. J., Carazo-Carazo, J. L., Font-Ugalde, P., & Requena-Tapia, M. J. (2015). Impact of 3D vision on mental workload and laparoscopic performance in inexperienced subjects. *Actas Urológicas Españolas (English Edition)*, *39*(4), 229-235. https://doi.org/10.1016/j.acuroe.2015.03.006

Goodale, M. A., Meenan, J. P., Bülthoff, H. H., Nicolle, D. A., Murphy, K. J., & Racicot, C. I. (1994). Separate neural pathways for the visual analysis of object shape in perception and prehension. *Current Biology*, *4*(7), 604-610. https://doi.org/10.1016/S0960-9822(00)00132-9

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature*, *349*(6305), 154-156. https://doi.org/10.1038/349154a0

Goodale, M. A., & Westwood, D. A. (2004). An evolving view of duplex vision : Separate but interacting cortical pathways for perception and action. *Current Opinion in Neurobiology*, *14*(2), 203-211. https://doi.org/10.1016/j.conb.2004.03.002

Gu, Y., Chi, J., Liu, J., Yang, L., Zhang, B., Yu, D., Zhao, Y., & Lu, X. (2021). A survey of computer-aided diagnosis of lung nodules from CT scans using deep learning. *Computers in Biology and Medicine*, *137*, 104806. https://doi.org/10.1016/j.compbiomed.2021.104806

Guang, Y., Yanfang, Y., Yongdi, Z., Songhe, W., & Yang, W. (2014). A dental model segmentation algorithm for Invisalign software. *Journal of Chemical and Pharmaceutical Research*, *6*(7), 401-408.

Guchet, X. (2014). Le patient « actionnable » de la médecine personnalisée. *Socio-anthropologie*, *29*, 37-51. https://doi.org/10.4000/socio-anthropologie.1648

Guresen, E., & Kayakutlu, G. (2011). Definition of artificial neural networks with comparison to other networks. *Procedia Computer Science*, *3*, 426-433. https://doi.org/10.1016/j.procs.2010.12.071

Gweon, H., Tenenbaum, J. B., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences*, *107*(20), 9066-9071. https://doi.org/10.1073/pnas.1003095107

Haji, F. A., Rojas, D., Childs, R., de Ribaupierre, S., & Dubrowski, A. (2015). Measuring cognitive load : Performance, mental effort and simulation task complexity. *Medical Education*, *49*(8), 815-827. https://doi.org/10.1111/medu.12773

Hamstra, S. J., Brydges, R., Hatala, R., Zendejas, B., & Cook, D. A. (2014). Reconsidering Fidelity in Simulation-Based Training. *Academic Medicine*, *89*(3), 387-392. https://doi.org/10.1097/ACM.0000000000000130

Harley, E. M., Pope, W. B., Villablanca, J. P., Mumford, J., Suh, R., Mazziotta, J. C., Enzmann, D., & Engel, S. A. (2009). Engagement of Fusiform Cortex and Disengagement of Lateral Occipital Cortex in the Acquisition of Radiological Expertise. *Cerebral Cortex*, *19*(11), 2746-2754. https://doi.org/10.1093/cercor/bhp051

Harris, L. R., Carnevale, M. J., D'Amour, S., Fraser, L. E., Harrar, V., Hoover, A. E. N., Mander, C., & Pritchett, L. M. (2015). How our body influences our perception of the world. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00819

Hays, R. T., & Singer, M. J. (1989). *Simulation Fidelity in Training System Design : Bridging the Gap Between Reality and Training*. Springer New York. https://doi.org/10.1007/978-1-4612-3564-4

He, J., Li, L., Xu, J., & Zheng, C. (2018). ReLU deep neural networks and linear finite elements. *arXiv preprint arXiv:1807.03973*.

Hegarty, M., Keehner, M., Khooshabeh, P., & Montello, D. R. (2009). How spatial abilities enhance, and are enhanced by, dental education. *Learning and Individual Differences*, *19*(1), 61-70.

Heinrich, F., Apilla, V., Lawonn, K., Hansen, C., Preim, B., & Meuschke, M. (2021). Estimating depth information of vascular models : A comparative user study between a virtual reality and a desktop application. *Computers & Graphics*, *98*, 210-217. https://doi.org/10.1016/j.cag.2021.05.014

Hennessy, J., & Al-Awadhi, E. A. (2016). Clear aligners generations and orthodontic tooth movement. *Journal of Orthodontics*, *43*(1), 68-76. https://doi.org/10.1179/1465313315Y.0000000004

Hoang, L. N. (2018). *La formule du savoir : Une philosophie unifiée du savoir fondée sur le théorème de Bayes*. EDP sciences.

Huang, C. L., Luo, Y. F., Yang, S. C., Lu, C. M., & Chen, A.-S. (2020). Influence of Students' Learning Style, Sense of Presence, and Cognitive Load on Learning Outcomes in an Immersive Virtual Reality Learning Environment. *Journal of Educational Computing Research*, *58*(3), 596-615. https://doi.org/10.1177/0735633119867422

Huckauf, A. (2005). Virtual and real visual depth. *Proceedings of the 2nd Symposium on Appied Perception in Graphics and Visualization - APGV '05*, 172. https://doi.org/10.1145/1080402.1080456

Interrante, V., Ries, B., & Anderson, L. (2006). Distance Perception in Immersive Virtual Environments, Revisited. *IEEE Virtual Reality Conference (VR 2006)*, 3-10. https://doi.org/10.1109/VR.2006.52

Irie, R., Otsuka, Y., Hagiwara, A., Kamagata, K., Kamiya, K., Suzuki, M., ... & Aoki, S. (2020). A novel deep learning approach with a 3D convolutional ladder network for differential diagnosis of idiopathic normal pressure hydrocephalus and Alzheimer's Disease. *Magnetic Resonance in Medical Sciences*, mp-2019.

Ismael, H. R., Abdulazeez, A. M., & Hasan, D. A. (2021). Detection of Diabetic Retinopathy Based on Convolutional Neural Networks : A Review. *Asian Journal of Research in Computer Science*, 1-15. https://doi.org/10.9734/ajrcos/2021/v8i330200

Jacko, J. A., & Ward, K. G. (1996). Toward establishing a link between psychomotor task complexity and human information processing. *Computers & Industrial Engineering*, *31*(1), 533-536. https://doi.org/10.1016/0360-8352(96)00192-1

Jaqueira, L. M. F., Armond, M. C., Pereira, L. J., Alcântara, C. E. P. de, & Marques, L. S. (2010). Determining skeletal maturation stage using cervical vertebrae : Evaluation of three diagnostic methods. *Brazilian Oral Research*, *24*(4), 433-437. https://doi.org/10.1590/S1806-83242010000400010

Jaquess, K. J., Gentili, R. J., Lo, L.-C., Oh, H., Zhang, J., Rietschel, J. C., Miller, M. W., Tan, Y. Y., & Hatfield, B. D. (2017). Empirical evidence for the relationship between cognitive workload and

attentional reserve. *International Journal of Psychophysiology*, *121*, 46-55. https://doi.org/10.1016/j.ijpsycho.2017.09.007

Jauregui, D. A. G., Argelaguet, F., & Lecuyer, A. (2012). Design and evaluation of 3D cursors and motion parallax for the exploration of desktop virtual environments. *2012 IEEE Symposium on 3D User Interfaces (3DUI)*, 69-76. https://doi.org/10.1109/3DUI.2012.6184186

Jaworek-Korjakowska, J., Brodzicki, A., Cassidy, B., Kendrick, C., & Yap, M. H. (2021). Interpretability of a Deep Learning Based Approach for the Classification of Skin Lesions into Main Anatomic Body Sites. *Cancers*, *13*(23), 6048. https://doi.org/10.3390/cancers13236048

Jaynes, E. T., & Bretthorst, G. L. (2003). *Probability theory : The logic of science*. Cambridge University Press.

Jeong, S. H., Yun, J. P., Yeom, H.-G., Kim, H. K., & Kim, B. C. (2021). Deep-Learning-Based Detection of Cranio-Spinal Differences between Skeletal Classification Using Cephalometric Radiography. *Diagnostics*, *11*(4), 591. https://doi.org/10.3390/diagnostics11040591

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare : Past, present and future. *Stroke and Vascular Neurology*, *2*(4), 230-243. https://doi.org/10.1136/svn-2017-000101

Johal, A. S., & Battagel, J. M. (1997). Dental crowding : A comparison of three methods of assessment. *The European Journal of Orthodontics*, *19*(5), 543-551. https://doi.org/10.1093/ejo/19.5.543

Johnson, J. E. V., & Bruce, A. C. (1998). Risk strategy under task complexity : A multivariate analysis of behaviour in a naturalistic setting. *Journal of Behavioral Decision Making*, *11*(1), 1-17. https://doi.org/10.1002/(SICI)1099-0771(199803)11:1<1::AID-BDM271>3.0.CO;2-T

Johnson, N. F. (1970). The Role of Chunking and Organization in The Process of Recall11The research reported herein was supported through the Cooperative Research Program of the Office of Education, U. S. Department of Health, Education and Welfare; Grant MH11236 from the National Institute of Mental Health, United States Public Health Service; and Grant GN 534.1 from the Office of Science Information Service, National Science Foundation, to the Computer and Information Science Research Center, The Ohio State University. In G. H. Bower (Éd.), *Psychology of Learning and Motivation* (Vol. 4, p. 171-247). Academic Press. https://doi.org/10.1016/S0079-7421(08)60432-6

Jones, N. (2014). Computer science : The learning machines. *Nature*, *505*(7482), 146-148. https://doi.org/10.1038/505146a

Jones, S. E. (2006). *Against technology : From the Luddites to Neo-Luddism*. Routledge.

José Viñas, M., Pie de Hierro, V., & M. Ustrell-Torrent, J. (2018). Superimposition of 3D digital models : A case report. *International Orthodontics*, *16*(2), 304-313. https://doi.org/10.1016/j.ortho.2018.03.017

Jost, P., Cobb, S., & Hämmerle, I. (2020). Reality-based interaction affecting mental workload in virtual reality mental arithmetic training. *Behaviour & Information Technology*, *39*(10), 1062-1078. https://doi.org/10.1080/0144929X.2019.1641228

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise : A failure to disagree. *American Psychologist*, *64*(6), 515-526. https://doi.org/10.1037/a0016755

Kalantari, S., & Neo, J. R. J. (2020). Virtual Environments for Design Research : Lessons Learned From Use of Fully Immersive Virtual Reality in Interior Design Research. *Journal of Interior Design*, *45*(3), 27-42. https://doi.org/10.1111/joid.12171

Kamel Boulos, M. N., & Toth-Cohen, S. (2009). The University of Plymouth Sexual Health SIM experience in Second Life® : Evaluation and reflections after 1 year: University of Plymouth SH-SIM experience in Second Life. *Health Information & Libraries Journal*, *26*(4), 279-288. https://doi.org/10.1111/j.1471-1842.2008.00831.x

Kapasiawala, N. N., Raval, N. R., & Patil, A. S. (2016). COMPARATIVE ANALYSIS OF DIMENSION OF FRONTAL SINUS WITH DIFFERENT SKELETAL PATTERNS. *International Journal of Current Research*, *08*, 5.

Karargyris, A., Kashyap, S., Wu, J. T., Sharma, A., Moradi, M., & Syeda-Mahmood, T. (2019). Age prediction using a large chest X-ray dataset. *arXiv:1903.06542 [cs]*. http://arxiv.org/abs/1903.06542

Katz, M. I. (1992). Angle classification revisited 1 : Is current use reliable? *American Journal of Orthodontics and Dentofacial Orthopedics*, *102*(2), 173-179. https://doi.org/10.1016/0889-5406(92)70030-E

Kelly, J. W., Hammel, W. W., Siegel, Z. D., & Sjolund, L. A. (2014). Recalibration of Perceived Distance in Virtual Environments Occurs Rapidly and Transfers Asymmetrically Across Scale. *IEEE Transactions on Visualization and Computer Graphics*, *20*(4), 588-595. https://doi.org/10.1109/TVCG.2014.36

Kenyon, R. V., Phenany, M., Sandin, D., & Defanti, T. (2008). Accommodation and Size-Constancy of Virtual Objects. *Annals of Biomedical Engineering*, *36*(2), 342-348. https://doi.org/10.1007/s10439-007-9414-7

Ker, J., Bai, Y., Lee, H. Y., Rao, J., & Wang, L. (2019). Automated brain histology classification using machine learning. *Journal of Clinical Neuroscience*, *66*, 239-245. https://doi.org/10.1016/j.jocn.2019.05.019

Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M. K., Pei, J., Ting, M. Y. L., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., … Zhang, K. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*, *172*(5), 1122-1131.e9. https://doi.org/10.1016/j.cell.2018.02.010

Kersten-Oertel, M., Chen, S., & Collins, L. (2014). An Evaluation of Depth Enhancing Perceptual Cues for Vascular Volume Visualization in Neurosurgery. *IEEE transactions on visualization and computer graphics*, *20*, 391-403. https://doi.org/10.1109/TVCG.2013.240

Khot, Z., Quinlan, K., Norman, G. R., & Wainman, B. (2013). The relative effectiveness of computer-based and traditional resources for education in anatomy : Virtual Reality Versus Reality in Anatomy. *Anatomical Sciences Education*, *6*(4), 211-215. https://doi.org/10.1002/ase.1355

Kim, I., Rajaraman, S., & Antani, S. (2019). Visual Interpretation of Convolutional Neural Network Predictions in Classifying Medical Image Modalities. *Diagnostics*, *9*(2), 38. https://doi.org/10.3390/diagnostics9020038

Kim, J., Kim, I., Kim, Y., Kim, M., Cho, J., Hong, M., Kang, K., Lim, S., Kim, S., Kim, Y. H., Kim, N., Sung, S., & Baek, S. (2021). Accuracy of automated identification of lateral cephalometric landmarks using cascade convolutional neural networks on lateral cephalograms from nationwide multi-centres. *Orthodontics & Craniofacial Research*, *24*(S2), 59-67. https://doi.org/10.1111/ocr.12493

Kjelldahl, L., & Prime, M. (1995). A study on how depth perception is affected by different presentation methods of 3D objects on a 2D display. *Computers & Graphics*, *19*(2), 199-202. https://doi.org/10.1016/0097-8493(94)00143-M

Klein, E., Swan, J. E., Schmidt, G. S., Livingston, M. A., & Staadt, O. G. (2009). Measurement Protocols for Medium-Field Distance Perception in Large-Screen Immersive Displays. *2009 IEEE Virtual Reality Conference*, 107-113. https://doi.org/10.1109/VR.2009.4811007

Knapp, J., & Loomis, J. (2003). Visual Perception of Egocentric Distance in Real and Virtual Environments. In L. Hettinger & M. Haas (Éds.), *Virtual and Adaptive Environments* (p. 21-46). CRC Press. https://doi.org/10.1201/9781410608888.pt1

Kockro, R. A., Killeen, T., Ayyad, A., Glaser, M., Stadie, A., Reisch, R., Giese, A., & Schwandt, E. (2016). Aneurysm Surgery with Preoperative Three-Dimensional Planning in a Virtual Reality

Environment : Technique and Outcome Analysis. *World Neurosurgery*, *96*, 489-499. https://doi.org/10.1016/j.wneu.2016.08.124

Kockro, R., Amaxopoulou, C., Killeen, T., Wagner, w, Reisch, R., Schwandt, E., Gutenberg, E., Giese, E., Stofft, E., & Stadie, A. (2015). Stereoscopic neuroanatomy lectures using a three-dimensional virtual reality environment. *Annals of Anatomy = Anatomischer Anzeiger : Official Organ of the Anatomische Gesellschaft*, *201*. https://doi.org/10.1016/j.aanat.2015.05.006

Korf, H.-W., Wicht, H., Snipes, R. L., Timmermans, J.-P., Paulsen, F., Rune, G., & Baumgart-Vogt, E. (2008). The dissection course – necessary and indispensable for teaching anatomy to medical students. *Annals of Anatomy - Anatomischer Anzeiger*, *190*(1), 16-22. https://doi.org/10.1016/j.aanat.2007.10.001

Koumaditis, K., Chinello, F., Mitkidis, P., & Karg, S. (2020). Effectiveness of Virtual Versus Physical Training : The Case of Assembly Tasks, Trainer's Verbal Assistance, and Task Complexity. *IEEE Computer Graphics and Applications*, *40*(5), 41-56. https://doi.org/10.1109/MCG.2020.3006330

Kravitz, N., Groth, C., Jones, P., Graham, J., & Redmond, W. (2014). Intraoral digital scanners. *Journal of clinical orthodontics : JCO*, *48*, 337-347.

Kula, K., & Ghoneima, A. (Éds.). (2018). *Cephalometry in orthodontics : 2D and 3D*. Quintessence Publishing Co,Inc.

Kulik, A., Hochstrate, J., Kunert, A., & Froehlich, B. (2009). The influence of input device characteristics on spatial perception in desktop-based 3D applications. *2009 IEEE Symposium on 3D User Interfaces*, 59-66. https://doi.org/10.1109/3DUI.2009.4811206

Kumar, A. A., Phillip, A., Kumar, S., Rawat, A., Priya, S., & Kumaran, V. (2015). Digital model as an alternative to plaster model in assessment of space analysis. *Journal of Pharmacy And Bioallied Sciences*, *7*(6), 465. https://doi.org/10.4103/0975-7406.163506

Kuntz, S., Krieghoff-Henning, E., Kather, J. N., Jutzi, T., Höhn, J., Kiehl, L., Hekler, A., Alwers, E., von Kalle, C., Fröhling, S., Utikal, J. S., Brenner, H., Hoffmeister, M., & Brinker, T. J. (2021). Gastrointestinal cancer classification and prognostication from histology using deep learning : Systematic review. *European Journal of Cancer*, *155*, 200-215. https://doi.org/10.1016/j.ejca.2021.07.012

Kunz, B. R., Wouters, L., Smith, D., Thompson, W. B., & Creem-Regehr, S. H. (2009). Revisiting the effect of quality of graphics on distance judgments in virtual environments : A comparison of verbal reports and blind walking. *Attention, Perception, & Psychophysics*, *71*(6), 1284-1293. https://doi.org/10.3758/APP.71.6.1284

Kuwada, C., Ariji, Y., Kise, Y., Funakoshi, T., Fukuda, M., Kuwada, T., Gotoh, K., & Ariji, E. (2021). Detection and classification of unilateral cleft alveolus with and without cleft palate on panoramic radiographs using a deep learning system. *Scientific Reports*, *11*(1), 16044. https://doi.org/10.1038/s41598-021-95653-9

Kwon, H. J., Koo, H. I., Park, J., & Cho, N. I. (2021). Multistage Probabilistic Approach for the Localization of Cephalometric Landmarks. *IEEE Access*, *9*, 21306-21314. https://doi.org/10.1109/ACCESS.2021.3052460

Kwon, H.-B., Park, Y.-S., & Han, J.-S. (2018). Augmented reality in dentistry : A current perspective. *Acta Odontologica Scandinavica*, *76*(7), 497-503. https://doi.org/10.1080/00016357.2018.1441437

Lagravère, M. O., & Flores-Mir, C. (2005). The treatment effects of Invisalign orthodontic aligners. *The Journal of the American Dental Association*, *136*(12), 1724-1729. https://doi.org/10.14219/jada.archive.2005.0117

Lamadé, W., Glombitza, G., Fischer, L., Chiu, P., Cárdenas, C., Thorn, M., Meinzer, H., Grenacher, L., Bauer, H., Lehnert, T., & Herfarth, C. (2000). The impact of 3-dimensional reconstructions on operation planning in liver surgery. *Archives of Surgery (Chicago, Ill. : 1960)*, *135*(11). https://doi.org/10.1001/archsurg.135.11.1256

Lanier, J. A Vintage Virtual Reality Interview, 1988. In *Interview by Adam Heilbrun of Jaron Lanier, Published*; 1988; Available online: https://philpapers.org/rec/LANAVV (accessed on 1 September 2021)

Lawonn, K., Luz, M., & Hansen, C. (2017). Improving spatial perception of vascular models using supporting anchors and illustrative visualization. *Computers & Graphics*, *63*, 37-49. https://doi.org/10.1016/j.cag.2017.02.002

Lawonn, K., Luz, M., Preim, B., & Hansen, C. (2015). Illustrative Visualization of Vascular Models for Static 2D Representations. In N. Navab, J. Hornegger, W. M. Wells, & A. Frangi (Éds.), *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015* (Vol. 9350, p. 399-406). Springer International Publishing. https://doi.org/10.1007/978-3-319-24571-3_48

Lee, H. S., Park, M., & Kim, J. (2017). Cephalometric landmark detection in dental x-ray images using convolutional neural networks. *Medical Imaging*. https://doi.org/10.1117/12.2255870

Lee, K.-S., Jung, S.-K., Ryu, J.-J., Shin, S.-W., & Choi, J. (2020). Evaluation of Transfer Learning with Deep Convolutional Neural Networks for Screening Osteoporosis in Dental Panoramic Radiographs. *Journal of Clinical Medicine*, *9*(2), 392. https://doi.org/10.3390/jcm9020392

Levinson, A. J., Weaver, B., Garside, S., McGinn, H., & Norman, G. R. (2007). Virtual reality and brain anatomy : A randomised trial of e-learning instructional designs. *Medical Education*, *41*(5), 495-501. https://doi.org/10.1111/j.1365-2929.2006.02694.x

Lewi, H. (2001). *Isaac Bashevis Singer : La génération du déluge*. Cerf.

Li, A., Montaño, Z., Chen, V. J., & Gold, J. I. (2011). Virtual reality and pain management : Current trends and future directions. *Pain Management*, *1*(2), 147-157. https://doi.org/10.2217/pmt.10.15

Li, J., & Kim, J.-E. (2021). The Effect of Task Complexity on Time Estimation in the Virtual Reality Environment : An EEG Study. *Applied Sciences*, *11*(20), 9779. https://doi.org/10.3390/app11209779

Li, L., Yu, F., Shi, D., Shi, J., Tian, Z., Yang, J., Wang, X., & Zhang, P. (2017). Application of virtual reality technology in clinical medicine. *American Journal of Translational Research*, *9*, 3867-3880.

Lieberman, D. E., Bramble, D. M., Raichlen, D. A., & Shea, J. J. (2009). Brains, Brawn, and the Evolution of Human Endurance Running Capabilities. In F. E. Grine, J. G. Fleagle, & R. E. Leakey (Éds.), *The First Humans – Origin and Early Evolution of the Genus Homo* (p. 77-92). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9980-9_8

Lieberman, D. E., Pearson, O. M., & Mowbray, K. M. (2000). Basicranial influence on overall cranial shape. *Journal of Human Evolution*, *38*(2), 291-315. https://doi.org/10.1006/jhev.1999.0335

Lieberman, D. E., Ross, C. F., & Ravosa, M. J. (2000). The primate cranial base : Ontogeny, function, and integration. *American Journal of Physical Anthropology*, *Suppl 31*, 117-169. https://doi.org/10.1002/1096-8644(2000)43:31+<117::aid-ajpa5>3.3.co;2-9

Lichtenberg, N., Hansen, C., & Lawonn, K. (2017). Concentric Circle Glyphs for Enhanced Depth-Judgment in Vascular Models. *Eurographics Workshop on Visual Computing for Biology and Medicine*, 10 pages. https://doi.org/10.2312/VCBM.20171252

Lichtenberg, N., & Lawonn, K. (2019). Auxiliary Tools for Enhanced Depth Perception in Vascular Structures. *Biomedical Visualisation*, 103-113. https://doi.org/10.1007/978-3-030-14227-8_8

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO : Common Objects in Context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Éds.), *Computer Vision – ECCV 2014* (p. 740-755). Springer International Publishing. https://doi.org/10.1007/978-3-319-10602-1_48

Lindner, C., & Cootes, T. F. (2015). Fully Automatic Cephalometric Evaluation using Random Forest Regression-Voting. *Host Publication*. IEEE International Symposium on Biomedical Imaging (ISBI) 2015 – Grand Challenges in Dental X-ray Image Analysis – Automated Detection and Analysis for Diagnosis in Cephalometric X-ray Image.

https://www.research.manchester.ac.uk/portal/en/publications/fully-automatic-cephalometric-evaluation-using-random-forest-regressionvoting(b42c658f-0a66-4d1e-99c7-9cb67fb282a0).html

Link, S., & Poulton, E. C. (1991). Bias in Quantifying Judgments. *The American Journal of Psychology*, *104*(2), 295. https://doi.org/10.2307/1423161

Liu, J., Chen, Y., Li, S., Zhao, Z., & Wu, Z. (2021). Machine learning in orthodontics : Challenges and perspectives. *Advances in Clinical and Experimental Medicine*, *30*(10), 1065-1074. https://doi.org/10.17219/acem/138702

Liu, N. (2021). Chronological age estimation of lateral cephalometric radiographs with deep learning. *arXiv:2101.11805 [cs, eess]*. http://arxiv.org/abs/2101.11805

Liu, P., & Li, Z. (2012). Task complexity : A review and conceptualization framework. *International Journal of Industrial Ergonomics*, *42*(6), 553-568. https://doi.org/10.1016/j.ergon.2012.09.001

Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., Mahendiran, T., Moraes, G., Shamdas, M., Kern, C., Ledsam, J. R., Schmid, M. K., Balaskas, K., Topol, E. J., Bachmann, L. M., Keane, P. A., & Denniston, A. K. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging : A systematic review and meta-analysis. *The Lancet Digital Health*, *1*(6), e271-e297. https://doi.org/10.1016/S2589-7500(19)30123-2

Lum, H. C., Greatbatch, R., Waldfogle, G., & Benedict, J. (2018). How Immersion, Presence, Emotion, & Workload Differ in Virtual Reality and Traditional Game Mediums. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *62*(1), 1474-1478. https://doi.org/10.1177/1541931218621334

Macredie, R. D., & Morar, S. S. (2000). Cognitive Issues in the Perception of Desktop Virtual Environments. *Journal of Intelligent Systems*, *10*(5-6). https://doi.org/10.1515/JISYS.2000.10.5-6.421

Makaremi, M. (2016). The role of cognitive sciences in orthodontic treatments : Optimizing the interface between practitioners and new technologies. *Journal of Dentofacial Anomalies and Orthodontics*, *19*(4), 410. https://doi.org/10.1051/odfen/2018116

Makaremi, M. (2018). Les enjeux d'une nouvelle interface. *Revue d'Orthopédie Dento-Faciale*, *52*(4), 319-324. https://doi.org/10.1051/odf/2018026

Makaremi, M., de Brondeau, F., Lacaule, C., & N'Kaoua, B. (2020). The digital workflow at the service of the practitioner's clinical learning. *l Orthodontie Française*, *91*(1-2), 93-99. https://doi.org/10.1684/orthodfr.2020.3

Makaremi, M., Lacaule, C., & Mohammad-Djafari, A. (2019). Deep Learning and Artificial Intelligence for the Determination of the Cervical Vertebra Maturation Degree from Lateral Radiography. *Entropy*, *21*(12), 1222. https://doi.org/10.3390/e21121222

Makaremi, M., & N'Kaoua, B. (2021). Estimation of Distances in 3D by Orthodontists Using Digital Models. *Applied Sciences*, *11*(18), 8285. https://doi.org/10.3390/app11188285

Makin, L., Barnaby, G., & Roudaut, A. (2019). Tactile and kinesthetic feedbacks improve distance perception in virtual reality. *Proceedings of the 31st Conference on l'Interaction Homme-Machine - IHM '19*, 1-9. https://doi.org/10.1145/3366550.3372248

Marcel, T. J. (2001). Three-dimensional on-screen virtual models. *American Journal of Orthodontics and Dentofacial Orthopedics*, *119*(6), 666-668. https://doi.org/10.1067/mod.2001.116502

March, J. G., & Simon, H. A. (1958). *Organizations*. John Wiley & Sons.

Marsh, W. E., Chardonnet, J.-R., & Merienne, F. (2014). Virtual Distance Estimation in a CAVE. In C. Freksa, B. Nebel, M. Hegarty, & T. Barkowsky (Éds.), *Spatial Cognition IX* (Vol. 8684, p. 354-369). Springer International Publishing. https://doi.org/10.1007/978-3-319-11215-2_25

Masud, M., Eldin Rashed, A. E., & Hossain, M. S. (2020). Convolutional neural network-based models for diagnosis of breast cancer. *Neural Computing and Applications*, 1-12.

Mastouri, R., Khlifa, N., Neji, H., & Hantous-Zannad, S. (2020). Deep learning-based CAD schemes for the detection and classification of lung nodules from CT images : A survey. *Journal of X-Ray Science and Technology*, *28*(4), 591-617. https://doi.org/10.3233/XST-200660

Mathieu, T., Bermont, L., Boyer, J.-C., Versuyft, C., Evrard, A., Cuvelier, I., Couderc, R., & Peoc'h, K. (2013). Lexical fields of predictive and personalized medicine1). *Drug Metabolism and Drug Interactions*, *28*(2). https://doi.org/10.1515/dmdi-2013-0024

Maurice, Bastien. (2018, septembre 15). Fonctionnement du neurone artificiel. *Deeply Learning*. https://deeplylearning.fr/tag/neurone/

McGloin, R., Farrar, K. M., & Krcmar, M. (2011). The Impact of Controller Naturalness on Spatial Presence, Gamer Enjoyment, and Perceived Realism in a Tennis Simulation Video Game. *Presence: Teleoperators and Virtual Environments*, *20*(4), 309-324. https://doi.org/10.1162/PRES_a_00053

Medellín-Castillo, H. I., Govea-Valladares, E. H., Pérez-Guerrero, C. N., Gil-Valladares, J., Lim, T., & Ritchie, J. M. (2016). The evaluation of a novel haptic-enabled virtual reality approach for computer-aided cephalometry. *Computer Methods and Programs in Biomedicine*, *130*, 46-53. https://doi.org/10.1016/j.cmpb.2016.03.014

Meng, J., Rieser, J. J., & Bodenheimer, B. (2006). Distance estimation in virtual environments using bisection. *Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization - APGV '06*, 146. https://doi.org/10.1145/1140491.1140523

Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments : A ten-year review of empirical research (1999–2009). *Computers & Education*, *56*(3), 769-780. https://doi.org/10.1016/j.compedu.2010.10.020

Miller, G. A. (1956). The magical number seven plus or minus two : Some limits on our capacity for processing information. *Psychological review*. https://doi.org/10.1037/h0043158

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*.

Mohammad-Rahimi, H., Nadimi, M., Rohban, M. H., Shamsoddin, E., Lee, V. Y., & Motamedian, S. R. (2021). Machine learning and orthodontics, current trends and the future opportunities : A scoping review. *American Journal of Orthodontics and Dentofacial Orthopedics*, *160*(2), 170-192.e4. https://doi.org/10.1016/j.ajodo.2021.02.013

Morar, S. S., Macredie, R. D., & Cribbin, T. (2002). An Investigation of Visual Cues used to Create and Support Frames of Reference and Visual Search Tasks in Desktop Virtual Environments. *Virtual Reality*, *6*(3), 140-150. https://doi.org/10.1007/s100550200015

Moro, C., Štromberga, Z., Raikos, A., & Stirling, A. (2017). The effectiveness of virtual and augmented reality in health sciences and medical anatomy : VR and AR in Health Sciences and Medical Anatomy. *Anatomical Sciences Education*, *10*(6), 549-559. https://doi.org/10.1002/ase.1696

Mubarrat, S. T., Opafunso, O., & Chowdhury, S. K. (2020). The Evaluation of User Experience and Functional Workload of a Physically Inter- active Virtual Reality System. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *64*(1), 2084-2086. https://doi.org/10.1177/1071181320641505

Naceri, D., Chellali, R., Dionnet, F., & Toma, S. (2010). Depth Perception Within Virtual Environments : Comparison Between two Display Technologies. *International Journal on Advances in Intelligent Systems*, *3*, 51-64.

Nahid, A.-A., & Kong, Y. (2017). Involvement of Machine Learning for Breast Cancer Image Classification : A Survey. *Computational and Mathematical Methods in Medicine*, *2017*, 1-29. https://doi.org/10.1155/2017/3781951

Naish, H., Dunbar, C., Crouch-Baker, J., Shah, K., Wallis, C., Atack, N. E., Sherriff, M., Sandy, J. R., & Ireland, A. J. (2016). Does a true knowledge of dental crowding affect orthodontic treatment decisions? *The European Journal of Orthodontics*, *38*(1), 66-70. https://doi.org/10.1093/ejo/cjv011

Napieralski, P. E., Altenhoff, B. M., Bertrand, J. W., Long, L. O., Babu, S. V., Pagano, C. C., Kern, J., & Davis, T. A. (2011). Near-field distance perception in real and virtual environments using both verbal and action responses. *ACM Transactions on Applied Perception*, *8*(3), 18:1-18:19. https://doi.org/10.1145/2010325.2010328

Nguyen, T. D. H. (2018). Imagerie tridimensionnelle de surface en orthopédie dento-faciale: revue de la littérature.

Nishimoto, S., Sotsuka, Y., Kawai, K., Ishise, H., & Kakibuchi, M. (2019). Personal Computer-Based Cephalometric Landmark Detection With Deep Learning, Using Cephalograms on the Internet: *Journal of Craniofacial Surgery*, *30*(1), 91-95. https://doi.org/10.1097/SCS.0000000000004901

Nowak, K. L., & Biocca, F. (2003). The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *12*(5), 481-494. https://doi.org/10.1162/105474603322761289

Ouramdane, N., Otmane, S., & Mallem, M. (2009). Interaction 3D en réalité virtuelle. Etat de l'art. *Techniques et sciences informatiques*, *28*(8), 1017-1049. https://doi.org/10.3166/tsi.28.1017-1049

Oviatt, S. (2006). Human-centered design meets cognitive load theory : Designing interfaces that help people think. *Proceedings of the 14th Annual ACM International Conference on Multimedia - MULTIMEDIA '06*, 871. https://doi.org/10.1145/1180639.1180831

Padalino, S., Sfondrini, M. F., Chenuil, L., Scudeller, L., & Gandini, P. (2014). Fiabilité de l'analyse de la maturité squelettique selon la méthode CVM des vertèbres cervicales faite par un logiciel dédié. *International Orthodontics*, *12*(4), 483-493. https://doi.org/10.1016/j.ortho.2014.10.004

Panwar, H., Gupta, P. K., Siddiqui, M. K., Morales-Menendez, R., Bhardwaj, P., & Singh, V. (2020). A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images. *Chaos, Solitons & Fractals*, *140*, 110190.

Parasuraman, R. (2003). Neuroergonomics : Research and practice. *Theoretical Issues in Ergonomics Science*, *4*(1-2), 5-20. https://doi.org/10.1080/14639220210199753

Patcas, R., Herzog, G., Peltomäki, T., & Markic, G. (2016). New perspectives on the relationship between mandibular and statural growth. *The European Journal of Orthodontics*, *38*(1), 13-21. https://doi.org/10.1093/ejo/cjv002

Pavani, F., Boscagli, I., Benvenuti, F., Rabuffetti, M., & Farnè, A. (1999). Are perception and action affected differently by the Titchener circles illusion? *Experimental Brain Research*, *127*(1), 95-101. https://doi.org/10.1007/s002210050777

Peck, E. M., Afergan, D., Yuksel, B. F., Lalooses, F., & Jacob, R. J. K. (2014). Using fNIRS to Measure Mental Workload in the Real World. In S. H. Fairclough & K. Gilleade (Éds.), *Advances in Physiological Computing* (p. 117-139). Springer London. https://doi.org/10.1007/978-1-4471-6392-3_6

Peillard, E., Argelaguet, F., Normand, J.-M., Lecuyer, A., & Moreau, G. (2019). Studying Exocentric Distance Perception in Optical See-Through Augmented Reality. *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 115-122. https://doi.org/10.1109/ISMAR.2019.00-13

Peillard, E., Thebaud, T., Normand, J.-M., Argelaguet, F., Moreau, G., & Lecuyer, A. (2019). Virtual Objects Look Farther on the Sides : The Anisotropy of Distance Perception in Virtual Reality. *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 227-236. https://doi.org/10.1109/VR.2019.8797826

Pensieri, C., & Pennacchini, M. (2016). Virtual Reality in Medicine. In Y. Sivan (Éd.), *Handbook on 3D3C Platforms : Applications and Tools for Three Dimensional Systems for Community, Creation and Commerce* (p. 353-401). Springer International Publishing. https://doi.org/10.1007/978-3-319-22041-3_14

Petersson, H., Sinkvist, D., Wang, C., & Smedby, Ö. (2009). Web-based interactive 3D visualization as a tool for improved anatomy learning. *Anatomical Sciences Education*, *2*(2), 61-68. https://doi.org/10.1002/ase.76

Philippe, J. (2013). Du nouveau en céphalométrie. *L'Orthodontie Française*, *84*(4), 343-346. https://doi.org/10.1051/orthodfr/2013064

Phulari, B. S. (2013). *History of orthodontics : A glance at an exciting path, the oldest specialty of dentistry has treaded so far* (1. ed).

Pirovano, A., Almeida, L. G., Ladjal, S., Bloch, I., & Berlemont, S. (2021). Computer-aided diagnosis tool for cervical cancer screening with weakly supervised localization and detection of abnormalities using adaptable and explainable classifier. *Medical Image Analysis*, *73*, 102167. https://doi.org/10.1016/j.media.2021.102167

Płotka, S., Włodarczyk, T., Szczerba, R., Rokita, P., Bartkowska, P., Komisarek, O., Matthews-Brzozowski, A., & Trzciński, T. (2021). Convolutional Neural Networks in Orthodontics : A review. *arXiv:2104.08886 [cs, eess]*. http://arxiv.org/abs/2104.08886

Poedjiastoeti, W., & Suebnukarn, S. (2018). Application of Convolutional Neural Network in the Diagnosis of Jaw Tumors. *Healthcare Informatics Research*, *24*(3), 236. https://doi.org/10.4258/hir.2018.24.3.236

Prashar, A., Vp, S., Singh, G., Sharma, N., & Singh, H. (2012). A Cephalometric Study Of Frontal Sinus And Its Relation With Craniofacial Patterns. *Undefined*. https://www.semanticscholar.org/paper/A-Cephalometric-Study-Of-Frontal-Sinus-And-Its-With-Journal-Prashar/1d6de20d2307c58f391a28755f26684ec9930c58

Preece, D., Williams, S. B., Lam, R., & Weller, R. (2013). "Let's Get Physical" : Advantages of a physical model over 3D computer models and textbooks in learning imaging anatomy: Let's Get Physical. *Anatomical Sciences Education*, *6*(4), 216-224. https://doi.org/10.1002/ase.1345

Preim, B., Baer, A., Cunningham, D., Isenberg, T., & Ropinski, T. (2016). A Survey of Perceptually Motivated 3D Visualization of Medical Image Data. *Computer Graphics Forum*, *35*(3), 501-525. https://doi.org/10.1111/cgf.12927

Priot, A.-E., Charbonneau, M., & Paillé, D. (2008). *Spatial constraints for 3D perception in helmet-mounted displays* (R. W. Brown, P. L. Marasco, T. H. Harding, & S. A. Jennings, Éds.; p. 69550G). https://doi.org/10.1117/12.790387

Proffitt, D. R. (2006). Embodied Perception and the Economy of Action. *Perspectives on Psychological Science*, *1*(2), 110-122. https://doi.org/10.1111/j.1745-6916.2006.00008.x

Proffitt, D. R., Stefanucci, J., Banton, T., & Epstein, W. (2003). The Role of Effort in Perceiving Distance. *Psychological Science*, *14*(2), 106-112. https://doi.org/10.1111/1467-9280.t01-1-01427

Puech, P.-F., Albertini, H., & Tyrand, H. (1996). CORRELATED DENTO-FACIAL PROGRESSION AND THE ORIGIN OF MAN. *Anthropologie (1962-)*, *34*(1/2), 35-38.

Qian, J., Cheng, M., Tao, Y., Lin, J., & Lin, H. (2019). CephaNet : An Improved Faster R-CNN for Cephalometric Landmark Detection. *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 868-871. https://doi.org/10.1109/ISBI.2019.8759437

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education : Design elements, lessons learned, and research agenda. *Computers & Education*, *147*, 103778. https://doi.org/10.1016/j.compedu.2019.103778

Ranzato, M. A., Poultney, C., Chopra, S., & LeCun, Y. (2007). Efficient learning of sparse representations with an energy-based model : 20th Annual Conference on Neural Information

Processing Systems, NIPS 2006. *Advances in Neural Information Processing Systems 19 - Proceedings of the 2006 Conference*, 1137-1144.

Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster R-CNN : Towards Real-Time Object Detection with Region Proposal Networks. *arXiv:1506.01497 [cs]*. http://arxiv.org/abs/1506.01497

Renner, R. S., Velichkovsky, B. M., & Helmert, J. R. (2013). The perception of egocentric distances in virtual environments—A review. *ACM Computing Surveys*, *46*(2), 1-40. https://doi.org/10.1145/2543581.2543590

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning : The effectiveness of reinforcement and non-reinforcement. *Classical Conditioning: Current Research and Theory*.

Reuschl, R. P., Heuer, W., Stiesch, M., Wenzel, D., & Dittmer, M. P. (2016). Reliability and validity of measurements on digital study models and plaster models. *The European Journal of Orthodontics*, *38*(1), 22-26. https://doi.org/10.1093/ejo/cjv001

Reyes, M., Meier, R., Pereira, S., Silva, C. A., Dahlweid, F.-M., Tengg-Kobligk, H. von, Summers, R. M., & Wiest, R. (2020). On the Interpretability of Artificial Intelligence in Radiology : Challenges and Opportunities. *Radiology: Artificial Intelligence*, *2*(3), e190043. https://doi.org/10.1148/ryai.2020190043

Rheude, B., Lionel Sadowsky, P., Ferriera, A., & Jacobson, A. (2005). An Evaluation of the Use of Digital Study Models in Orthodontic Diagnosis and Treatment Planning. *The Angle Orthodontist*, *75*(3), 300-304. https://doi.org/10.1043/0003-3219(2005)75[300:AEOTUO]2.0.CO;2

Ribbens, W., Malliet, S., Van Eck, R., & Larkin, D. (2016). Perceived realism in shooting games : Towards scale validation. *Computers in Human Behavior*, *64*, 308-318. https://doi.org/10.1016/j.chb.2016.06.055

Richler, J., Palmeri, T., & Gauthier, I. (2012). Meanings, Mechanisms, and Measures of Holistic Processing. *Frontiers in Psychology*, *3*. https://www.frontiersin.org/article/10.3389/fpsyg.2012.00553

Richman, H. B., Staszewski, J. J., & Simon, H. A. (1995). Simulation of expert memory using EPAM IV. *Psychological Review*, *102*(2), 305-330. https://doi.org/10.1037/0033-295X.102.2.305

Ries, B., Interrante, V., Kaeding, M., & Anderson, L. (2008). The effect of self-embodiment on distance perception in immersive virtual environments. *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, 167-170. https://doi.org/10.1145/1450579.1450614

Riley, J. M., & Kaber, D. B. (1999). The Effects of Visual Display Type and Navigational Aid on Performance, Presence, and Workload in Virtual Reality Training of Telerover Navigation.

Proceedings of the Human Factors and Ergonomics Society Annual Meeting, *43*(22), 1251-1255. https://doi.org/10.1177/154193129904302218

Roisin, L.-C., Brézulier, D., & Sorel, O. (2016). Remotely-controlled orthodontics : Fundamentals and description of the Dental Monitoring system. *Journal of Dentofacial Anomalies and Orthodontics*, *19*(4), 408. https://doi.org/10.1051/odfen/2016021

Ropinski, T., Steinicke, F., & Hinrichs, K. (2006). Visually Supporting Depth Perception in Angiography Imaging. *Smart Graphics*, 93-104. https://doi.org/10.1007/11795018_9

Rosenstone, R. A. (1981). *Romantic revolutionary : A biography of John Reed* (1st Vintage Books ed). Vintage Books.

Ross, A. (2004). Procrustes analysis. *Course report, Department of Computer Science and Engineering, University of South Carolina*, *26*, 1-8.

Ross, D., Tamber-Rosenau, B., Palmeri, T., Zhang, J., Xu, Y., & Gauthier, I. (2015). High Resolution fMRI Reveals Holistic Car Representations in the anterior FFA of Car Experts. *Journal of Vision*, *15*(12), 614. https://doi.org/10.1167/15.12.614

Rossini, G., Parrini, S., Castroflorio, T., Deregibus, A., & Debernardi, C. L. (2016). Diagnostic accuracy and measurement sensitivity of digital models for orthodontic purposes : A systematic review. *American Journal of Orthodontics and Dentofacial Orthopedics*, *149*(2), 161-170. https://doi.org/10.1016/j.ajodo.2015.06.029

Rousian, M., Koning, A. H. J., van der Spek, P. J., Steegers, E. A. P., & Exalto, N. (2011). Virtual reality for embryonic measurements requiring depth perception. *Fertility and Sterility*, *95*(2), 773-774. https://doi.org/10.1016/j.fertnstert.2010.12.032

Rousset, T. (2018). *Distorsion de la perception visuelle spatiale en réalité virtuelle : Mythe ou réalité ?* [These de doctorat, Aix-Marseille]. https://www.theses.fr/2018AIXM0719

Saghafi, N., Heaton, L. J., Bayirli, B., Turpin, D. L., Khosravi, R., & Bollen, A.-M. (2017). Influence of clinicians' experience and gender on extraction decision in orthodontics. *The Angle Orthodontist*, *87*(5), 641-650. https://doi.org/10.2319/020117-80.1

Sakata, S., Grove, P. M., Hill, A., Watson, M. O., & Stevenson, A. R. L. (2017). Impact of simulated three-dimensional perception on precision of depth judgements, technical performance and perceived workload in laparoscopy. *British Journal of Surgery*, *104*(8), 1097-1106. https://doi.org/10.1002/bjs.10528

Sakellariou, S., Ward, B. M., Charissis, V., Chanock, D., & Anderson, P. (2009). Design and Implementation of Augmented Reality Environment for Complex Anatomy Training : Inguinal Canal

Case Study. In R. Shumaker (Éd.), *Virtual and Mixed Reality* (Vol. 5622, p. 605-614). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-02771-0_67

Sakib, M. N., Chaspari, T., Ahn, C., & Behzadan, A. (2020, juillet 1). *An Experimental Study of Wearable Technology and Immersive Virtual Reality for Drone Operator Training*.

Satava, R. M., & Jones, S. B. (1997). Virtual Environments for Medical Training and Education. *Presence: Teleoperators and Virtual Environments*, *6*(2), 139-146. https://doi.org/10.1162/pres.1997.6.2.139

Scerbo, M. W., & Dawson, S. (2007). High Fidelity, High Performance? *Simulation in Healthcare*, *2*(4), 224-230. https://doi.org/10.1097/SIH.0b013e31815c25f1

Schell, B. H., & Martin, C. (2006). *Webster's new world hacker dictionary*. Wiley Pub.

Selfridge, O. (1959). Pandemonium: A Paradigm for Learning, Proceedings of Symposium on the Mechanization of Thought Processes. *National Physics Laboratory.*

Selvaraju, R. R., Das, A., Vedantam, R., Cogswell, M., Parikh, D., & Batra, D. (2017). Grad-CAM : Why did you say that? *arXiv:1611.07450 [cs, stat]*. http://arxiv.org/abs/1611.07450

Shanthi, T., Sabeenian, R. S., & Anand, R. (2020). Automatic diagnosis of skin diseases using convolution neural network. *Microprocessors and Microsystems*, *76*, 103074. https://doi.org/10.1016/j.micpro.2020.103074

Sheridan, H., & Reingold, E. M. (2017). Chess players' eye movements reveal rapid recognition of complex visual patterns : Evidence from a chess-related visual search task. *Journal of Vision*, *17*(3), 4. https://doi.org/10.1167/17.3.4

Shokri, A., Miresmaeili, A., Ahmadi, A., Amini, P., & Falah-kooshki, S. (2018). Comparison of pharyngeal airway volume in different skeletal facial patterns using cone beam computed tomography. *Journal of Clinical and Experimental Dentistry*, 0-0. https://doi.org/10.4317/jced.55033

Silva, C., Pinh, A., & Ferreira, Ã. (s. d.). *Frankfort Plane vs. Natural Head Posture in Cephalometric Diagnosis*. 6.

Sirazitdinov, I., Kholiavchenko, M., Mustafaev, T., Yixuan, Y., Kuleev, R., & Ibragimov, B. (2019). Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database. *Computers & Electrical Engineering*, *78*, 388-399. https://doi.org/10.1016/j.compeleceng.2019.08.004

Slater, M., & Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE) : Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *6*(6), 603-616. https://doi.org/10.1162/pres.1997.6.6.603

Slice, D. E. (2005). *Modern morphometrics in physical anthropology*. Kluwer Academic/Plenum Publishers.

Smahel, Z., & Skvaŕilová, B. (1988). Multiple correlations between craniofacial characteristics : An x-ray study. *American Journal of Physical Anthropology*, *77*(2), 221-229. https://doi.org/10.1002/ajpa.1330770210

Soler, L., Nicolau, S., Pessaux, P., Mutter, D., & Marescaux, J. (2014). Real-time 3D image reconstruction guidance in liver resection surgery. *Hepatobiliary Surgery and Nutrition*, *3*(2). https://doi.org/10.3978/j.issn.2304-3881.2014.02.03

Song, G., Jiang, R., Zhang, X., Liu, S., Yu, X., Chen, Q., Weng, X., Wu, W., Su, H., Ren, C., Shan, R., Geng, Z., Xu, T., & Research Group of Establishing Chinese Evaluation Standard of Orthodontic Treatment Outcome. (2015). [Validation of subjective and objective evaluation methods for orthodontic treatment outcome]. *Beijing Da Xue Xue Bao. Yi Xue Ban = Journal of Peking University. Health Sciences*, *47*(1), 90-97.

Sonka, M., Hlavac, V., & Boyle, R. (2015). *Image processing, analysis, and machine vision* (Fourth edition). Cengage Learning.

Stefanucci, J. K., Creem-Regehr, S. H., Thompson, W. B., Lessard, D. A., & Geuss, M. N. (2015). Evaluating the accuracy of size perception on screen-based displays: Displayed objects appear smaller than real objects. *Journal of Experimental Psychology: Applied*, *21*(3), 215.

Steiner, C. (1969). [Importance of cephalometry in orthodontic treatment]. *Informationen aus Orthodontie und Kieferorthopadie : mit Beitragen aus der internationalen Literatur*.

Steinicke, F., Bruder, G., Hinrichs, K., Steed, A., & Gerlach, A. L. (2009). Does a Gradual Transition to the Virtual World increase Presence? *2009 IEEE Virtual Reality Conference*, 203-210. https://doi.org/10.1109/VR.2009.4811024

Steinicke, F., Bruder, G., & Kuhl, S. (2011). Realistic perspective projections for virtual objects and environments. *ACM Transactions on Graphics*, *30*(5), 1-10. https://doi.org/10.1145/2019627.2019631

Steuer, J. (1992). Defining Virtual Reality : Dimensions Determining Telepresence. *Journal of Communication*, *42*(4), 73-93. https://doi.org/10.1111/j.1460-2466.1992.tb00812.x

Stevens, J., Mondesire, S., Maraj, C., Badillo-Urquiola, K., & Maxwell, D. (2016). *Workload Analysis of Virtual World Simulation for Military Training*.

Sudheer, K. P., & Jain, A. (2004). Explaining the internal behaviour of artificial neural network river flow models. *Hydrological Processes*, *18*(4), 833-844. https://doi.org/10.1002/hyp.5517

Suhermi, N., Suhartono, Rahayu, S. P., Prastyasari, F. I., Ali, B., & Fachruddin, M. I. (2019). Feature and Architecture Selection on Deep Feedforward Network for Roll Motion Time Series Prediction. In B. W. Yap, A. H. Mohamed, & M. W. Berry (Éds.), *Soft Computing in Data Science* (p. 58-71). Springer. https://doi.org/10.1007/978-981-13-3441-2_5

Swenson, H. A. (1932). The Relative Influence of Accommodation and Convergence in the Judgment of Distance. *The Journal of General Psychology*, *7*(2), 360-380. https://doi.org/10.1080/00221309.1932.9918473

Takahashi, K., Meilinger, T., Watanabe, K., & Bülthoff, H. H. (2013). Psychological influences on distance estimation in a virtual reality environment. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00580

Takeda, S., Mine, Y., Yoshimi, Y., Ito, S., Tanimoto, K., & Murayama, T. (2021). Landmark annotation and mandibular lateral deviation analysis of posteroanterior cephalograms using a convolutional neural network. *Journal of Dental Sciences*, *16*(3), 957-963. https://doi.org/10.1016/j.jds.2020.10.012

Takemiya, R., Kido, S., Hirano, Y., & Mabu, S. (2019). Detection of pulmonary nodules on chest x-ray images using R-CNN. In H. Fujita, F. Lin, & J. H. Kim (Éds.), *International Forum on Medical Imaging in Asia 2019* (p. 58). SPIE. https://doi.org/10.1117/12.2521652

Tehranchi, A., Motamedian, S. R., Saedi, S., Kabiri, S., & Shidfar, S. (2017). Correlation between frontal sinus dimensions and cephalometric indices : A cross-sectional study. *European Journal of Dentistry*, *11*(01), 064-070. https://doi.org/10.4103/1305-7456.202630

Teixeira, L. O., Pereira, R. M., Bertolini, D., Oliveira, L. S., Nanni, L., Cavalcanti, G. D. C., & Costa, Y. M. G. (2021). Impact of Lung Segmentation on the Diagnosis and Explanation of COVID-19 in Chest X-ray Images. *Sensors*, *21*(21), 7116. https://doi.org/10.3390/s21217116

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to Grow a Mind : Statistics, Structure, and Abstraction. *Science*, *331*(6022), 1279-1285. https://doi.org/10.1126/science.1192788

THILLOY (G.) 1982, L'apophyse ptérygoïde : les socles ptérygoïdiens, Université de Picardie, Thèse de doctorat d'état, 205 p.

Thomas, S. M., Lefevre, J. G., Baxter, G., & Hamilton, N. A. (2021). Interpretable deep learning systems for multi-class segmentation and classification of non-melanoma skin cancer. *Medical Image Analysis*, *68*, 101915. https://doi.org/10.1016/j.media.2020.101915

Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., & Beall, A. C. (2004). Does the Quality of the Computer Graphics Matter when Judging Distances in Visually Immersive Environments? *Presence: Teleoperators and Virtual Environments*, *13*(5), 560-571. https://doi.org/10.1162/1054746042545292

TootyVR—Application de VR Dentaire. 2020. Available online: https://www.tootyvr.com/ (accessed on 1 September 2021).

Tourancheau, S., Sjöström, M., Olsson, R., Persson, A., Ericson, T., Rudling, J., & Norén, B. (2012). *Subjective evaluation of user experience in interactive 3D visualization in a medical context* (C. K. Abbey & C. R. Mello-Thoms, Éds.; p. 831814). https://doi.org/10.1117/12.910828

Trabucco, J. T., Rottigni, A., Cavallo, M., Bailey, D., Patton, J., & Marai, G. E. (2019). User perspective and higher cognitive task-loads influence movement and performance in immersive training environments. *BMC Biomedical Engineering*, *1*(1), 21. https://doi.org/10.1186/s42490-019-0021-0

Tseng, Y.-C., Tsai, F.-C., Chou, S.-T., Hsu, C.-Y., Cheng, J.-H., & Chen, C.-M. (2021). Evaluation of pharyngeal airway volume for different dentofacial skeletal patterns using cone-beam computed tomography. *Journal of Dental Sciences*, *16*(1), 51-57. https://doi.org/10.1016/j.jds.2020.07.015

Tumer, K., & Ghosh, J. (1996). Estimating the Bayes error rate through classifier combining. *Proceedings of 13th International Conference on Pattern Recognition*. https://doi.org/10.1109/ICPR.1996.546912

Turing, A. M., & Girard, J.-Y. (1999). *La Machine de Turing*. Editions du Seuil.

Tweed, C. H. (1966). *Clinical Orthodontics Volume 2*. C.V. Mosby Company. https://oaktrust.library.tamu.edu/handle/1969.1/169330

Uysal, T., Ramoglu, S. I., Basciftci, F. A., & Sari, Z. (2006). Chronologic age and skeletal maturation of the cervical vertebrae and hand-wrist : Is there a relationship? *American Journal of Orthodontics and Dentofacial Orthopedics*, *130*(5), 622-628. https://doi.org/10.1016/j.ajodo.2005.01.031

van den Braber, W., van der Glas, H., van der Bilt, A., & Bosman, F. (2004). Masticatory function in retrognathic patients, before and after mandibular advancement surgery. *Journal of Oral and Maxillofacial Surgery: Official Journal of the American Association of Oral and Maxillofacial Surgeons*, *62*(5), 549-554. https://doi.org/10.1016/j.joms.2003.06.016

van der Hoort, B., Guterstam, A., & Ehrsson, H. H. (2011). Being Barbie : The Size of One's Own Body Determines the Perceived Size of the World. *PLoS ONE*, *6*(5), e20195. https://doi.org/10.1371/journal.pone.0020195

Varrela, J. (1992). Dimensional variation of craniofacial structures in relation to changing masticatory-functional demands. *European Journal of Orthodontics*, *14*(1), 31-36. https://doi.org/10.1093/ejo/14.1.31

Vassura, G., D'aolja, U., Vassura, M., & Venier, O. (2001). Proposition pour un nouvel algorithme diagnostique dans un système orthodontique fermé. *Le journal de l'Edgewise*, *44*, 1-12.

Vassura, G., Vassura, M., Bazzacchi, A., & Gracco, A. (2010). A shift of force vector from arm to brain : 3D computer technology in orthodontic treatment management. *International Orthodontics*, *8*(1), 46-63. https://doi.org/10.1016/j.ortho.2009.12.004

Villani, C. (2018, mars). *Donner un sens à l'intelligence artificielle. Mission parlementaire*. https://fichiers.acteurspublics.com/redac/pdf/2018/2018-03-28_Rapport-Villani.pdf

Viñas, M. J., de Hierro, V. P., & Ustrell-Torrent, J. M. (2018). Superposition de modèles numériques 3D: à propos d'un cas. *International Orthodontics*, *16*(2), 304-313.

von Cramon-Taubadel, N. (2014). The microevolution of modern human cranial variation : Implications for hominin and primate evolution. *Annals of Human Biology*, *41*(4), 323-335. https://doi.org/10.3109/03014460.2014.911350

Wainman, B., Aggarwal, A., Birk, S. K., Gill, J. S., Hass, K. S., & Fenesi, B. (2020). Virtual Dissection : An Interactive Anatomy Learning Tool. *Anatomical Sciences Education*, ase.2035. https://doi.org/10.1002/ase.2035

Wainman, B., Wolak, L., Pukas, G., Zheng, E., & Norman, G. R. (2018). The superiority of three-dimensional physical models to two-dimensional computer presentations in anatomy learning. *Medical Education*, *52*(11), 1138-1146. https://doi.org/10.1111/medu.13683

Waller, D., Knapp, D., & Hunt, E. (2001). Spatial Representations of Virtual Mazes : The Role of Visual Fidelity and Individual Differences. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *43*(1), 147-158. https://doi.org/10.1518/001872001775992561

Wallis, C., McNamara, C., Cunningham, S. J., Sherriff, M., Sandy, J. R., & Ireland, A. J. (2014). How good are we at estimating crowding and how does it affect our treatment decisions? *The European Journal of Orthodontics*, *36*(4), 465-470. https://doi.org/10.1093/ejo/cjt080

Wang, H., Wang, Z., Du, M., Yang, F., Zhang, Z., Ding, S., Mardziel, P., & Hu, X. (2020). Score-CAM : Score-Weighted Visual Explanations for Convolutional Neural Networks. *arXiv:1910.01279 [cs]*. http://arxiv.org/abs/1910.01279

Wanger, L., Ferwerda, J., & Greenberg, D. (1992). Perceiving spatial relationships in computer-generated images. *Undefined*. https://www.semanticscholar.org/paper/Perceiving-spatial-relationships-in-images-Wanger-Ferwerda/377dc9e8bff80d9c3746b1c76efead378289f5f8

Wann, J., & Mon-Williams, M. (1996). What does virtual reality NEED? : Human factors issues in the design of three-dimensional computer environments. *International Journal of Human-Computer Studies*, *44*(6), 829-847. https://doi.org/10.1006/ijhc.1996.0035

Wartenberg, C., & Wiborg, P. (2003). Precision of Exocentric Distance Judgments in Desktop and Cube Presentation. *Presence: Teleoperators and Virtual Environments*, *12*(2), 196-206. https://doi.org/10.1162/105474603321640941

Watson, D. S., Krutzinna, J., Bruce, I. N., Griffiths, C. E., McInnes, I. B., Barnes, M. R., & Floridi, L. (2019). Clinical applications of machine learning algorithms : Beyond the black box. *BMJ*, l886. https://doi.org/10.1136/bmj.l886

Weber, S., Mast, F. W., & Weibel, D. (2020). Body size illusions influence perceived size of objects : A validation of previous research in virtual reality. *Virtual Reality*, *24*(3), 385-397. https://doi.org/10.1007/s10055-019-00402-z

Weber, S., Weibel, D., & Mast, F. W. (2021). How to Get There When You Are There Already? Defining Presence in Virtual Reality and the Importance of Perceived Realism. *Frontiers in Psychology*, *12*, 628298. https://doi.org/10.3389/fpsyg.2021.628298

Willemsen, P., & Gooch, A. A. (2002). Perceived egocentric distances in real, image-based, and traditional virtual environments. *Proceedings IEEE Virtual Reality 2002*, 275-276. https://doi.org/10.1109/VR.2002.996536

Winther, F., Ravindran, L., Svendsen, K. P., & Feuchtner, T. (2020). Design and Evaluation of a VR Training Simulation for Pump Maintenance Based on a Use Case at Grundfos. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 738-746. https://doi.org/10.1109/VR46266.2020.00097

Wismer, A., Reinerman-Jones, L., Teo, G., Willis, S., McCracken, K., & Hackett, M. (2018). A Workload Comparison During Anatomical Training with a Physical or Virtual Model. In D. D. Schmorrow & C. M. Fidopiastis (Éds.), *Augmented Cognition : Users and Contexts* (Vol. 10916, p. 240-252). Springer International Publishing. https://doi.org/10.1007/978-3-319-91467-1_20

Wöber, W., Curto, M., Tibihika, P., Meulenbroek, P., Alemayehu, E., Mehnen, L., Meimberg, H., & Sykacek, P. (2021). Identifying geographically differentiated features of Ethopian Nile tilapia (Oreochromis niloticus) morphology with machine learning. *PLOS ONE*, *16*(4), e0249593. https://doi.org/10.1371/journal.pone.0249593

Wood, R. E. (1986). Task complexity : Definition of the construct. *Organizational Behavior and Human Decision Processes*, *37*(1), 60-82. https://doi.org/10.1016/0749-5978(86)90044-0

Wood, R. E., Mento, A. J., & Locke, E. A. (1987). Task complexity as a moderator of goal effects : A meta-analysis. *Journal of Applied Psychology*, *72*(3), 416-425. https://doi.org/10.1037/0021-9010.72.3.416

Wu, F., Chen, X., Lin, Y., Wang, C., Wang, X., Shen, G., Qin, J., & Heng, P.-A. (2014). A virtual training system for maxillofacial surgery using advanced haptic feedback and immersive workbench : VR-MFS using advanced haptic feedback and immersive workbench. *The International Journal of Medical Robotics and Computer Assisted Surgery*, *10*(1), 78-87. https://doi.org/10.1002/rcs.1514

Yap, B. P., & Ng, B. K. (2021). Semi-weakly Supervised Contrastive Representation Learning for Retinal Fundus Images. *arXiv:2108.02122 [cs]*. http://arxiv.org/abs/2108.02122

Yassaei, S., Emami, A., & Mirbeigi, S. (2018). Cephalometric association of mandibular size/length to the surface area and dimensions of the frontal and maxillary sinuses. *European Journal of Dentistry*, *12*(02), 253-261. https://doi.org/10.4103/ejd.ejd_345_17

Yu, H. J., Cho, S. R., Kim, M. J., Kim, W. H., Kim, J. W., & Choi, J. (2020). Automated Skeletal Classification with Lateral Cephalometry Based on Artificial Intelligence. *Journal of Dental Research*, *99*(3), 249-256. https://doi.org/10.1177/0022034520901715

Zamora, N., Llamas, Jm., Cibrian, R., Gandia, J., & Paredes, V. (2012). A study on the reproducibility of cephalometric landmarks when undertaking a three-dimensional (3D) cephalometric analysis. *Medicina Oral Patología Oral y Cirugia Bucal*, e678-e688. https://doi.org/10.4317/medoral.17721

Zeiler, M. D., & Fergus, R. (2013). Stochastic Pooling for Regularization of Deep Convolutional Neural Networks. *arXiv:1301.3557 [cs, stat]*. http://arxiv.org/abs/1301.3557

Zhang, J.-Y., Shen, Z.-H., Wang, B.-P., Liu, F., & Li, J. (2021). Influence of 3D laparoscopic surgery on surgeon's visual pattern and mental workload. *Journal of Medical Engineering & Technology*, *45*(5), 375-379. https://doi.org/10.1080/03091902.2021.1907466

Zhang, X., Zhou, F., Lin, Y., & Zhang, S. (2016). Embedding Label Structures for Fine-Grained Feature Representation. *arXiv:1512.02895 [cs]*. http://arxiv.org/abs/1512.02895

Zhong, J., & Negre, E. (2021, juin). AI : To interpret or to explain? *Congrès Inforsid ((INFormatique des ORganisations et Systèmes d'Information et de Décision) 2021*. https://hal.archives-ouvertes.fr/hal-03529203

Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2015). Learning Deep Features for Discriminative Localization. *arXiv:1512.04150 [cs]*. http://arxiv.org/abs/1512.04150

Ziemer, C. J., Plumert, J. M., Cremer, J. F., & Kearney, J. K. (2009). Estimating distance in real and virtual environments : Does order make a difference? *Attention, Perception, & Psychophysics*, *71*(5), 1095-1106. https://doi.org/10.3758/APP.71.5.1096

Annexes

Annexes-1 :

- Makaremi, M., & N'Kaoua, B. (2021). Estimation of Distances in 3D by Orthodontists Using Digital Models. *Applied Sciences*, *11*(18), 8285.

- Applied sciences impact factor : 2.679

applied sciences

MDPI

Article

Estimation of Distances in 3D by Orthodontists Using Digital Models

Masrour Makaremi [1,2,*] and Bernard N'Kaoua [2]

[1] Handicap Activity Cognition Health, Bordeaux Population Health (BPH) U1219, Université de Bordeaux,146 rue Léo Saignat, 33076 Bordeaux, France
[2] Dentofacial Orthopedics Department, UFR des Sciences Odontologiques, 146 rue Léo Saignat, 33076 Bordeaux, France; bernard.nkaoua@u-bordeaux.fr
* Correspondence: masrour@makaremi.fr

Abstract: In the field of orthodontics, digital dental arch models are increasingly replacing plaster models in orthodontic diagnostics. This change in interface from physical plaster models to digital image-based models raises the question of how orthodontists interpret intra- and inter-arch relation- ships from an image on a screen. In particular, the issue of the accuracy of the distances estimated on numerical models is crucial since the estimation of distances is the basis of the therapeutic decision- making process. Studies have shown that distances were well estimated on plaster models, but no study has verified this point on numerical models. This is the question that our study addresses. The experimental protocol consisted of collecting estimates of measurements made by orthodontists using digital models. The reliability of these measurements was then assessed by comparing them to the actual physical distances. We asked 31 orthodontists (19 women and 12 men; an average age of 37 years) to generate 3D model-based measurements of seven different elements: mandibular congestion, the maxillary intermolar distance, Spee's curve, 16/26 symmetry, the right canine class, overbite, and overjet. These values were then compared to the actual measurements calculated using Insignia® software (ORMCO Corporation: Brea, CA, USA), using single sample *t*-tests. This test makes it possible to compare a distance estimated by the participants with a reference value, which corresponds here to the real distance. The results indicate that, overall, the distance estimates made on the 3D models differ significantly from the actual distances measured using the Insignia® software. This was particularly so for mandibular crowding (test value = 0; t (30) = 10.74; $p \leq 0.01$), test value = 1; t (30) = 6.23; $p \leq 0.01$). Although no study has focused on distance estimation on nu- merical models in the field of orthodontics, our results agree with the conclusions of studies showing that distances are not estimated in the same way in real environments and virtual environments. Additional studies will make it possible to identify more clearly the parameters (individual factors, equipment, etc.), which make it possible to improve the estimation of distances in the practice of orthodontics. In any case, these studies are necessary to improve the training of future practitioners in the use of virtual models for decision-making and to support them in the digital transition.

Keywords: orthodontics; digital models; virtual reality; distance estimation

check for updates

Citation: Makaremi, M.; N'Kaoua, B. Estimation of Distances in 3D by Orthodontists Using Digital Models. *Appl. Sci.* **2021**, *11*, 8285. https://doi.org/10.3390/app11188285

Academic Editors: Gianrico Spagnuolo and Marco Tatullo

Received: 13 June 2021
Accepted: 2 September 2021
Published: 7 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affil- iations

Introduction

1.1. VR (Virtual Reality) and Medical Applications

For Ellis [1], VR is a human-machine interface that simulates a realistic environnement,allowing the participants to interact with it. Ouramdane et al [2] consider VR to be a technology that immerses one or more users in an artificial world representing a real or imaginary environnement and wich allows users to be actors capable of changing the properties of the environement and to interact with the various entities that comprise the simulated universe.

VR was initially considered to be a new gaming technology [3]. It has, however, proven to be useful in many areas such as commerce, education, the military, and architecture; among others. Today, medicine is one of the most important areas of application of virtual reality [4], for example, for diagnosis or therapeutic planning. Indeed, advances in medical imaging and advances in computing power have made simulated images much more realistic and much faster to generate, thus allowing the reconstruction of three-dimensional anatomical entities and offering practitioners the possibility to interact directly with this imagery [4].

Many fields of health are concerned, such as, for example, the treatment of a painful chronic handicap. Techniques using progressive motor imagery and VR can help manage the fear of moving or help make a movement that would normally be too painful [5]. VR is also used to treat anxiety disorders such as phobias [6], where an interactive space is created in which the patient is confronted with his situation or his phobic object. In addition, in the field of surgery, a remote control makes it possible to act on a virtual model by means of a virtual robot: the actions are first carried out in the virtual environment before being transmitted to the operator performing the operation, thereby allowing the procedure to be tested before being performed on the patient [7]. There are also various applications for the learning and training of medical personnel based on surgical simulators and softwareexploration of the human body [8].

1.2. VR in Orthopedic Dentofacial Orthodontics

The use of VR has also seen developments in the fields of orthopedic dento-facial orthodontics and orthognathic surgery. It played a major role in the initial and continuing training of practitioners. Indeed, the 3D acquisition of the different tissues of the head and neck provides a realistic platform for training in facial and dento-facial orthopedics [9]. For example, a new approach to virtual reality has been introduced and validated for cephalometric assessment by lateral view teleradiology [10]. A series of case studies using haptically activated computer-assisted cephalometry was performed. The authors showed that by providing tactile sensory feedback, errors in the cephalometric analysis were reduced, and tracking became more feasible and more intuitive. Another example is the use of the Voxel-Man simulator [11] in virtual apicoectomy, which revealed that, out of 53 medical dentistry students who performed a virtual apicoectomy, 51 had a positive evaluation of the impact of the simulation virtual as an additional modality in dental education.

Among clinical applications, techniques have been developed to obtain data on the soft and hard tissues of the maxillofacial area in order to produce virtual 3D models for analysis and surgical planning [12]. These techniques have overcome the disadvantages of 2D photographs and X-rays. Four main types of 3D imaging systems have been used to capture dental and orofacial structures, in particular conical beam tomography or CBCT, laser scanners, structured light scanners, and stereophotogrammetry [13].

Among these applications, one of the most important, in terms of impact and develop- ment, is computer-assisted 3D modeling of dental arches. Indeed, dental arch models are very useful for orthodontists [14]. They are an essential part of diagnostic information and are used to document the original condition, to design and measure the effect of treatment. They reduce the difficulty of 3D visualization of dental axes and positions during the clini- cal examination. Dental arch models have a central role in orthodontic practices around the world.

Until recently, plaster casts obtained from an imprint of the dental arches (Figure 1) were the only means to generate three-dimensional models to accurately represent the dental arches and malocclusions. 3D digital models have recently become a dematerialized alternative [15].

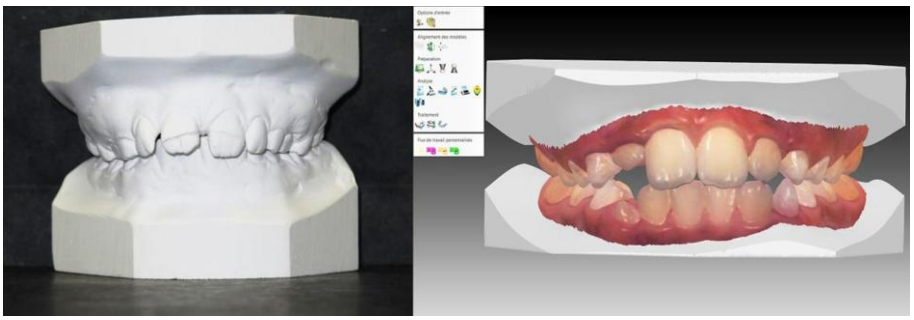


Figure 1. Plaster and 3D digital orthodontic working models

These models are electronic recordings of teeth and of the occlusion of a patient obtained with an intra-oral optical camera (Figure 1). The optical imprint of the dental arches is treated by software that draws a 3D projection, obtained from stereo-lithography (STL) files, on the computer screen. The features of the software allow the observer to change the view and to pivot or incline the image of the model. The software also allows measurement and storage of the circumferences of the arch, the diameter of the teeth, etc. Digital models have the advantage of electronic storage and remote access from multiple locations, thereby reducing the risk of loss or damage to the models. This VR technology can be considered to be part of VR as it represents two of its key characteristics:

- Manipulate the imprints virtually on the screen. There is, therefore, a real-time interaction, which allows the user to modify the properties of the environment withwhich he interacts [11,16].
- Deploy, thanks to the practitioner's diagnosis and therapeutic projection, a cognitive activity in space created digitally by 3D models (which is another key element of VR) [2].

In these 3D models, the virtual environment is not immersive because the models are viewed on a computer screen. However, there are applications that allow you to create an immersive environment using a head-mounted virtual reality screen to view models [17]. Interaction with digital work models is performed using a mouse, most often a 2D mouse, although it is also possible to use a 3D mouse. Generally, 3D digital models of dental arches are, therefore, non-immersive with a low degree of interaction.

1.3. VR and the Evaluation of Distances

Although the use of digital models has increased greatly, there are still a number of unresolved questions when it comes to comparisons with real-life models. In particular, to ensure that the applications developed in Virtual Reality are efficient, it is important that the spatial properties of the virtual environment are perceived with precision. This finding is even more important when the actions performed in VR are supposed to be transferred to the real world. If the actions carried out in VR require recalibration in the real environment, this will lead to an increase in training and cost [18].

In this framework, research indicates that distance [19–25] and size [18,26–28] are underperceived in VR, in contrast to relatively accurate perception in the real world. Waller and Richardson [29] realized a review of 14 independent studies that examined distance estimation in immersive virtual environments [23,30–41]. They showed that the estimationof distances in immersive VEs were, on average, only 71% 8.2 of the modeled distance.

As pointed out by Waller and Richardson [29], although the causes of distance com- pression in immersive VEs were not fully understood [39,42], several recent investigations documented experimental manipulations that can attenuate or alleviate it. These manipu-lations include (a) asking users to estimate distances in a known familiar environment [32]; but see [23], (b) presenting the environment on a large projection screen instead of an HMD [35], (c) providing users with explicit feedback about their distance errors [36], and(d) providing users a brief period of interaction with the environment before having them make distance judgments [34,43].

Other authors show that the task used can eliminate this effect of compressing dis- tances in VR. For example, Ref. [44] used an affordance judgment task and observed conser- vative answers or overestimation in both real and virtual environments. The methodology used by these authors involves a verbal indication that a particular action can or cannot be performed in a viewed environment. The results show that this type of VR task, which consists in deciding whether elements present in the environment are relevant to carry out a given action, gives estimates much closer to reality than the simpler and more passive tasks of size or distance estimate.

Indeed, in the study of [45], a task was carried out in two environments: a poor cue one with limited background cues and a rich cue one with textured background surfaces. The results show that the richness of the environment improves the estimation of distances and size of objects in a VR task.

All these studies show that the estimation of distances in a virtual environment is a complex problem, which depends on many parameters such as the more or less immersive nature of the environment, the familiarity with the environment, the more or less realistic character of the presented objects, the type of task to be performed (passive or actually actions to be performed), etc.

As we have indicated, in the medical field of orthodontics, distance estimation is at the center of the diagnosis and therapeutic follow-up of patients. The accuracy of these estimates is, therefore, a crucial issue from a quality-of-care perspective. Song et al. [46] published a study in which 69 orthodontists had to estimate, from 108 plaster molds, different variables, which included all the variables usually estimated by an orthodontist. The values were compared to the actual values measured with calipers. The results obtained revealed a statistically significant correlation between the objective measurements and the estimations carried out by the orthodontists. However, very few studies were carried outon virtual models, which are now widely used by professionals.

1.4. Objective of the Study

As we have seen, in the field of orthodontics, the digital transition means that pro- fessionals make therapeutic decisions by making distance estimations on digital models. However, no study has focused on the relevance of the estimates made, and in particularon the difference between estimated distances and actual distances.

In this context, the objective of our study was to use the methodology of Song et al. [46] (which these authors implemented on plaster models) to assess the relevance of estimating distances (intra and inter-arcs) using digital models. The experimental protocol consisted of collecting estimates of measurements made by orthodontists on digital models, then assessing the reliability of these measurements by comparing them to real distances.

Method

1.5. Participants

The measurements were estimated by 31 orthodontists whose characteristics were as follows: all the practitioners began their specialization in dento-facial orthopedics and orthodontics in French universities. The distribution of the sample by sex was 19 women and 12 men, the average age of the participants was 37 years (3.7). The number of years of exercise of the profession was distributed as follows: 0 to 5 years of professional practice (42%), 5 to 10 years (32%), and more than 10 years (26%). The objective of this research was explained to the practitioners, who gave their consent. The participants were recruited by an email sent to all orthodontic practices in the region. The participation was voluntary, and the experimentation was carried out in the participant's office.

1.6. Materials

The estimations of distances were carried out on 5 3D digital models (obtained with Insignia® software), derived from 5 clinical cases (Figure 2). The choice of these 5 cases (also used in the study by Song et al.) was based on the objective of allowing the practitioners to make estimations on a varied panel of orthodontic malocclusions comprising all of the types of dental-maxillary dysmorphias in the three dimensions of space. The characteristics of the different malocclusions are reported Table 1.

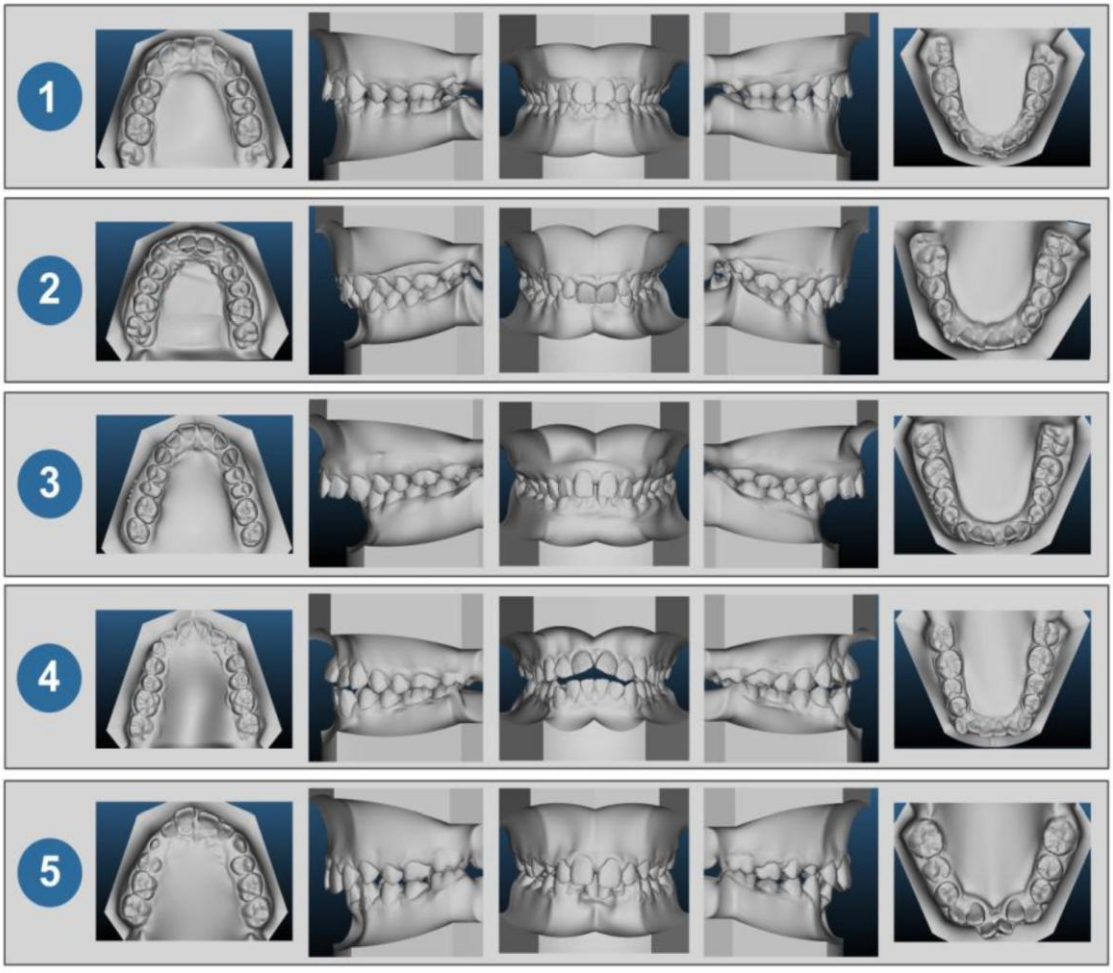


Figure 2. The five clinical cases used for distance estimation.

Table 1. Presentation of the five clinical cases used in our study.

Case no. 1	Class I with moderate crowding of the arch

Case no. 2	Class II with pronounced incisor supraclusion
Case no. 3	Class II with a significantly increased overjet
Case no. 4	Class II incisor infraclusion
Case no. 5	Class II with malposition of the mandibular incisors

The digital models were displayed on the same computer, by the same operator, and under the same conditions for each of the 31 practitioners. The participants had the option to change the orientation of the view of the model. The collection of distance estimates was carried out as follows: after observing the virtual models of each clinical case on the screen, the practitioner verbally gives their distance estimates for each of the 7 variables to the operator, who writes them down in a table.1

The variables to be estimated:
For each of the models, each practitioner had to estimate 7 quantifiable variables. These estimates were carried out on the 5 virtual models mentioned (5 × 7 = 35 estimates by orthodontist) without the possibility of making direct measurements (they were onlyestimates). The variables selected were similar to those of the study by Song et al. [46]: These variables were the following (Figure 3):

- Crowding of the mandibular arch: the amount of space required for proper alignment of the teeth in the arch (volumetric reconstitution, intra-arch measurement);
- Spee's curve—the curve formed by the projection in a sagittal plane of the buccal cusps of the mandibular teeth (vertical dimension, intra-arch measurement);
- Antero-posterior symmetry of the first permanent maxillary molars (teeth 16 and 26) while taking their mesial side as a reference (sagittal dimension, intra-arch measurement); Inter-molar distance of 16/26 (horizontal dimension, intra-arch measurement): dis-tance separating the mesiobuccal cusp of the first permanent maxillary molars (teeth 16 and 26);
- Right canine class according to Angle's classification [47] based on the mesiodistal relationship of the teeth (sagittal dimension, inter-arch measurement);
- Overjet of 11/41—the gap as assessed on to the occlusal plane between the buccal side of the mandibular incisors and the occlusal edges of the central maxillary incisors (sagittal dimension, inter-arch measurement);
- Alignment of 11/41—Alignment of the mandibular incisors by their opposing maxil- laries (vertical dimension, inter-arch measurement).

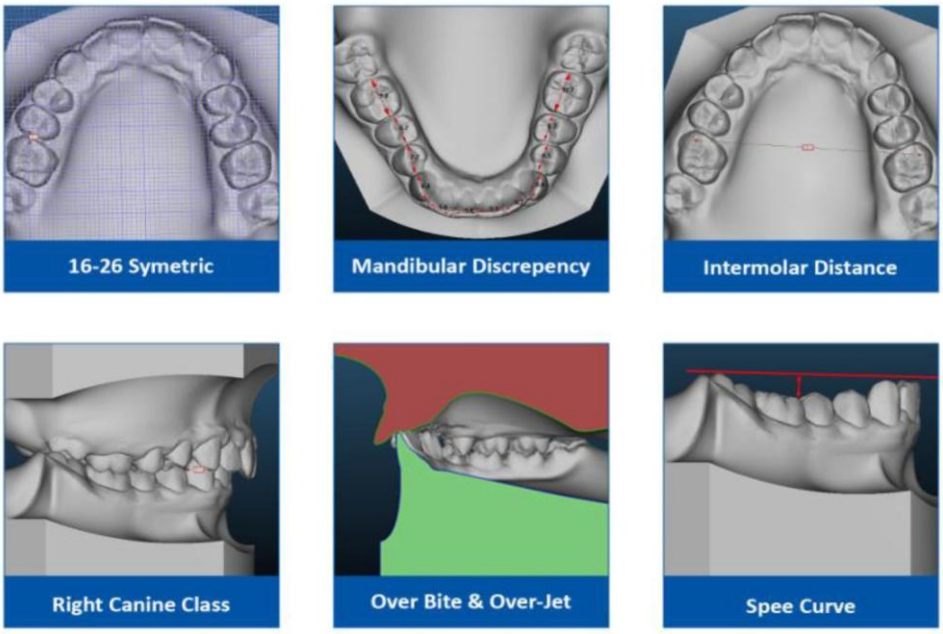


Figure 3. Presentation of the seven estimated variables.

The question that we wanted to address was the performance of the orthodontist's estimate compared to the measurements made with software whose reliability in this regard has been demonstrated and which can, therefore, serve as a reference. Indeed, the reliability and precision of the measurements made using different software have already been demonstrated [47–49]. The actual measurements were calculated for each case by the same operator using the Insignia® software (ORMCO Corporation: Brea, CA, USA) and then verified by a second operator.
For each of the 7 variables, the difference in the absolute value between the estimation of the measurement by the practitioner and the measurement of the variable carried using the Insignia® software was determined.

A flowchart of the procedure is shown in Figure 4.

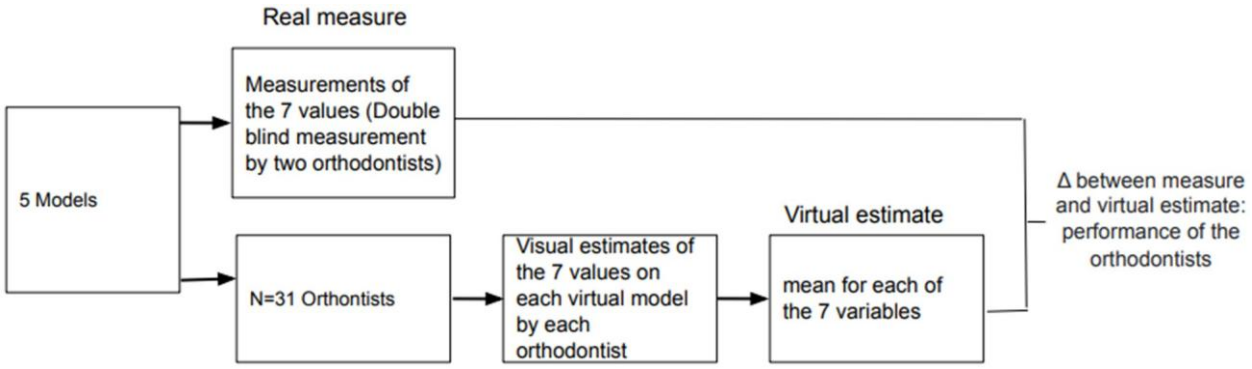


Figure 4. Flowchart of the procedure.

1.7. Statistical Analyses

As the objective was to compare the estimates made by orthodontists with the mea- surements made with the Insignia® software, we performed a single sample *t*-test. This test is generally used to compare the average of a sample for a specific variable (here, the average of the orthodontist's estimates) with a population average or with a theoretical reference average (here, the measurement carried out by the Insignia® software).

For each estimated variable (i.e., the 7 variables presented above), the difference between the estimate made by the practitioners and the actual distance was tested. The actual distance was reduced to 0 (since the estimated measurements have been subtracted from the actual measurements). For each variable measured, the difference between the estimated measurements and the real distance (here reduced to 0) was assessed using a single sample t-test. First, the theoretical value of 0 was used. If the t-test was not significant, it meant that the estimated value was not significantly different from the actual value. If the t-test was significant, it meant that the estimated value was significantly different from the actual value. Secondly, we performed the same analysis using a reference value equal to 1 (i.e., 1 cm more than the actual value equal to 0). Again, if the t-test was significant, it meant that the estimation error made by the practitioner was greater than1 cm.

Results

The results obtained for each model are reported in Table 2.On the basis of the single *t*-test results, the 7 variables fell into 3 groups: (1) variables for which the estimates were not significantly different from the real values; in this case, the estimated value was not significantly different from the actual value; (2) variables for which the estimates were significantly different from the actual value 0 but not from the value 1; in this case, the estimation error was greater than 0 but less than 1 cm; (3) variables for which the estimation error was significantly different from the value 1; in this case, the estimation error was greater than 1 cm (Table 3). Table 3 (power column) also shows the power of each result for a β risk of 20%. The column values indicate the minimum sample size for the conclusion at H0 to be strong enough. The column values indicate the minimum sample size for the conclusion at H0 to be strong enough. This power analysis show that the size samples were sufficient for all the significant results, with the exception of the right canine class variable for a theoretical value of 0 (but for this same variable, the size sample is sufficient for a value theoretical of 1).

Table 2. Means and standard deviation for each model. ×1 represent the machine values, ×2 is the mean of the measurement. σ is the standard deviation.

	Case No. 1	Case No. 2	Case No. 3	Case No. 4	Case No. 5
Mandibular Discrepancy	×1: −3 ×2: −2.73 σ: 1.43	×1: 0.5 ×2: −2.35 σ: 3.21	×1: −3 ×2: −3.55 ×1: −3	×1: −3.5 ×2: −4.44 ×1: −3.5	×1: 2.5 ×2: −4.60 ×1: 2.5
Intermolar Distance	× 1: 36 × 2: 38.10	× 1: 37 × 2: 38.6 σ: 7.82	× 1: 36.5 × 2: 38 σ: 8.22	× 1: 40 × 2: 36.6 σ: 8.73	× 1: 40 × 2: 41 σ: 8.σ: 8.36
Spee Curve	× 1: 2 × 2: 2.42 σ: 0.81	× 1: 1.5 × 2: 1.66 σ: 0.90	× 1: 2.5 × 2: 2.47 σ: 0.86	× 1: 2 × 2: 1.03 σ: 1.07	× 1: 2.5 × 2: 2.2 σ: 0.87
16/26 Symetric	×1: 3 ×2: 1.52 σ: 1.80	×1: 1 ×2: 0.32 σ: 0.97	×1: 1 ×2: 0.11 σ: 1.19	×1: 0 ×2: −0.32 σ: 0.77	×1: 1.5 ×2: 0.19 σ: 1.60
Right Canine Class	×1: 6 ×2: 5.45 σ: 1.47	×1: 1 ×2: 3.24 σ: 2.57	×1: 8 ×2: 7.34 σ: 2.19	×1: −4 ×2: −2.75 σ: 3.08	×1: 2 ×2: 2.95 σ: 1.36
Over Jet	×1: 1.5 ×2: 1.45 σ: 1.23	×1: 3.5 ×2: 4.32 σ: 1.81	×1: 14 ×2: 10.94 σ: 4.47	×1: 3 ×2: 2.97 σ: 1.04	×1: 4.5 ×2: 5.08 σ: 1.96
Over Bite	×1: 7 ×2: 6.27 σ: 1.60	×1: 5 ×2: 3.52 σ: 1.80	×1: 5 ×2: 9.55 σ: 1.11	×1: −5 ×2: 4.18 σ: 2.05	×1: 4 ×2: 3.63 σ: 0.97

Table 3. Results of the one-sample test for all of the variables. This table shows the mean value, standard deviation, *t*-test value, and the significance for each variable. The upper table presents results with 0 cm as a reference value. The lower table presents results with 1 cm as a reference value (only significant variables with the value 0 are tested with the value 1).

t = 0	**Mean**	**Standard Deviation**	**t**	**Sig.**	**Power (0.8)**
Mandibular Discrepancy	−2.23	1.15	−10.74	≤0.01	5
Right Canine Class	0.62	1.08	3.19	≤0.01	48
Spee Curve	−0.13	0.58	−1.33	0.19	313
Intermolar Distance	0.56	7.79	0.39	0.69	3049
Over Bite	−0.65	0.66	−5.40	≤0.01	17
Over Jet	−0.34	1.20	−1.60	0.11	197
16/26 Symetric	−0.93	0.53	−9.65	≤0.01	6
t = 1	**Mean**	**Standard Deviation**	**t**	**Sig.**	**Power (0.8)**
Mandibular Discrepancy	2.25	1.11	6.23	≤0.01	4
Right Canine Class	0.92	0.84	−0.51	0.61	14
Over Bite	0.80	0.44	0.64	0.52	5
16/26 Symetric	0.93	0.53	−0.66	0.51	6

-Variables for which the estimates were not significantly different from the real values. This was the case for the following 3 variables: The overjet (t (30) = 1.60; $p = 0.11$); The curve of Spee (t (30) = −1.33; $p = 0.19$); The inter-molar distance (t (29) = 0.39;$p = 0.69$).

-Variables for which the estimates were significantly different from the actual value 0 but not from the value 1. This was the case for the following 3 variables: The 16/26 symmetry: value = 0 (t (30) = −9.65; $p \leq 0.01$), but with a test value = 1 (t (30) = −0.66;

p = 0.51); The over bite: test value = 0 (t (30) = 5.40; p 0.01), but with a test value = 1 (t (30) = 0.64; p = 0.52); The right canine class: test value = 0 (t (30) = 3.19; p = 0.01), but with a test value = 1 (t (30) = 0.51; p = 0.61).

-Variables for which the estimation error was significantly different from the value

This was the case only for the mandibular discrepancy variable: test value = 0(t (30) = −10.74; $p \leq$ 0.01), test value = 1 (t (30) = 6.23; $p \leq 0.01$).

Discussion

The objective of our study was to find out if, for seven variables of interest, the estimation of distances carried out on numerical models was significantly different (or not) from the real distance (here, reduced to the value 0). For the variables whose estimates were significantly different from the real distance (comparison between the measurements carried out by the practitioners and the theoretical mean of 0), we also checked whether this difference between the estimates and the real value was greater than 1 cm (comparison between the measurements made by practitioners and the theoretical average of 1).

Of the seven variables studied, three were associated with distances estimated to be significantly greater than the real distance, and for one variable, the difference was greater than 1 cm.

These results question the reliability of the use of numerical models for decision- making, in particular in borderline cases. The relevant elements for this discussion are the difference in the estimation of distances between virtual and real models, the way in which our results can be discussed in the light of the work on the estimation of distances in virtual environments, and the perspectives of this study as part of further investigations.

The results obtained in the study by Song et al. [46] on real models revealed a statisti- cally significant correlation between the objective measurements and the estimates made by orthodontists for all the variables (the same variables as those used in our study).

Thus, with only the virtual model (i.e., other than a situation where the practitioner can manipulate real physical models), the estimate was less reliable for four of the seven variables, in particular the estimate of the mandibular crowding.

The results of our study are in keeping with the various studies in the literature regarding the estimation of distances in a digital environment in terms of the presence of bias in the estimation of distances in a virtual environment [25,50–53].

In terms of crowding of the mandibular arch, which is the amount of space required for proper alignment of the teeth on the arch, this is the variable estimated in the least precise way. One of the explanations could be the difficulty for the practitioner to perform a three- dimensional mental reconstruction from the virtual model to estimate the space required for aligning the teeth from the malposed teeth. It is of particular interest to note that it was in case no. 5 that the accuracy was less good, that is to say, where the estimation of the crowding requires reconstitution of the shape of the arch based on the initial malpositions that were substantial before estimation of the lack of space for alignment of the teeth on the arch. For this particular case, the performance of the practitioners was particularly poor, with a relatively large standard deviation (σ = 7.47), while the value of the variable was low (Table 2). The large standard deviation between the estimated value and the actual value in case no. 5 can be explained by the study by [54], who showed the difficulty of maintaining attention and, therefore, the efficacy of the process beyond a certain degree of complexity of a 3D image. This appears to be illustrated by the difficulty of the estimationof the crowding.

Future work should also study the role of calibration. This is a very important point.The contribution of calibration to visual perception in VR has been established by numerousworks [43,55,56]. For example, Ref. [56] showed the effects of calibration to visual and proprioceptive information on near filed distance judgments in 3D user interface. One method of calibration that could be experienced would allow viewing the error that they made on the different estimation of distance, and they allow the participant to correct the response. If this work were conclusive, the calibration training programs could havea positive impact on the performance of practitioners. Studies on these different points will allow the development of educational materials to employ a wider range of stimuli to support the training of practitioners in the use of virtual models for decision-making and also to optimize the interaction of the practitioner with the virtual environment. Forexample, one image with a particular mandibular discrepancy can be transformed to create a number of stimuli with a continuum of different mandibular discrepancies.

Another aspect of estimating values in a virtual environment that needs to be explored in our study is the

mental workload necessary to estimate distances in a virtual environ- ment [57] and, therefore, in this context, on 3D digital orthodontic models. According to the experience of users and practitioners, the mental workload required for the analysis of virtual models seems to be greater than that required for physical models. This increase in mental load accentuated the difficulty of transition between real models and virtual models. This feeling is widely shared (interview with the participants in our study after the experiment) and needs to be documented.

In this context, the main limitations of our study are, therefore, that, as we have just seen, we did not measure the mental load associated with the estimates to be made, but also, we did not take into account the number of years of professional experience (evaluation of the effect of expertise on the estimation of distances), and we did not compare the estimates made on the numerical models with the estimates made on the same plaster models (our comparison focused on the real distance and not on the estimate made on plaster models). The future investigations, which we are in the process of carrying out, are intended to answer all of these questions.

The future investigations, which we are in the process of carrying out, are intended to answer these main questions.

In addition, the difficulties related to digital models also appear to be due to the fact that practitioners handle the actual models to estimate distances. As a result, the transition of real models to virtual models involves a paradigm shift, with a loss of haptic range for the practitioner who can no longer manipulate the models with his hands. This observation may be linked to the different studies showing dissociation of neuronal routes related to the seizure or perception of objects [58–61]. Another interesting perspective would then be to study the estimation of distances on digital models by proposing to practitioners a more realistic manipulation device than that used in our study (computer mouse), such as a glove or an articulated arm.

Conclusions

3D digital models of dental arches can be considered a non-immersive virtual reality characterized by a low degree of interaction. These models are now widely used to carry out diagnoses, but the estimation of distances by the user (orthodontist) carried out on this type of digital model, and which constitutes a determining factor in the therapeutic decision-making, had not yet been studied.

It is within this framework that we conducted our study. Our results of distance estimations on numerical models showed that four of the seven variables tested were associated with a lack of precision in the estimation of distances, the size of the mandibular arch being the least well-estimated variable.

Although no studies have been carried out in the field of orthodontia, the lack of precision in the estimation of measurements made on virtual 3D models corresponds to what has been described in studies relating to distance estimations in digital environments. Other studies in preparation will be necessary to understand better the variables (related to professionals, equipment, type of models, etc.) likely to affect the estimation of distances and, therefore, the therapeutic decision-making. Such studies are fundamental in particular for improving the training of professionals in the use of new digital tools. These objectives can be best addressed through a multidisciplinary approach [62], in order to help keep the expertise and perception of the practitioner at the center of decision-making, thus ensuring a successful digital transition.

Author Contributions: Conceptualization, M.M. and B.N.; methodology, M.M.; software, M.M.; validation, M.M. and B.N.; formal analysis, B.N.; investigation, M.M.; resources, M.M.; data curation, M.M.; writing—original draft preparation, M.M. and B.N.; writing—review and editing, B.N.; visual- ization, M.M.; supervision, B.N.; project administration, M.M.; funding acquisition, M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Ethics Committee of University of Bordeaux.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The study data can be obtained by email request to the authors.

Acknowledgments: Charpentier Valentine; Marcy Benoit; Gibaudon Lionel.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Stephen, R.E. What are virtual environments? *IEEE Comput. Graph. Appl.* **1994**, *1*, 17–22.
2. Ouramdane, N.; Otmane, S.; Mallem, M. Interaction 3D en Réalité Virtuelle-Etat de l'art. *Tech. Sci. Inform.* **2009**, *28*, 1017–1049.[CrossRef]
3. Engler, C.E. Affordable VR by 1994. *Comput. Gaming World* **1992**, *1*, 80–82.
4. Li, L.; Yu, F.; Shi, D.; Shi, J.; Tian, Z.; Yang, J.; Wang, X.; Jiang, Q. Application of virtual reality technology in clinical medicine. *Am. J. Transl. Res.* **2017**, *9*, 3867–3880. [PubMed]
5. Baldominos, A.; Saez, Y.; del Pozo, C.G. An approach to physical rehabilitation using state-of-the-art virtual reality and motiontracking technologies. *Procedia Comput. Sci.* **2015**, *64*, 10–16. [CrossRef]
6. Li, A.; Montaño, Z.; Chen, V.J.; Gold, J.I. Virtual reality and pain management: Current trends and future directions. *Pain Manag.* **2011**, *1*, 147–157. [CrossRef] [PubMed]
7. Satava, R.M.; Jones, S.B. Virtual environments for medical training and education. *Presence Teleoperators Virtual Environ.* **1997**, *6*,139–146. [CrossRef]
8. Ackerman, M.J. The visible human project. *Proc. IEEE* **1998**, *86*, 504–511. [CrossRef]
9. Ayoub, A.; Pulijala, Y. The application of virtual reality and augmented reality in Oral & Maxillofacial Surgery. *BMC Oral. Health* **2019**, *19*, 1–8.
10. Medellín-Castillo, H.I.; Govea-Valladares, E.H.; Pérez-Guerrero, C.N.; Gil-Valladares, J.; Lim, T.; Ritchie, J.M. The evaluation of a novel haptic-enabled virtual reality approach for computer-aided cephalometry. *Comput. Methods Programs Biomed.* **2016**, *130*, 46–53. [CrossRef]
11. Wu, F.; Chen, X.; Lin, Y.; Wang, C.; Wang, X.; Shen, G.; Qin, J.; Heng, P.-A. A virtual training system for maxillofacial surgery using advanced haptic feedback and immersive workbench. *Int. J. Med Robot. Comput. Assist. Surg.* **2014**, *10*, 78–87. [CrossRef]
12. Kwon, H.B.; Park, Y.S.; Han, J.S. Augmented reality in dentistry: A current perspective. *Acta Odontol. Scand.* **2018**, *76*, 497–503. [CrossRef] [PubMed]
13. Ayoub, A.F.; Xiao, Y.; Khambay, B.; Siebert, J.P.; Hadley, D. Towards building a photo-realistic virtual human face for craniomax- illofacial diagnosis and treatment planning. *Int. J. Oral Maxillofac. Surg.* **2007**, *36*, 423–428. [CrossRef] [PubMed]
14. Marcel, T.J. Three-dimensional on-screen virtual models. *Am. J. Orthod. Dentofac. Orthop.* **2001**, *119*, 666–668. [CrossRef] [PubMed]
15. Kravitz, N.D.; Groth, C.; Jones, P.E.; Graham, J.W.; Redmond, W.R. Intraoral digital scanners. *J. Clin. Orthod.* **2014**, *48*, 337–347.
16. Lanier, J. A Vintage Virtual Reality Interview, 1988. In *Interview by Adam Heilbrun of Jaron Lanier, Published*; 1988; Available online:https://philpapers.org/rec/LANAVV (accessed on 1 September 2021).
17. TootyVR—Application de VR Dentaire. 2020. Available online: https://www.tootyvr.com/ (accessed on 1 September 2021).
18. Siegel, Z.D.; Kelly, J.W. Walking through a virtual environment improves perceived size within and beyond the walked space. *Atten. Percept. Psychophys.* **2017**, *79*, 39–44. [CrossRef]
19. El Jamiy, F.; Marsh, R. Distance estimation in virtual reality and augmented reality: A survey. In Proceedings of the 2019 IEEE International Conference on Electro Information Technology (EIT), Brookings, SD, USA, 20–22 May 2019; IEEE: New York, NY,USA, 2019; pp. 063–068.
20. Swan, J.E.; Kuparinen, L.; Rapson, S.; Sandor, C. Visually perceived distance judgments: Tablet-based augmented reality versusthe real world. *Int. J. Hum.-Comput. Interact.* **2017**, *33*, 576–591. [CrossRef]
21. Bodenheimer, B.; Meng, J.; Wu, H.; Narasimham, G.; Rump, B.; McNamara, T.P.; Carr, T.H.; Rieser, J.J. Distance estimation in virtual and real environments using bisection. In Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization, Tubingen, Germany, 25–27 July 2007; pp. 35–40.
22. Kuhl, S.A.; Thompson, W.B.; Creem-Regehr, S.H. HMD calibration and its effects on distance judgments. *ACM Trans. Appl. Percept.* **2009**, *6*, 1–20.
23. Messing, R.; Durgin, F.H. Distance perception and the visual horizon in head-mounted displays. *ACM Trans. Appl. Percept.* **2005**, *2*, 234–250. [CrossRef]
24. Steinicke, F.; Bruder, G.; Hinrichs, K.; Lappe, M.; Ries, B.; Interrante, V. Transitional environments enhance distance perception in immersive virtual reality systems. In Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization,Chania, Greece, 30 September–2 October 2009; pp. 19–26.
25. Ziemer, C.J.; Plumert, J.M.; Cremer, J.F.; Kearney, J.K. Estimating distance in real and virtual environments: Does order make a difference? *Atten. Percept. Psychophys.* **2009**, *71*, 1095–1106. [CrossRef] [PubMed]
26. Kelly, J.W.; Donaldson, L.S.; Sjolund, L.A.; Freiberg, J.B. More than just perception-action recalibration:

Walking through a virtual environment causes rescaling of perceived space. *Atten. Percept. Psychophys.* **2013**, *75*, 1473–1485. [CrossRef] [PubMed]
27. Kunz, B.R.; Creem-Regehr, S.H.; Thompson, W.B. Does perceptual-motor calibration generalize across two different forms of locomotion? Investigations of walking and wheelchairs. *PLoS ONE* **2013**, *8*, e54446. [CrossRef]
28. Stefanucci, J.K.; Creem-Regehr, S.H.; Thompson, W.B.; Lessard, D.A.; Geuss, M.N. Evaluating the accuracy of size perception on screen-based displays: Displayed objects appear smaller than real objects. *J. Exp. Psychol. Appl.* **2015**, *21*, 215. [CrossRef] [PubMed]
29. Waller, D.; Richardson, A.R. Correcting distance estimates by interacting with immersive virtual environments: Effects of task and available sensory information. *J. Exp. Psychol. Appl.* **2008**, *14*, 61. [CrossRef] [PubMed]
30. Durgin, F.H.; Gigone, K.; Scott, R. Perception of visual speed while moving. *J. Exp. Psychol. Hum. Percept. Perform.* **2005**, *31*, 339.
[CrossRef] [PubMed]
31. Henry, D.; Furness, T. Spatial perception in virtual environments: Evaluating an architectural application. In Proceedings of the IEEE Virtual Reality Annual International Symposium, Seattle, WA, USA, 18–22 September 1993; pp. 33–40.
32. Interrante, V.; Ries, B.; Anderson, L. Distance perception in immersive virtual environments, revisited. In Proceedings of the IEEE Virtual Reality Conference (VR 2006), Alexandria, VA, USA, 25–29 March 2006; pp. 3–10.
33. Knapp, J.M.; Loomis, J.M. Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence Teleoperators Virtual Environ.* **2004**, *13*, 572–577. [CrossRef]
34. Mohler, B.J.; Creem-Regehr, S.H.; Thompson, W.B. The influence of feedback on egocentric distance judgments in real and virtual environments. In Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization, Boston, MA, USA, 28–29 July 2006; pp. 9–14.
35. Plumert, J.M.; Kearney, J.K.; Cremer, J.F.; Recker, K. Distance perception in real and virtual environments. *ACM Trans. Appl. Percept.*
2005, *2*, 216–233. [CrossRef]
36. Richardson, A.R.; Waller, D. The effect of feedback training on distance estimation in virtual environments. *Appl. Cogn. Psychol.*
2005, *19*, 1089–1108. [CrossRef]
37. Sahm, C.S.; Creem-Regehr, S.H.; Thompson, W.B.; Willemsen, P. Throwing versus walking as indicators of distance perception in similar real and virtual environments. *ACM Trans. Appl. Percept.* **2005**, *2*, 35–45. [CrossRef]
38. Sinai, M.J.; Krebs, W.K.; Darken, R.P.; Rowland, J.H.; McCarley, J.S. Egocentric distance perception in a virutal environment using a perceptual matching task. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Los Angeles, CA, USA, 1 September 1999; SAGE Publications: Los Angeles, CA, USA, 1999; Volume 43, pp. 1256–1260.
39. Thompson, W.B.; Willemsen, P.; Gooch, A.A.; Creem-Regehr, S.H.; Loomis, J.M.; Beall, A.C. Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence* **2004**, *13*, 560–571. [CrossRef]
40. Willemsen, P.; Gooch, A.A. Perceived egocentric distances in real, image-based, and traditional virtual environments. In Proceedings of the IEEE Virtual Reality 2002, Orlando, FL, USA, 24–28 March 2002; pp. 275–276.
41. Witmer, B.G.; Sadowski, W.J., Jr. Nonvisually guided locomotion to a previously viewed target in real and virtual environments.
Hum. Factors **1998**, *40*, 478–488. [CrossRef]
42. Loomis, J.M.; Knapp, J.M. Visual perception of egocentric distance in real and virtual environments. *Virtual Adapt. Environ.* **2003**,
11, 21–46.
43. Richardson, A.R.; Waller, D. Interaction with an immersive virtual environment corrects users' distance estimates. *Hum. Factors*
2007, *49*, 507–517. [CrossRef]
44. Geuss, M.; Stefanucci, J.; Creem-Regehr, S.; Thompson, W.B. Can I pass? Using affordances to measure perceived size in virtual environments. In Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization, Los Angeles, CA, USA, 23–24 July 2010; pp. 61–64.
45. Murgia, A.; Sharkey, P.M. Estimation of distances in virtual environments using size constancy. *Int. J. Virtual Real.* **2009**, *8*, 67–74.
[CrossRef]
46. Song, G.Y.; Jiang, R.P.; Zhang, X.Y.; Liu, S.Q.; Yu, X.N.; Chen, Q.; Weng, X.R.; Wu, W.Z.; Su, H.; Ren, C.; et al. Validation of subjective and objective evaluation methods for orthodontic treatment outcome. *J. Peking Univ. Health Sci.* **2015**, *47*, 90–97.
47. Campbell, S.; Goldstein, G. Angle's Classification—A Prosthodontic Consideration: Best Evidence Consensus Statement.
J. Prosthodont. **2021**, *30* (Suppl. 1), 67–71. [CrossRef]
48. Rossini, G.; Parrini, S.; Castroflorio, T.; Deregibus, A.; Debernardi, C.L. Diagnostic accuracy and

measurement sensitivity of digital models for orthodontic purposes: A systematic review. *Am. J. Orthod. Dentofac. Orthop.* **2016**, *149*, 161–170. [CrossRef]
49. Kumar, A.A.; Phillip, A.; Kumar, S.; Rawat, A.; Priya, S.; Kumaran, V. Digital model as an alternative to plaster model in assessment of space analysis. *J. Pharm. Bioallied Sci.* **2015**, *7* (Suppl. 2), S465. [CrossRef] [PubMed]
50. Reuschl, R.P.; Heuer, W.; Stiesch, M.; Wenzel, D.; Dittmer, M.P. Reliability and validity of measurements on digital study models and plaster models. *Eur. J. Orthod.* **2016**, *38*, 22–26. [CrossRef]
51. Priot, A.E.; Charbonneau, M.; Paillé, D. Spatial constraints for 3D perception in Helmet-Mounted Displays. In *Head-and Helmet- Mounted Displays XIII: Design and Applications*; International Society for Optics and Photonics: Bellingham, WA, USA, 2008; Volume 6955, p. 69550G.
52. Aglioti, S.; DeSouza, J.F.; Goodale, M.A. Size-contrast illusions deceive the eye but not the hand. *Curr. Biol.* **1995**, *5*, 679–685. [CrossRef]
53. Pavani, F.; Boscagli, I.; Benvenuti, F.; Rabuffetti, M.; Farnè, A. Are perception and action affected differently by the Titchener circles illusion? *Exp. Brain Res.* **1999**, *127*, 95–101. [CrossRef]
54. Kenyon, R.V.; Phenany, M.; Sandin, D.; Defanti, T. Accommodation and size-constancy of virtual objects. *Ann. Biomed. Eng.* **2008**, *36*, 342–348. [CrossRef]
55. Frey, J.; Hachet, M.; Lotte, F. EEG-based neuroergonomics for 3D user interfaces: Opportunities and challenges. *Trav. Hum.* **2017**, *80*, 73–92. [CrossRef]
56. Kelly, J.W.; Hammel, W.W.; Siegel, Z.D.; Sjolund, L.A. Recalibration of perceived distance in virtual environments occurs rapidly and transfers asymmetrically across scale. *IEEE Trans. Vis. Comput. Graph.* **2014**, *20*, 588–595. [CrossRef]
57. Ebrahimi, E.; Altenhoff, B.M.; Pagano, C.C.; Babu, S.V. Carryover effects of calibration to visual and proprioceptive information on near field distance judgments in 3d user interaction. In Proceedings of the 2015 IEEE Symposium on 3D User Interfaces (3DUI), Arles, France, 23–24 March 2015; pp. 97–104.
58. Trabucco, J.T.; Rottigni, A.; Cavallo, M.; Bailey, D.; Patton, J.; Marai, G.E. User perspective and higher cognitive task-loads influence movement and performance in immersive training environments. *BMC Biomed. Eng.* **2019**, *1*, 21. [CrossRef]
59. Goodale, M.A.; Milner, A.D.; Jakobson, L.S.; Carey, D.P. A neurological dissociation between perceiving objects and grasping them. *Nature* **1991**, *349*, 154–156. [CrossRef]
60. Goodale, M.A.; Westwood, D.A. An evolving view of duplex vision: Separate but interacting cortical pathways for perception and action. *Curr. Opin. Neurobiol.* **2004**, *14*, 203–211. [CrossRef]
61. Goodale, M.A.; Meenan, J.P.; Bülthoff, H.H.; Nicolle, D.A.; Murphy, K.J.; Racicot, C.I. Separate neural pathways for the visual analysis of object shape in perception and prehension. *Curr. Biol.* **1994**, *4*, 604–610. [CrossRef]
62. Makaremi, M. The role of cognitive sciences in orthodontic treatments: Optimizing the interface between practitioners and newtechnologies. *J. Dentofac. Anom. Orthod.* **2016**, *19*, 410. [CrossRef]

Annexes-2 :

- Estimation of distances within real and virtual dental models as a function of task complexity.

- En cours de soumission

Estimation of distances within real and virtual dental models as a function of task complexity.

Abstract

Objectives: The shift from plaster models to virtual digital models has naturally induced new scientific questions within the orthodontic community. In particular, in orthodontic practice, most diagnoses are based on distance estimates (inter-canine distance, crowding, etc.), and these estimates very often condition the entire therapeutic intervention process. However, no study has directly compared distance estimation performed in real (on casts) or virtual (on a computer screen) environments by professionals for diagnostic purposes. Furthermore, in orthodontic practice, the models to be estimated to make therapeutic decisions are of varying complexity. However, to our knowledge, no study has focused on the estimation of distances (real or virtual) according to the complexity of the task to be performed.

Methods: A total of 50 orthodontists (31 women and 19 men), mean age 36 years (σ = 12.84; min = 23; max = 63) have participated in an experiment consisting to collect the estimation of distances on real and virtual models made by orthodontists. Each participant had to estimate 3 quantitative variables (mandibular crowding - inter-canine distance - inter-molar distance) on 6 dental models (3 real and 3 virtual). These models were categorized in 3 levels of complexity (easy - medium - difficult) with one real and one virtual model per level. We also evaluated the mental load associated with each model according to their nature (real or virtual) and their level of complexity.

Results: Results show that globally the distances are overestimated whatever the situation, but, this overestimation is more important in virtual than in real [$F(1, 49) = 36.61$, $p < .05$]. Moreover, the overestimation is larger in the easy condition, and decreases in the more complex conditions. Indeed, the mental load was greater in the virtual conditions than in the real conditions for the 3 levels of complexity [$F(1, 49) = 16.778$, $p < .05$], and for both situations

(real and virtual), the mental load increased with the complexity of the task [$F(2, 98) = 5.331$, $p < .05$].

Conclusion: These results show that there is a difference in the accuracy of distance estimation performed on virtual and real dental models, and that the complexity of the task also has an impact on distance estimation. It is important to continue these studies in order to better understand the parameters that impact the estimation of distances in virtual as, in the future, the use of digital tools for diagnosis or therapeutic intervention itself is widely expected to develop.

Key words: Distance estimation - Virtual environments - Real environments - Task complexity - Orthodontics

Introduction

Virtual Reality and Medicine

Virtual reality (VR) can be defined as a tool that allows "a person (or several people) to engage in sensorimotor and cognitive activity in an artificial, digitally created world that can be imaginary, symbolic or a simulation of some aspect of the real world" (Fuchs et al., 2006). This technology allows, to access and interact with 3D objects and 3D spatial representations named virtual environment (VE) (Mikropoulos & Natsis, 2011). While virtual environments are used in many different industries today, medicine is one of the main sectors (L. Li et al., 2017) and gives rise to 4 main application areas (Pensieri & Pennacchini, 2016): communication interfaces, psychotherapy, surgical planning, and medical training.

Physical models and Virtual models

VR also enables 3D anatomical reconstruction, providing practitioners with the ability to interact with 3D objects (L. Li et al., 2017). In particular, these anatomical models can be used for diagnostic and preoperative planning purposes as well as in learning anatomical concepts.

Although anatomical areas relevant to diagnosing and planning surgery can be identified with 2D images, surgeons may have difficulty perceiving the spatial relationships between structures prior to surgery. Thus, the manipulation of anatomical structures in 3D has been shown to have advantages over traditional 2D visualization. In the case of liver surgery, 3D visualization has allowed surgeons to improve understanding of the three-dimensional relationships of the tumor with intrahepatic vascular networks (Lamadé et al., 2000), liver volume, and tumor (Soler et al., 2014). These 3D reconstructions have led to a significant improvement in tumor localization and to relevant proposals for diagnosis and surgical interventions (Lamadé et al., 2000; Soler et al., 2014). Moreover, in the case of aneurysm surgery requiring a precise spatial understanding of the

vascular anatomy, VR has shown good possibilities of anticipating difficulties and potential complications allowing for improved preoperative planning (R. A. Kockro et al., 2016).

Also used in anatomy training and learning, studies show an increase in engagement, motivation, enjoyment of learning, and spatial understanding of anatomy compared to traditional learning methods (Azer & Azer, 2016; Birbara et al., 2020; R. Kockro et al., 2015; Sakellariou et al., 2009).

However, even though VR shows many advantages, a review of the literature, (Yammine and Violato, 2016) indicate that physical models are superior to virtual models for the acquisition of anatomical knowledge and spatial skills. For example, Khot et al., (2013) compared learning of pelvic system anatomy using physical models or VR and show that physical models result in better learning than virtual models. Another study (Preece et al., 2013) found similar results for learning the anatomy of a horse's foot by third-year veterinary school students. Furthermore, the authors observed that physical models outperformed virtual models in a subsequent MRI image comprehension task and increased students' confidence in their recognitions of anatomical structures in these images. In addition, Wainman et al. (2018) showed that students who learned on physical models performed better on a pelvic anatomy analysis task on either physical or virtual models compared to learning using VR. Finally, these authors, in a recent study (Wainman et al., 2020) showed that learners' visuo-spatial abilities influenced learning in VR but not learning using physical models. Low visuo-spatial ability resulted in poorer learning with VR than high visuo-spatial ability, while for physical models learners' visuo-spatial ability did not impact learning.

Virtual Reality in orthodontics or dentofacial orthopedics

Dentofacial orthopedics, or orthodontics, is a medical field in which the use of virtual reality has also developed, particularly due to the creation of virtual models of dental arches. Dental arch

models are of great importance to orthodontists (Marcel, 2001). They are an essential part of the diagnosis and are used to document the original condition and to plan and measure the effects of a therapeutic treatment. The generation of three-dimensional models to accurately represent dental arches and occlusions (how the mandibular teeth fit together with the maxillary teeth) was previously done using plaster casts obtained from impressions of the dental arch. Today, digital 3D models have become a dematerialized alternative (Kravitz et al., 2014). Two ways of obtaining these digital models are known. The first is the optical impression, which is made directly, using an intraoral camera (or intraoral scanner). The camera acquires numerous intraoral images which are then transmitted to a computer or an adapted acquisition unit, on which a software processes the data in order to create a virtual model in three dimensions. The contributions of such technology in orthodontics are multiple and have been described numerous times (Coblentz, 2012; Kravitz et al., 2014; Mangano et al., 2017). The second method is to obtain digital models from a 3D scan of a physical cast or impression. In this case, the term indirect method is used because all the "digital" steps are performed in the laboratory using an extra-oral scanner from a physical model and not directly on the patient.

Studies have shown the reliability of measurements performed on digital models compared to measurements on plaster models (Kumar et al., 2015; Reuschl et al., 2016; Rossini et al., 2016). These measurements are usually performed with a caliper for plaster models and with dedicated software (for calculating distances between 2 points) for digital models. In the study by Radeke et al. (2014), measurements were performed by participants classified according to their level of experience: laypersons, dentists (non-experts in orthodontics) and orthodontists. The authors concluded that in clinical practice, all measurements taken on plaster models are similar to measurements taken directly on real models. However, accuracy and reproducibility varied with the level of orthodontic expertise, with the highest accuracies obtained by orthodontists (Radeke et al., 2014).

Another study (Sousa et al., 2012), obtained similar results with the use of different measurement software and scanner. Czarnota et al. (2016) on the other hand, found an underestimation of distances on the numerical models when performing a similar protocol. However, these differences and systematic errors were small and clinically non-significant. The authors conclude that the measurement of parameters on digital models is a reliable method and can be used in clinical practice. However, in the cited studies, distances are measured using measurement tools and few studies have focused on the potential differences in distance estimation on virtual or physical models performed by practitioners without measurement tools. This procedure (distance estimation without measurement tools) is however part of the clinical routine of professionals.

Virtual reality and distance estimation

As digital models are increasingly used in professional practice, it is important that the perception of features within VR applications is reliable and accurate for these applications to be effective.

However, the perception of distances, depth and size of objects within VR is often incorrect. Indeed, in their review, Renner et al. (2013) report that the distances separating an object from an observer are generally underestimated. On the other hand, several studies show that distances separating two objects are overestimated. For example, Waller (1999) compared the perceived distance between 2 cubes presented in a VE using a VR headset or a desktop computer and showed that the distance was overestimated in both conditions. In another study, Wartenberg & Wiborg (2003) investigated the perception of distance between 2 points within three environments: the real environment, an EV presented with a CAVE system, and an EV presented using a desktop computer. Distances were misperceived regardless of the presented environment and the greater the distances to be estimated the greater the estimation errors. In the real environment and in the VE presented in a CAVE system, a prevalence of overestimation of distances was present, while in the VE presented with a desktop computer, distances were either

overestimated or underestimated. Perceived distances between several elements of a complex object also appear to be overestimated. Tourancheau et al. (2012) studied distance and depth perception within a 3D puzzle presented using a stereoscopic VR system and show that distances were overestimated by about 12%. Objects, on the other hand, would be perceived as smaller within VE (Stefanucci et al., 2015).

Distance estimation in non-immersive environments

Studies have also focused on distance estimation within non-immersive environments, i.e., a VE in which "the user is faced with a simple screen and some devices that allow him to interact with the VE" (Chellali, 2009).

While exocentric distance is generally overestimated within immersive virtual environments, one study has shown that underestimation and overestimation can occur equally within non-immersive VEs (Wartenberg & Wiborg, 2003). However, another study found that exocentric distance was overestimated regardless of the type of device involved, immersive or not (Waller, 1999). According to the latter study, the overestimation of the distance between two objects is higher with a desktop computer than with a VR headset. Furthermore, Watenberg and Wilborg (2003) observed that the error rate increases as the distance to be estimated becomes larger within VEs, an effect amplified by interacting with a non-immersive environment.

Thus, in non-immersive VEs the distance estimation is erroneous and this error would be greater with a non-immersive environment compared to an immersive environment.

Aims

In orthodontic practice, most diagnoses are based not on actual measurements but on distance estimates (inter-canine distance, crowding, etc.), and these estimates very often condition the entire therapeutic intervention process. However, even if, as we have seen, the estimation of distances performed in real or virtual environments can be different, no study has directly compared the estimation of distances performed in real (on casts) or virtual (on a computer

screen) environments by professionals for diagnostic purposes. Moreover, the environments used in orthodontic practice are non-immersive, which we have seen to increase estimation errors compared to immersive environments. Finally, in orthodontic practice, the models to be estimated to make therapeutic decisions are of variable complexity. However, to our knowledge, no study has investigated the estimation of distances (real or virtual) as a function of the complexity of the task to be performed. The complexity of the models presented in a non-immersive virtual environment or in a real environment will also be taken into account in our study.

Method

Subjects

A total of 50 orthodontists (31 women and 19 men), mean age 36 years (σ = 12.84; min = 23; max = 63) participated in this experiment. A part of the participants were in the process of specialization at the university (n = 23; first year = 7; second year = 7; third year = 9), and another part were graduated dental surgeons and specialized in orthodontics (n = 27). The training sites of the participants were the University of Bordeaux (n = 31), Rennes (n = 13), Toulouse (n = 3), Nantes (n = 1), Montpellier (n = 1), and Paris (n = 1). All participants were subject to all conditions.

Material

To carry out this study we used models of dental arches and questionnaires that the participants filled in after completing the tasks. The order of presentation of each model, the order of presentation of the levels of complexity as well as the choice of the models presented in real and virtual were chosen randomly beforehand.

Dental arch models

As a preamble to the research we identified dental arch models of different complexity (3 levels of complexity were selected). For this, we combined the intra-arch parameters of 3 indices used in the classification of case difficulty in orthodontics (Bellot-Arcis et al., 2012):

- The DHC (dental health component): morphological component of the IOTN (Index of Orthodontic Treatment Need);
- The PAR Index (Peer Assessment Rating Index);
- The DAI (Dental Aesthetic Index).

For methodological reasons, each model had to meet the following inclusion criteria: stable and complete adolescent, young adult or adult dentition; absence of agenesis or inclusion; and absence of extracted teeth.

This allowed us to divide the dental models into three categories of difficulty:

- Easy: anterior crowding only, with contact point shift of 1 to 2mm, without arch shape anomaly (normal inter-canine and inter-molar distances), nor dental anomaly;
- Medium: anterior crowding resulting in a contact point shift of 2 to 4mm, with rotation or ectopia of one tooth per sector; asymmetry of less than 2mm and/or pathological arch shape;
- Difficult: anterior and medium crowding resulting in contact point shifts of more than 4mm, and/or ectopic positions and rotations of several teeth per sector; asymmetry greater than 2mm and/or pathological arch form.

The selected mandibular models were taken from the clinical database of an orthodontic practice in Bergerac (24100). The search was carried out among the patients on whom initial impressions had been made using the 3Shape Trios3® camera.

After pre-selection, 64 cases were retained. These 64 cases were segmented and analyzed with the Orthoanalyzer® software to define the measurement of the parameters taken into account in

our study (mandibular crowding, inter-canine distance, inter-molar distance) and each case was categorized as easy, medium or difficult. For each level of difficulty, 2 models were selected for their clinical interest (6 models in total).

Then, the 6 mandibular models were printed at real scale with a DLP (Digital Light Processing) printer: the NextDent 5100® from 3DSystems and via the 3D software SprintBasic®. Plaster duplicates of these models were made after taking alginate impressions (Figure 1).

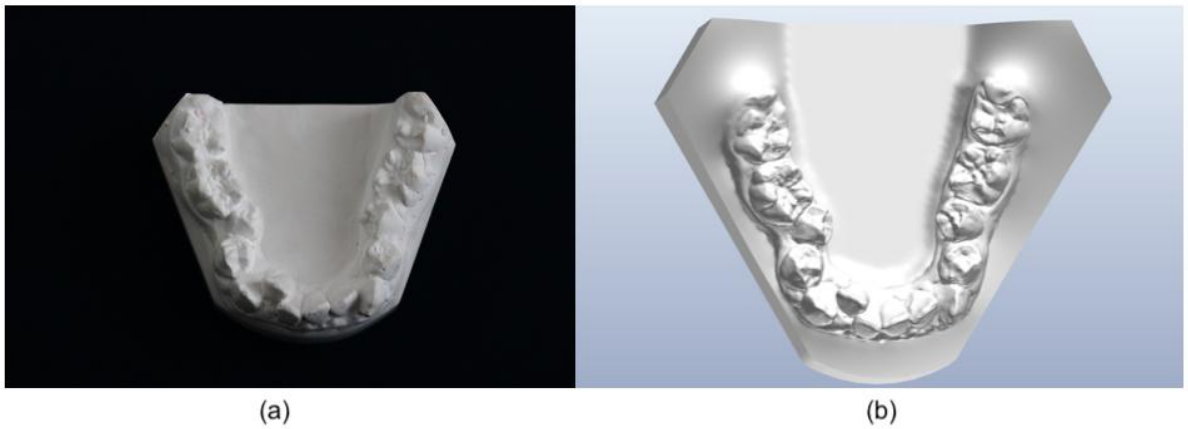


Figure 1: Example of (a) real mandibular model used and (b) virtual mandibular model used

The virtual 3D models were saved in dicom format on a single computer so that they could be presented to participants under the same experimental conditions. They were presented using a computer with a traditional 17-inch 2D monitor (DELL-TRPO030351 Intel(R) Core (TM) i7-7700HQ CPU 2.80 GHz) with no 3D stereo image.

The actual measurements of the footprint of mandibular crowding, inter-canine and inter-molar distance having already been performed on the Orthoanalyzer® software, were again performed on the plaster models using a digital caliper and brass wire (traditional measurement). This operation was performed by two different operators.

Questionnaires

Various questionnaires were submitted to each participant.

Demographic data

Prior to the completion of the tasks, a demographic data collection questionnaire was completed by the participants (gender, age, year they started orthodontic practice, year of residency if they were interns, place of training/specialization, and use of virtual and/or real models in their training/profession).

A questionnaire was also developed to collect the participants' experiences and computer skills.

Mental load

In order to verify a posteriori the relevance of the classification into easy, medium and difficult made before the experimentation, each participant had to evaluate the mental load associated with each of the models for which he had to estimate the distances.

The NASA-TLX is a questionnaire that uses a subjective measurement technique, i.e., it relies on the judgment of the person performing the task or action. It was developed in the 1980s by the Human Performance Group of the NASA-Ames Research Center, CA (Sandra Hart). The NASA-TLX is composed of 6 dimensions: mental demand, physical demand, temporal demand, performance, effort and frustration. These dimensions are each presented in the form of a Likert scale (ranging from 0 to 20), which provides a score for each dimension and an overall score per participant. The questionnaire was presented in its simplified version (Byers, Bittner & Hill, 1989).

Procedure

For each participant, a session was divided into three phases. The first was the "presentation" phase, which included signing the consent form, describing and explaining the task to the participants. Then, the participants were asked to fill in the questionnaire concerning their

experience with the computer tool. This questionnaire was presented on a tablet in the form of a Google Forms.

Then, the second phase took place: the estimation of distances on the dental models. The order of presentation of each model, the order of presentation of the levels of complexity as well as the choice of the models presented in real and virtual were randomly established beforehand.

Participants were asked to estimate three quantitative variables on each model, whether they were presented in real (plaster) or virtual (on a computer screen). The virtual model estimates were performed on the same computer for all participants and the plaster model estimates were performed on the same white desk for all participant. For each of the estimations the participants were allowed to manipulate the models. In the real condition the plaster models could be picked up manually while in the virtual condition the models were manipulated using a 2D mouse on the OrthoViewer® software.

The three variables to be estimated were mandibular crowding, inter-canine distance, and inter-molar distance (Figure 2). These values are expressed in millimeters (mm) and can be rounded to the nearest millimeter.

- The crowding is an intra-arch parameter, which can be evaluated on the maxillary and mandibular arches. It is taken into account in the assessment of dento-maxillary disharmony. In this study, we asked the participants to perform a crowding estimation only on the mandibular arch and on the ten anterior teeth (from the mesial side of 36 to the mesial side of 46). The actual measures were 2,37mm and 5,02mm for the easy complexity condition, 5,00mm and 5,15mm for the medium complexity condition and 9,77mm and 11,15mm for the difficult complexity condition;

- The inter-canine distance is the straight transverse distance measured between the top of the right and left canine tips). The actual measures were 24,00mm and 24,09mm for the easy complexity condition, 21,63mm and 25,35mm for the medium complexity condition and 26,46mm and 27,14mm for the difficult complexity condition;

- The inter-molar distance is the straight transverse distance joining the apex of the mesiovestibular cusps of the 1st permanent molars). The actual measures were 42,96mm and 45,22mm for the easy complexity condition, 37,75mm and 43,34mm for the medium complexity condition and 38,07mm and 41,35mm for the difficult complexity condition;

The first two values (mandibular crowding and inter-canine distance) are part of the diagnosis classically made by professionals. The inter-molar distance, on the other hand, is a control condition that is never estimated in practice.

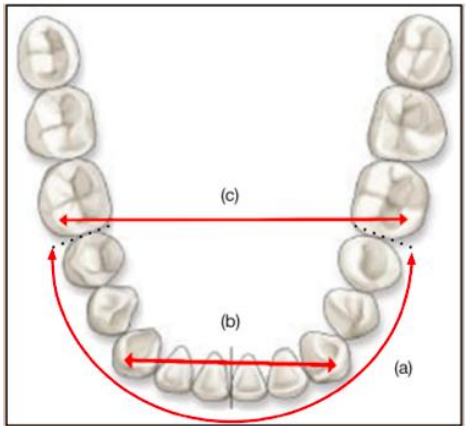


Figure 2: Indicative diagram of (a) mandibular crowding, (b) inter-canine distance and (c) inter-molar distance

After estimating the distances on a dental model, the participants were asked to complete the NASA-TLX questionnaire. They completed the questionnaire 6 times (Figure 3).

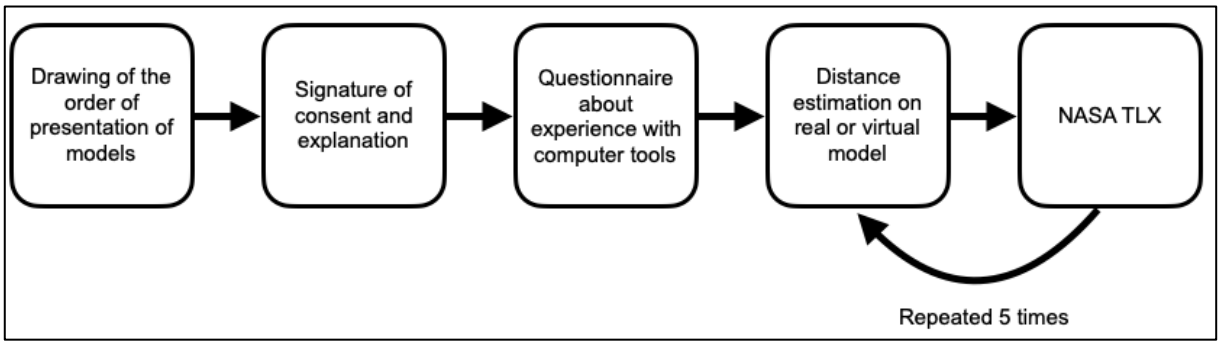


Figure 3: Experimental procedure

Statistical analysis

Results was analysed in two steps. The first was to perform repeated measures Anova's to analyze the participants' performance in estimating distances. In these Anova's, the two factors

were the "presentation" factor (real vs. virtual) and the "complexity" factor (3 levels), which were two between-subjects factors. The dependent (quantitative) variables were the distance estimate in the three situations mentioned above (mandibular crowding, inter-canine distance and inter-molar distance). More precisely, for each situation, the dependent variable corresponded to an indicator that was an error score = value (distance) estimated by the practitioner - real value (distance). The distance was overestimated when the indicator was greater than 0, and underestimated when the indicator was less than 0. For each participant, the presentation of the 6 conditions (3 levels of complexity x 2 presentation conditions) was counterbalanced.

The mental load is used here as a control. A two-factor anova (with presentation and complexity as factors and score at the mental load test as dependent variable) was used to validate the easy, medium and difficult classification carried out before the study, and more generally to evaluate the mental load according to the level of complexity and the virtual or real character of the estimated model.

Before each ANOVA was performed, the samples were subjected to the normality test (Shapiro test) as well as the homogeneity of variances test (Levene test). These analyses indicated that all data followed normal distributions and that the variances were homogeneous.

Results

Mental load

The results of the 2 ways analysis of variance showed a main effect of the type of presentation (real vs. virtual) [$F(1, 49) = 16.778$, $p < .05$], a main effect of the level of complexity [$F(2, 98) = 5.331$, $p < .05$], and no interaction effect between type of presentation and level of complexity [$F(2, 98) = 1.342$, $p < .05$].

The descriptive analyses (Table 1) show that mental load was greater in virtual ($M = 53.26$) than in real ($M = 47.91$) conditions for the 3 levels of complexity. And for both situations (real and

virtual), the mental load score was lowest in the easy condition (M = 48.35), then increased in the medium condition (M = 50.32), and was highest in the difficult condition (M = 53.08).

Mental Load						
	Easy		Medium		Difficult	
	M	σ	M	σ	M	σ
Real	44.9	17.68	47.4	19.19	51.42	20.06
Virtual	51.8	18.73	53.24	17.67	54.74	18.86

Table 1. Descriptive statistics (M = mean; σ = standard deviation) of mental load by task complexity level and presentation type (real vs. virtual).

Distance estimates

Distance estimate in function of the 6 actual measures

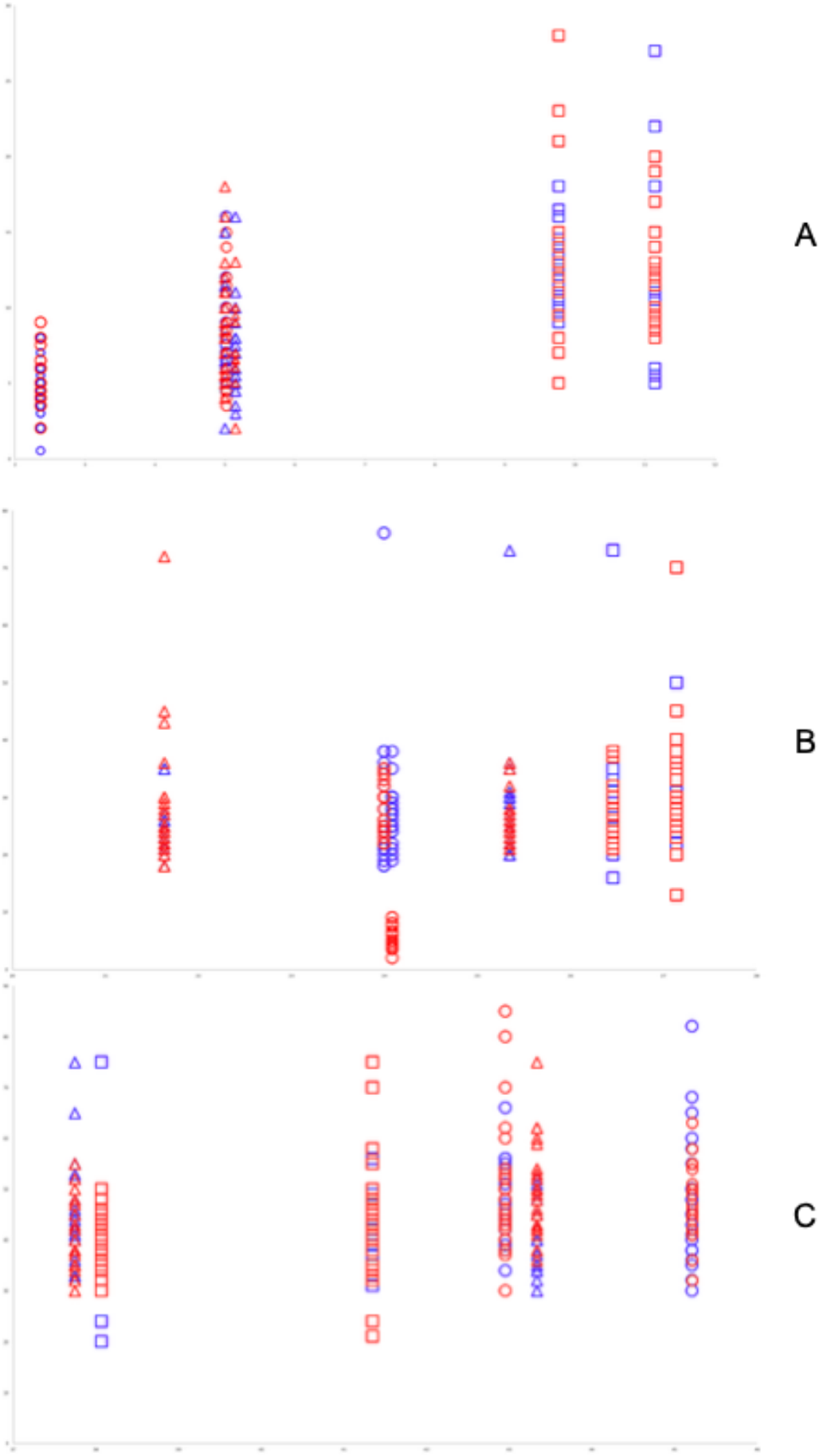


Figure 4: Distances estimated by the practitioners in function of the 6 actual measures for (A) the mandibular crowding, (B) the inter-canine distance and (C) the inter-molar distance.

The estimation made on:

- *the 2 easy complexity models are symbolized by **O** on the figure;*
- *the 2 medium complexity models are symbolized by ▵ on the figure;*
- *the 2 difficult complexity models are symbolized by □ on the figure.*

The distances represented in red are the distances estimated on the virtual models and the distances represented in blue are the distances estimated on the real models.

The figure 4 shows the distribution of the distances estimated by the participants in the real condition and in the virtual condition in function of the actual measure they did on the 2 easy complexity models, on the 2 medium complexity models and on the 2 difficult complexity models.

Global analysis

We first performed a global analysis of variance by averaging the scores obtained for the 3 estimation errors situations (mandibular crowding, the inter-canine distance and the inter-molar distance). The analysis shows a main effect of the type of presentation (real vs. virtual) [$F(1, 49) = 36.61$, $p < .05$], a main effect of the level of complexity [$F(2, 98) = 47.476$, $p < .05$], as well as an interaction between type of presentation and level of complexity [$F(2, 98) = 12.744$, $p < .05$].

The descriptive analyses (Table 2) show that globally the distances are overestimated whatever the situation. But, the overestimation is more in virtual ($M = 3.37$) than in real ($M = 2.67$). For the complexity levels, the overestimation is more important in the easy condition ($M = 3.58$), then decreases in the medium condition ($M = 3.23$), and is the lowest in the difficult condition ($M = 2.44$). Moreover, the interaction is due to the fact that the differences between real and virtual (higher overestimation in virtual than in real) disappear in the difficult condition (see also Figure 5).

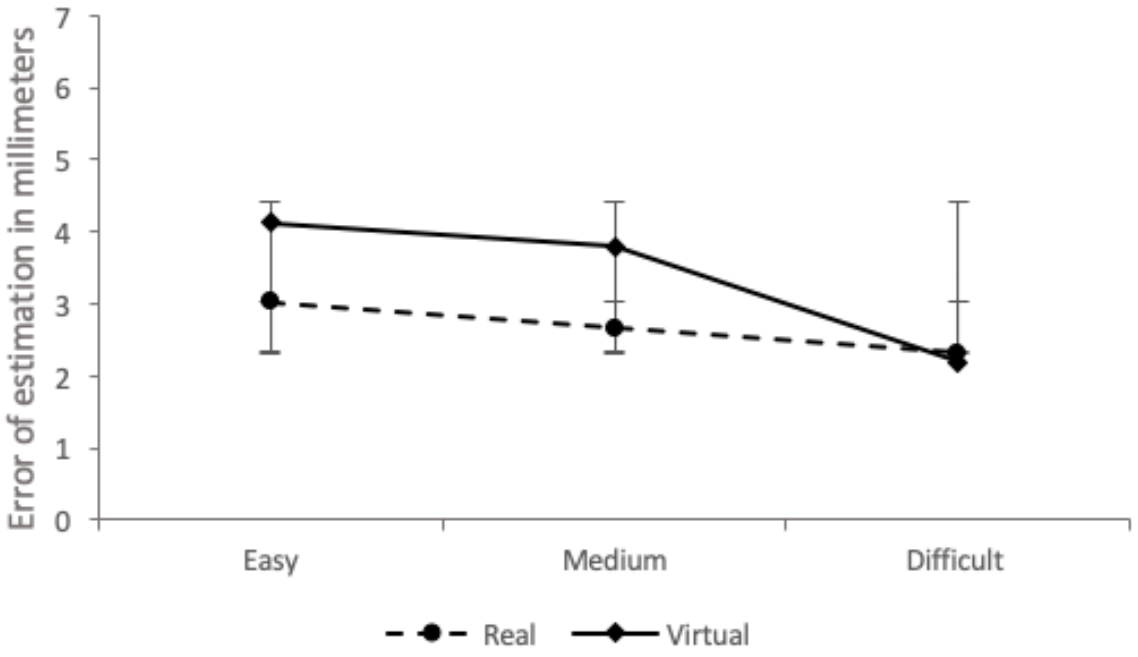


Figure 5: Graphical representation of the average of the errors associated with the estimates of *mandibular crowding, inter-canine distance and inter-molar distance.*

Mandibular crowding

The analysis of variance shows a main effect of presentation type (real vs. virtual) [F (1, 49) = 7.662, $p < .05$], as well as a main effect of complexity level [F (2, 98) = 5.139, $p < .05$]. In contrast, we did not observe an interaction between presentation type and complexity level [F (2, 98) = 0.951, $p < .05$].

Descriptive analyses (Table 2) show that mandibular crowding is overestimated in both real and virtual but that the overestimation is greater in virtual (M = 2.77) than in real (M = 1.19). For complexity levels, the overestimation is smaller in the difficult condition (M = 1.90) than in the easy (M = 2.72) and medium (M = 2.82) conditions.

Inter-canine distance

The analysis of variance shows a main effect of presentation type (real vs. virtual) [F (1, 49) = 6.053, $p < .05$], a main effect of complexity level [F (2, 98) = 4.078, $p < .05$], and an interaction between presentation type and complexity level [F (2, 98) = 3.154, $p < .05$].

Descriptive analyses (Table 2) show that the inter-canine distance is overestimated in both real and virtual, but that the overestimation is greater in virtual (M = 3.63) than in real (M = 2.56). For

the complexity levels, the overestimation is the most important in easy condition (M = 4.03) then decreases in medium condition (M = 2.76) and is the least important in difficult condition (M = 2.36). Finally, the interaction is due to the fact that the difference between virtual and real (overestimation more important in virtual than in real) is present in easy and medium condition but not in difficult condition.

Inter-molar distance

Analysis of variance shows no main effect of presentation type (real vs. virtual) [F (1, 49) = 0.441, p > .05], complexity level [F (2, 98) = 1.740, p > .05], and no interaction between presentation type and complexity level [F (2, 98) = 1.660, p > .05]. The results (no effect of presentation type, level of complexity, and the interaction between these two factors) for inter-molar distance are different from those obtained for the other 2 measures. It should be noted that this is the only measure that orthodontists do not usually use in their practice.

	Mandibular crowding						Inter-canine distance						Inter-molar distance					
	Easy		Medium		Difficult		Easy		Medium		Difficult		Easy		Medium		Difficult	
	M	σ	M	σ	M	σ	M	σ	M	σ	M	σ	M	σ	M	σ	M	σ
Real	2.61	2.18	2.34	2.76	1.62	4.16	3.35	8.75	1.83	7.93	2.48	8.76	3.13	9.79	3.82	9.03	2.83	11.12
Virtual	2.84	2.54	3.29	2.95	2.17	4.21	4.71	9.08	3.69	8.88	2.25	8.27	4.84	10.67	4.39	8.10	2.13	9.32

Table 2. Descriptive statistics (M = mean of the estimation errors; σ = standard deviation) of crowding, inter-canine distance, and inter-molar distance for virtual and real conditions.

Effect of Experience by using a regression analysis

We also tested the effect of the practitioner's orthodontic experience on distance estimation using a regression (figure 6). This regression did not show a significant correlation between estimation errors and the orthodontic experience of the practitioners. whatever the distance estimated [mean of the 3 distances (R= -0.053, p > .05), mandibular crowding (R= -0.023, p > .05), inter-canine distance (R= -0.046, p > .05) and inter-molar distance (R= -0.090, p > .05)].

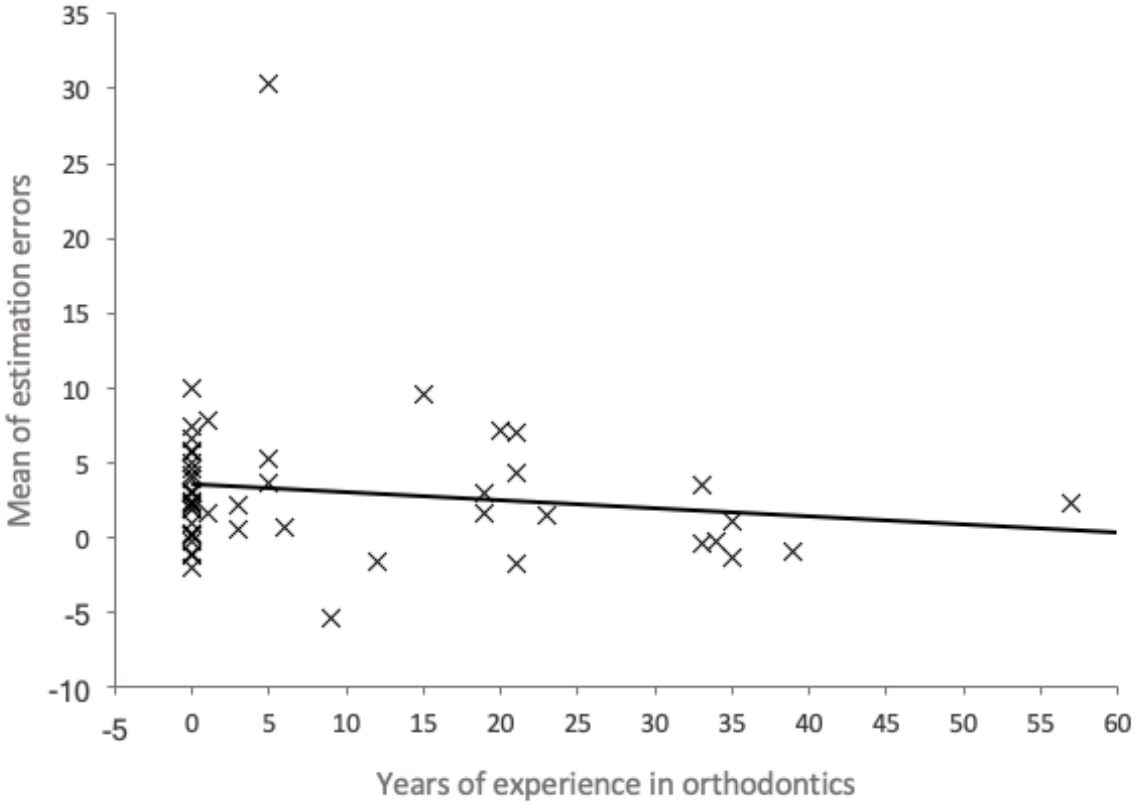


Figure 6: Linear regression of the mean of the estimation errors according to the experience level of the practitioners

It should be noted that a correlation is interpreted using the correlation coefficient but also by visualizing the distribution of subjects on the regression line. However, as shown in Figure 6 a large number of participants are still novices which does not allow an optimal distribution of the points on the regression line, making it difficult to interpret this result.

Discussion

The results of the analysis of variances performed on the average of the three estimated distances showed that the distances were overestimated in both real and virtual environments. These results are consistent with the results found by other authors. Several researches have shown overestimation of mandibular crowding distance when estimated by orthodontists viewing real models (Johal & Battagel, 1997; Naish et al., 2016; Wallis et al., 2014). Makaremi & N'Kaoua, (2021) showed an overestimation of mandibular crowding distances estimated by 31 orthodontists on virtual models. In addition, several studies have investigated the perception of

exocentric distances (distances between two objects, in our case between two elements of the same object) within non-immersive virtual environments. For example, Waller (1999) found that exocentric distance within non-immersive devices was overestimated. However, Wartenberg & Wiborg (2003) demonstrated that exocentric distance was generally overestimated within immersive VEs, while underestimation and overestimation can occur equally within non-immersive VEs. Furthermore, these authors observed that poor distance estimation was amplified by interacting with a non-immersive environment.

Our results show that distances were less well perceived with the virtual models than with the plaster models. Indeed, distances were more overestimated in the VR condition than in the real condition.

A possible explanation lies in the effect of various non-pictorial depth cues on distance estimation. These cues correspond, among other things, to information in the human binocular oculomotor system that allows depth perception. Thus, binocular disparity and accommodation are of great importance in distance perception and estimation, especially in the personal space from the observer to 2 meters (Cutting & Vishton, 1995).

Plaster models provide natural access to various non-pictorial depth cues, whereas in the case of virtual models, presented here using a computer screen, these cues provide depth information relative to the distance between the user and the screen and not between the user and the objects themselves (Stefanucci et al., 2015). Binocular disparity refers to the difference between images perceived by the right and left eyes (Cutting & Vishton, 1995). Wainman et al, (2018) showed, in a task requiring spatial understanding of female pelvic anatomy, the importance of this cue in the difference in performance between physical and virtual models. In this study, participants had the nondominant eye hidden (eliminating binocular vision) or not when estimating distance on virtual or physical models. When binocular vision was available, the physical model led to better performance than the virtual model. However, without binocular vision, performance on the

physical model decreased while it did not change for the virtual model (no significant difference between the physical and virtual models in the condition without binocular vision). These results indicate that performance in spatial analysis is diminished by the limited information provided by binocular vision on virtual models viewed with non-immersive VR devices compared to physical models.

Furthermore, in the virtual situation, accommodation occurs on the plane constituted by the computer screen rather than on the virtual objects themselves, which would decrease depth visual acuity (Huckauf, 2005). As a result, distance perception of virtual models would be misdirected by accommodation (Renner et al., 2013).
Thus, the lack of information from accommodation and binocular disparity on features presented within the non-immersive VE are therefore factors that may have decreased distance estimation.

Furthermore, in our study, participants used a keyboard and mouse to interact with the non-immersive VE. However, this method would not be optimal for distance perception. Indeed, the manipulation of 3D objects with a 2D cursor (an arrow that can only move in 2 dimensions) controlled with a mouse would lead to a conflict between the 2D space of the cursor and the 3D space of the content, in addition to not giving any clue about the depth of the selected objects (Jauregui et al., 2012). Thus, interaction methods may have also played a role in decreasing performance in distance estimation within virtual models.

Other factors, such as the type of learning performed may also be involved in the decrease in performance on virtual models compared to physical models. Indeed, Wainman et al, (2018) analyzed the effect of the type of learning (learning performed on physical model vs. virtual model) of the female pelvis on a task requiring spatial analysis on a computer screen. These authors showed that learning on a physical model led to better performance than learning on a

virtual model. The effect of the type of training (professionals were trained on either real models, virtual models, or both types of models during their studies) and visuospatial abilities could be investigated in future studies.

Finally, a last source of explanation is that tasks performed on a virtual model result in a higher mental load than those performed on a real model. Indeed, an increase in mental load is associated with a decrease in performance (Chanquoy et al., 2007; Jaquess et al., 2017). Yet, one of the results of our study was that mental load was higher when estimates were performed on a computer screen than when they were performed on the plaster models. This result is in agreement with the study of Stevens et al, (2016) which showed an increase in mental load in tasks performed in non-immersive VR compared to the same tasks performed in the real environment. Indeed, several authors show that interaction with an EV, due to its complex nature, could increase cognitive load (Fraser et al., 2012; Huang et al., 2020).

Thus, in our study, to compensate for the lack of non-pictorial visual cues in the virtual model condition, practitioners may have had to rely on their visuospatial skills and/or other types of cognitive processing to apprehend the virtual model and estimate distances in the model, resulting in increased mental load and decreased performance. Again, this type of hypothesis requires further investigation.

Regarding the effect of model complexity, our results indicate that as the level of complexity increases, the mental load increases for both plaster and virtual models. These results are in agreement with the literature showing that high mental load is associated with high task complexity (Chao et al., 2017; Das et al., 2020; Jacko & Ward, 1996; J. Li & Kim, 2021). For example, Haji et al. (2015) showed that mental load was higher in a complex task than in a simple task during a physical simulation of a knot to secure a chest tube. Di Stasi et al (2016), on

the other hand, found that mental load increased with the level of task complexity in different laparoscopy exercises performed in VR. Our results show that this is also the case in distance estimation tasks performed in real or virtual using non-immersive technologies, where increasing the complexity of the models to be estimated increases mental load. Our results also show that the increase in complexity is associated with an increase in mental load but with an improvement in distance estimation performance (for mandibular crowding and inter-canine). Yet, increased mental load is generally associated with decreased performance (Chanquoy et al., 2007; Jaquess et al., 2017). It should be noted, however, that some studies have shown that the mental load/performance relationship is more complex and may have an inverted U shape (Babiloni, 2019; Basahel et al., 2010; Bruggen, 2015). Performance can increase with mental load until it reaches a threshold at which it decreases as mental load increases.

As shown above, our results show that increasing model complexity (virtual and real) was associated with better distance estimates for mandibular crowding and inter-canine distance while no difference in estimates as a function of task complexity was found for inter-molar distance. Such a result could then be explained by the distance estimation questions performed in our study. Indeed, the mandibular crowding and the inter-canine distance have an important role in the diagnosis and thus in the treatment that will be implemented, whereas the inter-molar distance is a condition that is not encountered in professional practice. Authors have shown that task objectives mediate the relationship between performance and task complexity. Important goals in complex tasks can improve performance through motivational and cognitive processes that lead to the development of effective strategies (Campbell, 1988; Wood et al., 1987). Thus, in our study, distance estimation in complex cases may be associated with issues of more severe surgery or potentially more impactful subsequent management. These higher stakes when estimating complex models could be the cause of increased attention when taking distance information on both real and virtual models, thus improving performance. Such a hypothesis obviously remains to be tested in further studies, especially in terms of the impact of the therapeutic issue on the

accuracy of estimation responses in real and virtual environments. Another possible source of explanation for this complexity effect on distance estimation performance may be in terms of experience level. However, our study did not find any relationship between experience level and distance estimation accuracy (regardless of the estimation condition considered), although, as noted in the results section (Section 3.2.4), the correlation results should be taken with caution. Future studies that more systematically take into account the level of experience will also help to clarify the link between level of complexity, level of experience and distance estimation in real and virtual situations (for example by comparing different groups of practitioners with different levels of experience).

Finally, other complementary studies could also compare distance estimates using different VR devices, from the least immersive (like the one we used) to the most immersive (head-mounted display, for example).

Conclusion

The technological evolution in the field of orthodontics is gradually leading to a shift from real dental models to virtual models. This study showed that there is a difference in accuracy when estimating distances on dental models between the two environments, and that practitioners are more accurate when performing this task on real models than on virtual models. The complexity of the task to be performed also has an impact on the distance estimation performed by practitioners. It is important to continue these studies in order to better understand the impact of variables (related to professional goals and experiences, hardware, type of model, etc.) on estimation tasks as, in the future, the use of digital tools for diagnosis or intervention itself is likely to grow.

Summary table

Previous findings	Findings added by our study
• Distances are overestimated when estimated by orthodontists visualizing real models or virtual models (Johal & Battagel, 1997; Makaremi & N'Kaoua, 2021; Naish et al., 2016; Wallis et al., 2014). • Many factors such as technical factors, environmental composition factors and human factors have an impact on the estimation of distances (Renner et al., 2013). • The mental load increases with the complexity of the task.	• Our study showed that the distances were overestimated on real and virtual dental models. Moreover, it adds that this overestimation of distances is more important for virtual models. • Our results showed that the complexity of the task is a factor that have an effect on the estimation of distance. The more complex the model, the more accurate the estimation of distances within the model. • Our study showed that the increase in mental load with task complexity was also present in distance estimation tasks.

Table 3. Summary table of main finding of our study compared to previous finding.

Bibliography

- Azer, S. A., & Azer, S. (2016). 3D Anatomy Models and Impact on Learning : A Review of the Quality of the Literature. *Health Professions Education*, *2*(2), 80-98. https://doi.org/10.1016/j.hpe.2016.05.002
- Babiloni, F. (2019). Mental Workload Monitoring : New Perspectives from Neuroscience. In L. Longo & M. C. Leva (Éds.), *Human Mental Workload : Models and Applications* (Vol. 1107, p. 3-19). Springer International Publishing. https://doi.org/10.1007/978-3-030-32423-0_1
- Basahel, A. M., Young, M. S., & Ajovalasit, M. (2010). Impacts of physical and mental workload interaction on human attentional resources performance. *Proceedings of the 28th Annual European Conference on Cognitive Ergonomics - ECCE '10*, 215. https://doi.org/10.1145/1962300.1962344
- Birbara, N. S., Sammut, C., & Pather, N. (2020). Virtual Reality in Anatomy : A Pilot Study Evaluating Different Delivery Modalities. *Anatomical Sciences Education*, *13*(4), 445-457. https://doi.org/10.1002/ase.1921
- Bruder, G., Argelaguet, F., Olivier, A.-H., & Lécuyer, A. (2016). CAVE Size Matters : Effects of Screen Distance and Parallax on Distance Estimation in Large Immersive Display Setups. *Presence: Teleoperators and Virtual Environments*, *25*(1), 1-16. https://doi.org/10.1162/PRES_a_00241
- Bruggen, A. (2015). An empirical investigation of the relationship between workload and performance. *Management Decision*, *53*(10), 2377-2389. https://doi.org/10.1108/MD-02-2015-0063
- Burdea, G., & Coiffet, P. (1993). *La réalité virtuelle*. Hermès.
- Campbell, D. J. (1988). Task Complexity : A Review and Analysis. *Academy of Management Review*, *13*(1), 40-52. https://doi.org/10.5465/amr.1988.4306775
- Chao, C.-J., Wu, S.-Y., Yau, Y.-J., Feng, W.-Y., & Tseng, F.-Y. (2017). Effects of three-dimensional virtual reality and traditional training methods on mental workload and training performance. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *27*(4), 187-196. https://doi.org/10.1002/hfm.20702
- Coblentz, J.-F. (2012). Empreintes numériques : Mythes et réalité quotidienne. *L'Orthodontie Française*, *83*(1), 81-92. https://doi.org/10.1051/orthodfr/2011149
- Combe, T., Chardonnet, J.-R., Merienne, F., & Ovtcharova, J. (2021). CAVE vs. HMD in Distance Perception. *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 448-449. https://doi.org/10.1109/VRW52623.2021.00106
- Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances : The integration, relative potency, and contextual use of different information about depth. In

Perception of space and motion (p. 69-117). Academic Press. https://doi.org/10.1016/B978-012240530-3/50005-5

- Czarnota, J., Hey, J., & Fuhrmann, R. (2016). Measurements using orthodontic analysis software on digital models obtained by 3D scans of plaster casts. *Journal of Orofacial Orthopedics / Fortschritte Der Kieferorthopädie*, *77*(1), 22-30. https://doi.org/10.1007/s00056-015-0004-2
- Das, S., Maiti, J., & Krishna, O. B. (2020). Assessing mental workload in virtual reality based EOT crane operations : A multi-measure approach. *International Journal of Industrial Ergonomics*, *80*, 103017. https://doi.org/10.1016/j.ergon.2020.103017
- Di Stasi, L. L., Diaz-Piedra, C., Rieiro, H., Sánchez Carrión, J. M., Martin Berrido, M., Olivares, G., & Catena, A. (2016). Gaze entropy reflects surgical task load. *Surgical Endoscopy*, *30*(11), 5034-5043. https://doi.org/10.1007/s00464-016-4851-8
- Drascic, D., & Milgram, P. (1996). *Perceptual issues in augmented reality* (M. T. Bolas, S. S. Fisher, & J. O. Merritt, Éds.; p. 123-134). https://doi.org/10.1117/12.237425
- Ellis, S. R. (1994). What are virtual environments? *IEEE Computer Graphics and Applications*, *14*(1), 17-22. https://doi.org/10.1109/38.250914
- Fraser, K., Ma, I., Teteris, E., Baxter, H., Wright, B., & McLaughlin, K. (2012). Emotion, cognitive load and learning outcomes during simulation training : Emotion and cognitive load during simulation. *Medical Education*, *46*(11), 1055-1062. https://doi.org/10.1111/j.1365-2923.2012.04355.x
- Fuchs, P., Moreau, G., Auvray M. et École nationale supérieure des mines de Paris (2006). *Le traité de la réalité virtuelle*. Paris : Presses de l'Ecole des mines.
- Haji, F. A., Rojas, D., Childs, R., de Ribaupierre, S., & Dubrowski, A. (2015). Measuring cognitive load : Performance, mental effort and simulation task complexity. *Medical Education*, *49*(8), 815-827. https://doi.org/10.1111/medu.12773
- Huang, C. L., Luo, Y. F., Yang, S. C., Lu, C. M., & Chen, A.-S. (2020). Influence of Students' Learning Style, Sense of Presence, and Cognitive Load on Learning Outcomes in an Immersive Virtual Reality Learning Environment. *Journal of Educational Computing Research*, *58*(3), 596-615. https://doi.org/10.1177/0735633119867422
- Huckauf, A. (2005). Virtual and real visual depth. *Proceedings of the 2nd Symposium on Appied Perception in Graphics and Visualization - APGV '05*, 172. https://doi.org/10.1145/1080402.1080456
- Interrante, V., Ries, B., & Anderson, L. (2006). Distance Perception in Immersive Virtual Environments, Revisited. *IEEE Virtual Reality Conference (VR 2006)*, 3-10. https://doi.org/10.1109/VR.2006.52
- Jacko, J. A., & Ward, K. G. (1996). Toward establishing a link between psychomotor task complexity and human information processing. *Computers & Industrial Engineering*, *31*(1), 533-536. https://doi.org/10.1016/0360-8352(96)00192-1
- Jaquess, K. J., Gentili, R. J., Lo, L.-C., Oh, H., Zhang, J., Rietschel, J. C., Miller, M. W.,

Tan, Y. Y., & Hatfield, B. D. (2017). Empirical evidence for the relationship between cognitive workload and attentional reserve. *International Journal of Psychophysiology*, *121*, 46-55. https://doi.org/10.1016/j.ijpsycho.2017.09.007

- Jauregui, D. A. G., Argelaguet, F., & Lecuyer, A. (2012). Design and evaluation of 3D cursors and motion parallax for the exploration of desktop virtual environments. *2012 IEEE Symposium on 3D User Interfaces (3DUI)*, 69-76. https://doi.org/10.1109/3DUI.2012.6184186
- Johal, A. S., & Battagel, J. M. (1997). Dental crowding : A comparison of three methods of assessment. *The European Journal of Orthodontics*, *19*(5), 543-551. https://doi.org/10.1093/ejo/19.5.543
- Khot, Z., Quinlan, K., Norman, G. R., & Wainman, B. (2013). The relative effectiveness of computer-based and traditional resources for education in anatomy : Virtual Reality Versus Reality in Anatomy. *Anatomical Sciences Education*, *6*(4), 211-215. https://doi.org/10.1002/ase.1355
- Klein, E., Swan, J. E., Schmidt, G. S., Livingston, M. A., & Staadt, O. G. (2009). Measurement Protocols for Medium-Field Distance Perception in Large-Screen Immersive Displays. *2009 IEEE Virtual Reality Conference*, 107-113. https://doi.org/10.1109/VR.2009.4811007
- Kockro, R. A., Killeen, T., Ayyad, A., Glaser, M., Stadie, A., Reisch, R., Giese, A., & Schwandt, E. (2016). Aneurysm Surgery with Preoperative Three-Dimensional Planning in a Virtual Reality Environment : Technique and Outcome Analysis. *World Neurosurgery*, *96*, 489-499. https://doi.org/10.1016/j.wneu.2016.08.124
- Kockro, R., Amaxopoulou, C., Killeen, T., Wagner, w, Reisch, R., Schwandt, E., Gutenberg, E., Giese, E., Stofft, E., & Stadie, A. (2015). Stereoscopic neuroanatomy lectures using a three-dimensional virtual reality environment. *Annals of Anatomy = Anatomischer Anzeiger : Official Organ of the Anatomische Gesellschaft*, *201*. https://doi.org/10.1016/j.aanat.2015.05.006
- Koumaditis, K., Chinello, F., Mitkidis, P., & Karg, S. (2020). Effectiveness of Virtual Versus Physical Training : The Case of Assembly Tasks, Trainer's Verbal Assistance, and Task Complexity. *IEEE Computer Graphics and Applications*, *40*(5), 41-56. https://doi.org/10.1109/MCG.2020.3006330
- Kravitz, N., Groth, C., Jones, P., Graham, J., & Redmond, W. (2014). Intraoral digital scanners. *Journal of clinical orthodontics : JCO*, *48*, 337-347.
- Kumar, A. A., Phillip, A., Kumar, S., Rawat, A., Priya, S., & Kumaran, V. (2015). Digital model as an alternative to plaster model in assessment of space analysis. *Journal of Pharmacy And Bioallied Sciences*, *7*(6), 465. https://doi.org/10.4103/0975-7406.163506
- Kunz, B. R., Wouters, L., Smith, D., Thompson, W. B., & Creem-Regehr, S. H. (2009). Revisiting the effect of quality of graphics on distance judgments in virtual environments : A comparison of verbal reports and blind walking. *Attention, Perception, &*

Psychophysics, *71*(6), 1284-1293. https://doi.org/10.3758/APP.71.6.1284

- Lamadé, W., Glombitza, G., Fischer, L., Chiu, P., Cárdenas, C., Thorn, M., Meinzer, H., Grenacher, L., Bauer, H., Lehnert, T., & Herfarth, C. (2000). The impact of 3-dimensional reconstructions on operation planning in liver surgery. *Archives of Surgery (Chicago, Ill. : 1960)*, *135*(11). https://doi.org/10.1001/archsurg.135.11.1256
- Lanier, J. A Vintage Virtual Reality Interview, 1988. In *Interview by Adam Heilbrun of Jaron Lanier, Published*; 1988; Available online: https://philpapers.org/rec/LANAVV (accessed on 1 September 2021)
- Li, J., & Kim, J.-E. (2021). The Effect of Task Complexity on Time Estimation in the Virtual Reality Environment : An EEG Study. *Applied Sciences*, *11*(20), 9779. https://doi.org/10.3390/app11209779
- Li, L., Yu, F., Shi, D., Shi, J., Tian, Z., Yang, J., Wang, X., & Zhang, P. (2017). Application of virtual reality technology in clinical medicine. *American Journal of Translational Research*, *9*, 3867-3880.
- Makaremi, M. (2018). Les enjeux d'une nouvelle interface. *Revue d'Orthopédie Dento-Faciale*, *52*(4), 319-324. https://doi.org/10.1051/odf/2018026
- Makaremi, M., & N'Kaoua, B. (2021). Estimation of Distances in 3D by Orthodontists Using Digital Models. *Applied Sciences*, *11*(18), 8285. https://doi.org/10.3390/app11188285
- Mangano, F., Bakaj, R., Frezzato, I., Frezzato, A., Montini, S., & Mangano, C. (2017). Morse Taper Connection Implants Placed in Grafted Sinuses in 65 Patients : A Retrospective Clinical Study with 10 Years of Follow-Up. *International Journal of Dentistry*, *2017*, 1-9. https://doi.org/10.1155/2017/4573037
- Marcel, T. J. (2001). Three-dimensional on-screen virtual models. *American Journal of Orthodontics and Dentofacial Orthopedics*, *119*(6), 666-668. https://doi.org/10.1067/mod.2001.116502
- Martin, C. B., Chalmers, E. V., McIntyre, G. T., Cochrane, H., & Mossey, P. A. (2015). Orthodontic scanners : What's available? *Journal of Orthodontics*, *42*(2), 136-143. https://doi.org/10.1179/1465313315Y.0000000001
- Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments : A ten-year review of empirical research (1999–2009). *Computers & Education*, *56*(3), 769-780. https://doi.org/10.1016/j.compedu.2010.10.020
- Murgia, A., & Sharkey, P. M. (2009). Estimation of Distances in Virtual Environments Using Size Constancy. *International Journal of Virtual Reality*, *8*(1), 67-74. https://doi.org/10.20870/IJVR.2009.8.1.2714
- Naish, H., Dunbar, C., Crouch-Baker, J., Shah, K., Wallis, C., Atack, N. E., Sherriff, M., Sandy, J. R., & Ireland, A. J. (2016). Does a true knowledge of dental crowding affect orthodontic treatment decisions? *The European Journal of Orthodontics*, *38*(1), 66-70. https://doi.org/10.1093/ejo/cjv011

- Nowak, K. L., & Biocca, F. (2003). The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *12*(5), 481-494. https://doi.org/10.1162/105474603322761289
- Ouramdane, N., Otmane, S., & Mallem, M. (2009). Interaction 3D en réalité virtuelle. Etat de l'art. *Techniques et sciences informatiques*, *28*(8), 1017-1049. https://doi.org/10.3166/tsi.28.1017-1049
- Peillard, E., Thebaud, T., Normand, J.-M., Argelaguet, F., Moreau, G., & Lecuyer, A. (2019). Virtual Objects Look Farther on the Sides : The Anisotropy of Distance Perception in Virtual Reality. *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 227-236. https://doi.org/10.1109/VR.2019.8797826
- Pensieri, C., & Pennacchini, M. (2016). Virtual Reality in Medicine. In Y. Sivan (Éd.), *Handbook on 3D3C Platforms : Applications and Tools for Three Dimensional Systems for Community, Creation and Commerce* (p. 353-401). Springer International Publishing. https://doi.org/10.1007/978-3-319-22041-3_14
- Preece, D., Williams, S. B., Lam, R., & Weller, R. (2013). "Let's Get Physical" : Advantages of a physical model over 3D computer models and textbooks in learning imaging anatomy: Let's Get Physical. *Anatomical Sciences Education*, *6*(4), 216-224. https://doi.org/10.1002/ase.1345
- Radeke, J., von der Wense, C., & Lapatki, B. G. (2014). Comparison of orthodontic measurements on dental plaster casts and 3D scans. *Journal of Orofacial Orthopedics / Fortschritte Der Kieferorthopädie*, *75*(4), 264-274. https://doi.org/10.1007/s00056-014-0217-9
- Renner, R. S., Velichkovsky, B. M., & Helmert, J. R. (2013). The perception of egocentric distances in virtual environments—A review. *ACM Computing Surveys*, *46*(2), 1-40. https://doi.org/10.1145/2543581.2543590
- Reuschl, R. P., Heuer, W., Stiesch, M., Wenzel, D., & Dittmer, M. P. (2016). Reliability and validity of measurements on digital study models and plaster models. *European Journal of Orthodontics*, *38*(1), 22-26. https://doi.org/10.1093/ejo/cjv001
- Rossini, G., Parrini, S., Castroflorio, T., Deregibus, A., & Debernardi, C. L. (2016). Diagnostic accuracy and measurement sensitivity of digital models for orthodontic purposes : A systematic review. *American Journal of Orthodontics and Dentofacial Orthopedics*, *149*(2), 161-170. https://doi.org/10.1016/j.ajodo.2015.06.029
- Sakellariou, S., Ward, B. M., Charissis, V., Chanock, D., & Anderson, P. (2009). Design and Implementation of Augmented Reality Environment for Complex Anatomy Training : Inguinal Canal Case Study. In R. Shumaker (Éd.), *Virtual and Mixed Reality* (Vol. 5622, p. 605-614). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-02771-0_67
- Soler, L., Nicolau, S., Pessaux, P., Mutter, D., & Marescaux, J. (2014). Real-time 3D image reconstruction guidance in liver resection surgery. *Hepatobiliary Surgery and*

Nutrition, *3*(2). https://doi.org/10.3978/j.issn.2304-3881.2014.02.03

- Sousa, M. V. S., Vasconcelos, E. C., Janson, G., Garib, D., & Pinzan, A. (2012). Accuracy and reproducibility of 3-dimensional digital model measurements. *American Journal of Orthodontics and Dentofacial Orthopedics*, *142*(2), 269-273. https://doi.org/10.1016/j.ajodo.2011.12.028
- Stefanucci, J. K., Creem-Regehr, S. H., Thompson, W. B., Lessard, D. A., & Geuss, M. N. (2015). Evaluating the accuracy of size perception on screen-based displays : Displayed objects appear smaller than real objects. *Journal of Experimental Psychology: Applied*, *21*(3), 215-223. https://doi.org/10.1037/xap0000051
- Stevens, J., Mondesire, S., Maraj, C., Badillo-Urquiola, K., & Maxwell, D. (2016). *Workload Analysis of Virtual World Simulation for Military Training*.
- Swenson, H. A. (1932). The Relative Influence of Accommodation and Convergence in the Judgment of Distance. *The Journal of General Psychology*, *7*(2), 360-380. https://doi.org/10.1080/00221309.1932.9918473
- Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., & Beall, A. C. (2004). Does the Quality of the Computer Graphics Matter when Judging Distances in Visually Immersive Environments? *Presence: Teleoperators and Virtual Environments*, *13*(5), 560-571. https://doi.org/10.1162/1054746042545292
- Tourancheau, S., Sjöström, M., Olsson, R., Persson, A., Ericson, T., Rudling, J., & Norén, B. (2012). *Subjective evaluation of user experience in interactive 3D visualization in a medical context* (C. K. Abbey & C. R. Mello-Thoms, Éds.; p. 831814). https://doi.org/10.1117/12.910828
- Wainman, B., Aggarwal, A., Birk, S. K., Gill, J. S., Hass, K. S., & Fenesi, B. (2020). Virtual Dissection : An Interactive Anatomy Learning Tool. *Anatomical Sciences Education*, ase.2035. https://doi.org/10.1002/ase.2035
- Wainman, B., Wolak, L., Pukas, G., Zheng, E., & Norman, G. R. (2018). The superiority of three-dimensional physical models to two-dimensional computer presentations in anatomy learning. *Medical Education*, *52*(11), 1138-1146. https://doi.org/10.1111/medu.13683
- Waller, D. (1999). Factors Affecting the Perception of Interobject Distances in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *8*(6), 657-670. https://doi.org/10.1162/105474699566549
- Wallis, C., McNamara, C., Cunningham, S. J., Sherriff, M., Sandy, J. R., & Ireland, A. J. (2014). How good are we at estimating crowding and how does it affect our treatment decisions? *The European Journal of Orthodontics*, *36*(4), 465-470. https://doi.org/10.1093/ejo/cjt080
- Wartenberg, C., & Wiborg, P. (2003). Precision of Exocentric Distance Judgments in Desktop and Cube Presentation. *Presence: Teleoperators and Virtual Environments*, *12*(2), 196-206. https://doi.org/10.1162/105474603321640941

- Wood, R. E., Mento, A. J., & Locke, E. A. (1987). Task complexity as a moderator of goal effects : A meta-analysis. *Journal of Applied Psychology*, *72*(3), 416-425. https://doi.org/10.1037/0021-9010.72.3.416
- Yammine, K., & Violato, C. (2016). The effectiveness of physical models in teaching anatomy : A meta-analysis of comparative studies. *Advances in Health Sciences Education*, *21*(4), 883-895. https://doi.org/10.1007/s10459-015-9644-7

Annexes-3 :

- Makaremi, M., Lacaule, C., & Mohammad-Djafari, A. (2019). Deep learning and artificial intelligence for the determination of the cervical vertebra maturation degree from lateral radiography. *Entropy*, *21*(12), 1222.

- Entropy MDPI impact factor : 2.524

entropy

MDPI

Article

Deep Learning and Artificial Intelligence for the Determination of the Cervical Vertebra Maturation Degree from Lateral Radiography

Masrour Makaremi [1], Camille Lacaule [1] and Ali Mohammad-Djafari [2,*]

[1] Department of Orthodontics, University of Bordeaux, 33000 Bordeaux, France;masrour@makaremi.fr (M.M.); orthodontie@makaremi.fr (C.L.)

[2] International Science Consulting and Training (ISCT), 91440 Bures-sur-Yvette, France

* Correspondence: djafari@free.fr; Tel.: +33-6-2295-4233

Received: 29 September 2019; Accepted: 9 December 2019; Published: 14 December 2019

Abstract: Deep Learning (DL) and Artificial Intelligence (AI) tools have shown great success in different areas of medical diagnostics. In this paper, we show another success in orthodontics. In orthodontics, the right treatment timing of many actions and operations is crucial because many environmental and genetic conditions may modify jaw growth. The stage of growth is related to the Cervical Vertebra Maturation (CVM) degree. Thus, determining the CVM to determine the suitable timing of the treatment is important. In orthodontics, lateral X-ray radiography is used to determine it. Many classical methods need knowledge and time to look and identify some features. Nowadays, ML and AI tools are used for many medical and biological diagnostic imaging. This paper reports on the development of a Deep Learning (DL) Convolutional Neural Network (CNN) method to determine (directly from images) the degree of maturation of CVM classified in six degrees. The results show the performances of the proposed method in different contexts with different number of images for training, evaluation and testing and different pre-processing of these images. The proposed model and method are validated by cross validation. The implemented software is almost ready for use by orthodontists.

Keywords: Deep Learning (DL); Artificial Intelligence (AI); Convolutional Neural Network (CNN); classification; orthodontics; cervical vertebra maturation; machine learning

1. Introduction and the Organisation of The Paper

Artificial Intelligence (AI), Machine Learning (ML) and Artificial Neural Networks (ANN) and many of their sub-fields such as Convolutional Neural Networks (CNN) and Deep Learning (DL) have became common tools in different areas of our life, and in particular, in biological and medical diagnostics. In this paper we show one more success in orthodontics.

In computer vision and particularly in image classification, deep convolutional neural networks (DCNNs) have emerged as the master algorithm and provided state-of-the-art performance in most tasks. The main power of a CNN is attributed to its deep architecture, allowing extracting a set of discriminating features at multiple levels of abstraction.

This paper is organized as follows: First, the importance of the work and its interest for the orthodontics community is mentioned. Then, the classical manual classification methods of radiographic images are described. The difficulties of the labeling task by the specialists are mentioned and the need for automatisation and the utility of ML tools are shown. Then, the proposed method is described which is mainly an appropriate CNN-based classification method which has been developed for these radiographic images.

2. Importance of the Work and Its Interest for the Orthodontics Community

Specialists in orthodontics are responsible for the treatment of dentofacial dysmorphisms, from different functional, genetical and morphological aetiologies. As a child or teenager is still growing, orthodontic treatment consists in a combination of orthodontics (about tooth position) and dentofacial orthopedics (about the guidance and stimulation of facial, maxilla and mandible growth in the three dimensions).

Many environmental and genetic conditions may induce upper or lower jaws' lack of growth. Classically, to handle a treatment properly, every etiological condition that can be modified or corrected must be identified (diagnosis), normalized (treatment), and stabilized (retention). Specialists have to carefully examine and precisely analyze all the medical, functional, clinical and radiographic data in order to identify normal versus pathological conditions about tooth position, form or size, about lip, chin, cheeks, tongue and breathing functions, and about facial and jaws position and growing patterns. Adolescent orthodontic treatment also depends on proper management of jaws and facial growth to allow a balanced jaws position, maximize the airway and improve the facial appearance [1]. Treatment planning in orthodontics depends on a systematic diagnosis and prognosis.

Contemporary theories about craniofacial growth admit that the phenotype of the craniofacial complex is a result of a combination of genetic, epigenetic and environmental factors. The skeletal tissue of maxillomandibular complex grows due to sutures and osteogenic cartilages proliferation depending on genetic, intrinsic and extrinsic environment. So, facial growth can also be modified in amount and direction by extrinsic factors, including orthopedic and functional treatment. Thus, quantify facial and, in particular mandibular growth remaining, influences diagnosis, prognosis, treatment goals and planning. Indeed, apart choosing the good appliance needed to change the rate and direction of jaw growth, the right treatment timing is crucial. If high growth rate is about to occur, orthopedic treatment may permit to correct jaws, otherwise surgical correction of the jaw shift will be considered. The success of a dentofacial orthopedic treatment is linked to the determination of the interventional frame (periods of accelerated or intense growth) to maximize the chances to reach skeletal goals, with adapted methods and devices and in an optimized duration.

The most common dentofacial dysmorphism is the skeletal class II, corresponding to a short mandible. Study of normal mandibular growth and remodeling has shown different ways of bone formation that can be stimulated by functional and orthopedics treatments, in particular, condylar growth responsible for 80% of the mandible growth.

Numerous radiographic investigations have established that condylar/mandibular growth follows a similar growth curve as statural growth [2]. This growth pattern is characterized by variations of growth rate in four stages: first, a decrease of growth velocity from birth to 6 years old, then minor mid-growth spurt around 6–8 years, followed by a pre-pubertal plateau with decelerated growth rate, and finally, the facial growth curve describes a peak of growth velocity corresponding to the pubertal growth spurt, which coincides, precedes or follows (from 6 to 12 months) the statural growth peak (controversial) [3]. This spurt occurs approximately two years earlier in some girls than in some boys [4].

To estimate mandibular growth potential, the patient must be localized on is growth curve, and many biologic indicators have been proposed; increase in body height, menarche, breast and voice changes, dental development and eruption, middle phalanx maturation of the third finger, maturation of the wrist and cervical vertebral maturation [5–8].

3. The Classical Radiographic Manual Methods

3.1. Hand-Wrist Radiograph Method HWM:

The comparison method describes in the Atlas of Greulich et Pyle in 1959 or the Fishman's method in 1982, permits to identify specific ossification stages occurring before, during, or after mandibular growth peak, on left hand and wrist radiographs [9,10]. wrist radiographs have been used

as a gold standard in the assessment of skeletal maturation for many decades, but have presented several issues such as additional x-ray exposure, time and experience required (even if a digital software is now available [11]), and a sexual dimorphism and ethnic polymorphism in morphological modifications [12,13]

3.2. Vertebrae Maturation CVM:

The first to propose to predict skeletal age and growth potential by cervical vertebrae maturation (CVM) method is LAMPARSKI in 1972. Cervical vertebrae are available on the lateral cephalometric radiographs, prescribed routinely by orthodontists for each patient diagnosis and treatment planning [14]. He used measurements of mandibular length on several annual lateral cephalograms to describe individual mandibular growth curve, and correlated it with morphological description of vertebrae morphology at each stage. This method was modified several times first by Hassel and Farman (1995) [15], then twice by Baccetti et al. (2002 and 2005) for a more accurate assessment of cervical maturation, by six stages identified by morphological changes in the C2,C3,C4 vertebral bodies on a single lateral cephalogram, independently of patient gender [16].

This last version is the most used to detect the mandibular growth spurt, as it shows the best results in clinical applicability [17].

As most bones of the human body, vertebrae growth and maturation change from birth to full maturity. Cervical vertebrae are the first seven pieces of the spinal column. Vertebral growth in the cartilaginous layer of the superior and inferior surfaces of each vertebrae involves changes in size of vertebral bodies and shape of upper and lower borders of C2,C3,C4 vertebrae. These changes have been described into six stages, correlating with morphological modifications of the vertebral shapes and estimated time lapse from the mandibular growth peak. Both visual and cephalometric appraisals of morphological changes have been proposed. See Figure 1.

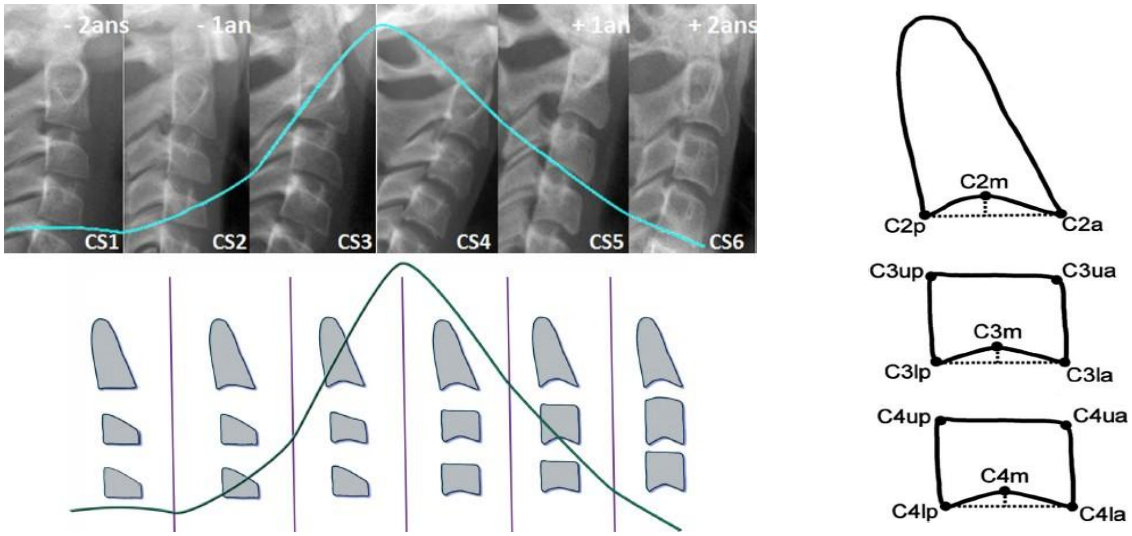


Figure 1. (**Left**) CVM radiological and morphological stages superposed with Björk growth curve [16];(**Right**) Cephalometric landmarks for CVM stages determination [1].

Visual analysis [1]:

- Cervical stage 1 (CS1) = 2 years before mandibular growth peak:
  Lower borders of C2 to C4 vertebrae are flat. C3 and C4 superior borders are tapered from posterior to anterior.
- Cervical stage 2 (CS2) = 1 year before mandibular growth peak:
  Lower border of C2 presents a concavity. Bodies of C3 and C4 are the same.

- Cervical stage 3 (CS3) = during the year of the mandibular growth peak:
  Lower borders of C2 and C3 present concavities. Vertebrae are growing so C3 and C4 may be either trapezoid or rectangular shape, as superior borders are less and less tapered.
- Cervical stage 4 (CS4) = 1 or 2 years after mandibular growth peak:
  Lower borders of C2, C3 and C4 present concavities. Both C3 and C4 bodies are rectangular with horizontal superior borders longer than higher.
- Cervical stage 5 (CS5) = 1 year after the end of mandibular growth peak:
  Still concavities of lower borders of C2, C3 and C4. At least one of C3 or C4 bodies are squared and spaces between bodies are reduced.
- Cervical stage 6 (CS6) = 2 years after the end of mandibular growth peak:
  The concavities of lower borders of C2 to C4 have deepened. C3 and C4 bodies are both square or rectangular vertical in shape (bodies higher than wide)

Cephalometric Appraisals:

Using the landmarks illustrated on Figure 1(right), cephalometric analysis consists in the measurement of:

- The concavity depth of the lower vertebral border (estimated by the distance of the middle point(Cm) from the line connecting posterior to anterior points (Clp-Cla))
- The tapering of upper border of vertebral C3 and C4 bodies (estimated by the ratio between posterior and anterior bodies heights (Cup-Clp)/(Cua-Cla))
- The lengthening of vertebral bodies (estimated by the ratio between the bases length and anterior bodies borders height (Clp-Cla)/Cua-Cla)

Many researchers found this method as valid and reliable as hand and wrist Xray [14]. The cervical vertebrae maturation stages have been demonstrated as a clinically useful maturation indicators for evaluation of pubertal growth height and mandibular velocities [18–20], by correlation between chronological age and cervical vertebrae maturation, between hand–wrist and cervical-vertebrae maturation [16,21–23].

Some studies underlined the need for association with other clinical assessments [24] in clinical practice, and a good reliability in differentiating pre and post mandibular growth spurt periods [25].

3.3. The Difficulties of the Labeling Task

Specific training is provided to assess CVM stages reliably, and repeatably at a satisfactory level [26,27]. Gabriel et al. minimized the risk of bias (radiographs without tracings, standardized training to private practice orthodontists...) and observed a moderate intra and inter-observer agreement (30%–62% of cases). These results confirm the expertise required to properly determine the CVM stage, and may be explained by the use of a qualitative method of assessment, and the lack in detecting exceptional cases (individual variations in size and morphology, outside the norms defined by the method). Moreover, for orthodontists, the cervical vertebrae area on the lateral cephalograms is outside their expertise of "visual fields". They have poor general knowledge and experiences about vertebrae observation, as they focus on maxillomandibular bones and teeth at a first glance. This would have been a difficulty in the labeling task of our radiographs.

In our work, all lateral radiographs have been labeled by a radiological technician, specialized in cephalometric tracing and over trained in CVM stages agreement (3 years full time), using a standardized morphological and cephalometrical protocol. Intra observer reproducibility must be estimated in further study.

3.4. The Need for Automatisation and the Help Which It Brings

Estimation of CVM stage represents only one single element influencing the patient orthodontictreatment. The practitioner must master the entire clinical, functional, bio-mechanical and cephalometric

data analysis in order to define proper diagnosis and treatment goals and planning. Even in being specialists, orthodontists require a very broad range of skills and a great deal of time for each patient to complete diagnosis. Considering that reproducibility of classifying CVM stages is superior at 98% by trained examiners [1], automatisation by expert eyes will provide time saving, efficiency, accuracy, repeatability in treatment planning and patient care.

Few studies have presented software programs for cephalometric determination of C2, C3 and C4 vertebrae proportions according reference points marked manually on the image, and automatically calculates the skeletal maturation stage. This computer-assisted analysis still depends on operator experience [28]. Padalino et al. ran a study comparing manual analysis of CVM stages and the analysis performed by a dedicated software. It has shown a concordance of 94% between the two methods but hand-tracing analysis was quicker of 28 s on average [29].

Deep Learning (DL) Convolutional Neural Networks (CNN) have already been used to diagnose metabolic disorders in pediatric endocrinology, in order to assess skeletal bone age on left hand-wrist radiographs. Deep learning approach proposes better accuracy than conventional methods in processing the image in less than 1 s. Our study aims to develop a fully automated deep learning assessment of CVM stages on lateral cephalograms in orthodontics.

4. Proposed Method

The proposed method is mainly a supervised classification method using CNN and DL. The main contributions of this paper are:

- Appropriate needed preprocessing of the images before starting the training, evaluation and testing steps;
- A Comparison of different appropriate CNN and DL existing methods on a reduced set of images in our data base;
- Appropriate choice of a DL CNN structure, criteria, optimization algorithm and setting of all the parameters and hyper parameters for our application;
- Showing the performances of the proposed method for a different number of training, evaluating and testing images;
- Pre-clinical evaluation of the implemented method.

These contributions are presented in the following subsections.

4.1. Preprocessing of the Data

For this classification task, we had an image data base of 2000 X-ray radiographic images. Each image has a size of 2012 × 2012. These images were extracted from the patients files and are anonymized.

A selection of 600 images were studied and labeled by the experts in six classes (CVS1, . . . , CVS6). These labeled data were divided in three sets of Training, Validation and Testing. We did different division of the data: First, we had started by 300, 200 and 100, respectively, for Training, Validation and Testing. We did some preliminary tests with this small dataset. Then, we obtained more data, 600, then, 900 and, finally, about 1900. We used these images, step by step, in different experiences and used a Cross Validation (CV) technique by division and permutation of these data sets in different experiences.

Also, as these images were from the whole head, only a specific part of the image was useful for this classification, we performed different preprocessing before feeding then to the DL input. Figure 2 shows one example of the interesting part of the images to be used for this application.

In a preprocessing step, each original image was first cropped to the interesting part (Test0: size 512 × 512), then resized to (Test1: size 256 × 256, Test2: 128 × 128 and Test3: size 64 × 64) or after resizing to 256 × 256, they are Sobel filtered to enhance the contours of the image (Test 4). These sets are used for studying the importance of the preprocessing the images.

We used other combinations of preprocessing with different cropping, resizing and filtering. Figure 3 shows in summary examples of these inputs.

Figure 2. Isolated part (vertebra) of a standard radiography.

4.2. Considered Deep Learning Networks

In a preliminary study, we used different Deep Learning network structures for this classification task. Several network architectures [30–39] have been proposed with success for ordinary photographic image classification. These different architectures have been proposed to deal with great problems of over fitting, vanishing or exploding gradients in the optimization parts of learning. Each one tries to provide a more efficient use of network parameters in order to offer ultimate improvement in the model efficiency in terms of computation time and accuracy.

The success that brought the breakthrough in the ImageNet competition in 2012 is related to the efficient use of Graphical Processing Units (GPU), data augmentation, rectified linear units, new dropout regularization and deep network architecture, where convolution operations were repeated multiple times between max-pooling operations. Since then, there has been a trend to make this style of network increasingly deeper through the use of small (3×3) convolutional filters such as VGG architecture [40–48].

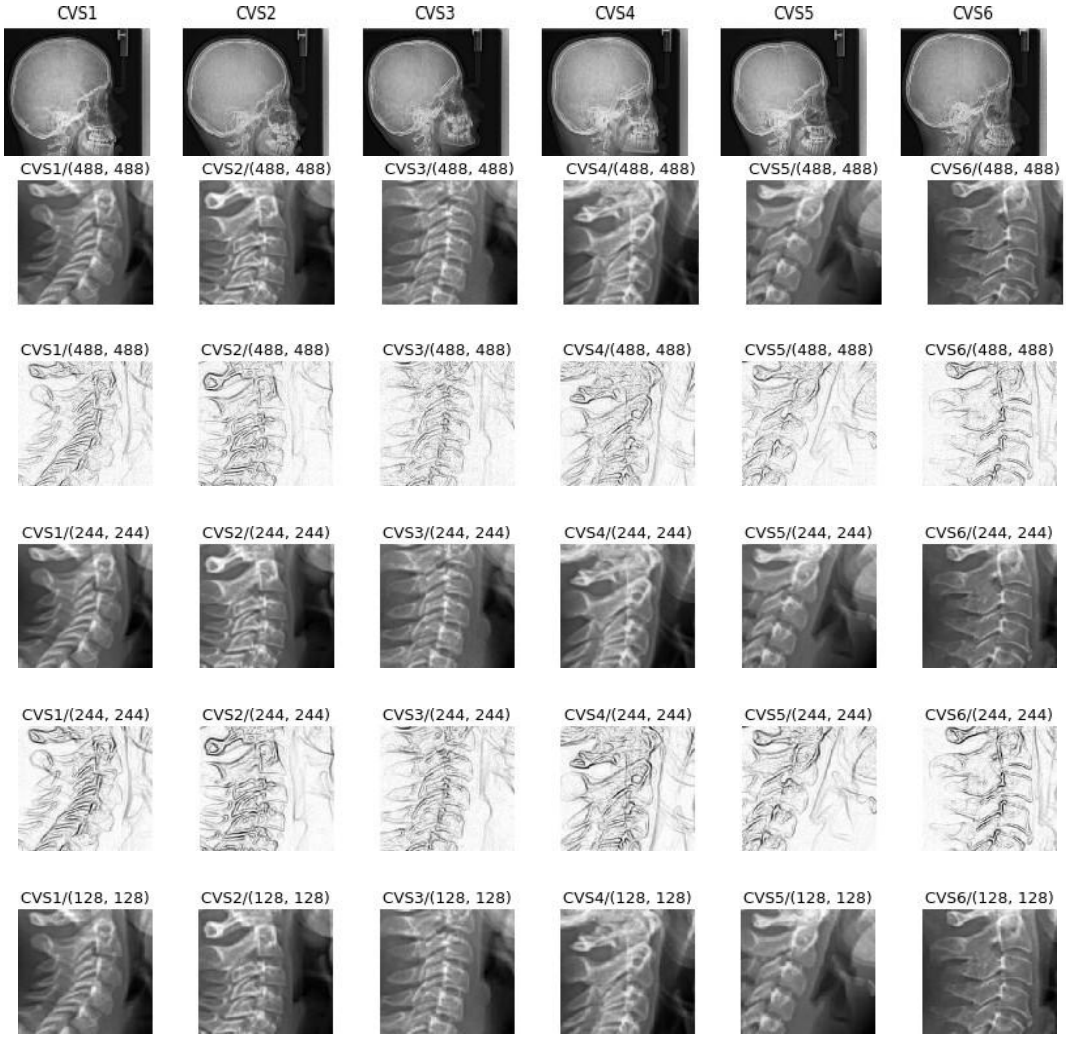


Figure 3. Originals and different preprocessing before training: (**a**) Originals (2012 × 2020), (**b**) test0: cropped images (488 × 488), (**c**) test1: cropped and sobel edge detector filter (488 × 488), (**d**) test2: cropped and resized (244 × 244), (**e**) test3: cropped, resized and sobel edge detector filter (244 × 244), (**f**) test4: cropped and resized (64 × 64).

However, very deep networks suffer from the problem of vanishing/exploding gradients during training process. To overcome this problem, He et al. [49] proposed residual networks termed ResNet by bypassing signal from one layer to the next via identity (or residual) connections. Recently, Huang et al. [50] took a step further by proposing dense networks (DenseNets) that connects each layer to every other layers in a feed-forward fashion. DenseNets offer several advantages, such as alleviating the vanishing gradient problem, encouraging feature reuse and substantially reducing the number of network parameters.

Almost all these networks have been designed and tested for ordinary color images on great data bases. Our main objective in this preliminary study was to compare these networks for our particular case of radiographic images.

We considered different classical networks:

- Resnet:

  Resnet was introduced in the paper “Deep Residual Learning for Image Recognition ” [49] There are several variants with different output zes, including Resnet18, Resnet34, Resnet50,

Resnet101, and Resnet152, all of which are available from torchvision models. As our dataset is small, we used Resnet18 that we adapted in our case for 6 classes.

- Alexnet:

  Alexnet was introduced in the paper "ImageNet Classification with Deep Convolutional Neural Networks" [51] and was the first very successful CNN on the ImageNet dataset.

- VGG:

  VGG was introduced in the paper "Very Deep Convolutional Networks for Large-Scale Image Recognition" [52]. Torchvision offers eight versions of VGG with various lengths and some that have batch normalizations layers.

- Squeezenet:

  The Squeeznet architecture is described in the paper "SqueezeNet: AlexNet-level accuracy with 50× fewer parameters and <0.5 MB model size" [53]. It uses a different output structure than the other models mentioned here. Torchvision has two versions of Squeezenet. We used version 1.0.

- Densenet:

  Densenet was introduced in the paper "Densely Connected Convolutional Networks" [50]. Torchvision has four variants of Densenet. Here we used Densenet-121 and modified the output layer, which is a linear layer with 1024 input features and 6 classes, for our case.

- Inception v3:

  Inception v3 was first described in "Rethinking the Inception Architecture for Computer Vision" [54]. This network is unique because it has two output layers when training. The second output is known as an auxiliary output and is contained in the AuxLogits part of the network. The primary output is a linear layer at the end of the network. Note that when testing, we only consider the primary output.

4.3. Comparison of a Simple Network on Different Preprocessed Images

In a first step, to decide for a preprocessing module for our images, we designed a simple CNN network and tested its performances on different set of preprocessed images.

As it can be seen on Figure 4, the structure of Deep Learning model is composed of an input convolutional layer and three or four other convolutional nets (CNN) layers and a fully connected of (32 × 32) to 6 classes. Each of the three CNNs is followed by a normalization, pooling and dropout layers with different dropout coefficients.

For this step, the models are trained with different partitions of the images in Training, Validation and Testing sets. The following figure shows 300 images which have been prepared for the training, then validated on 200 images and saved to be used for the testing step. A set of 120 images is used for testing and the average score was 80 percent.

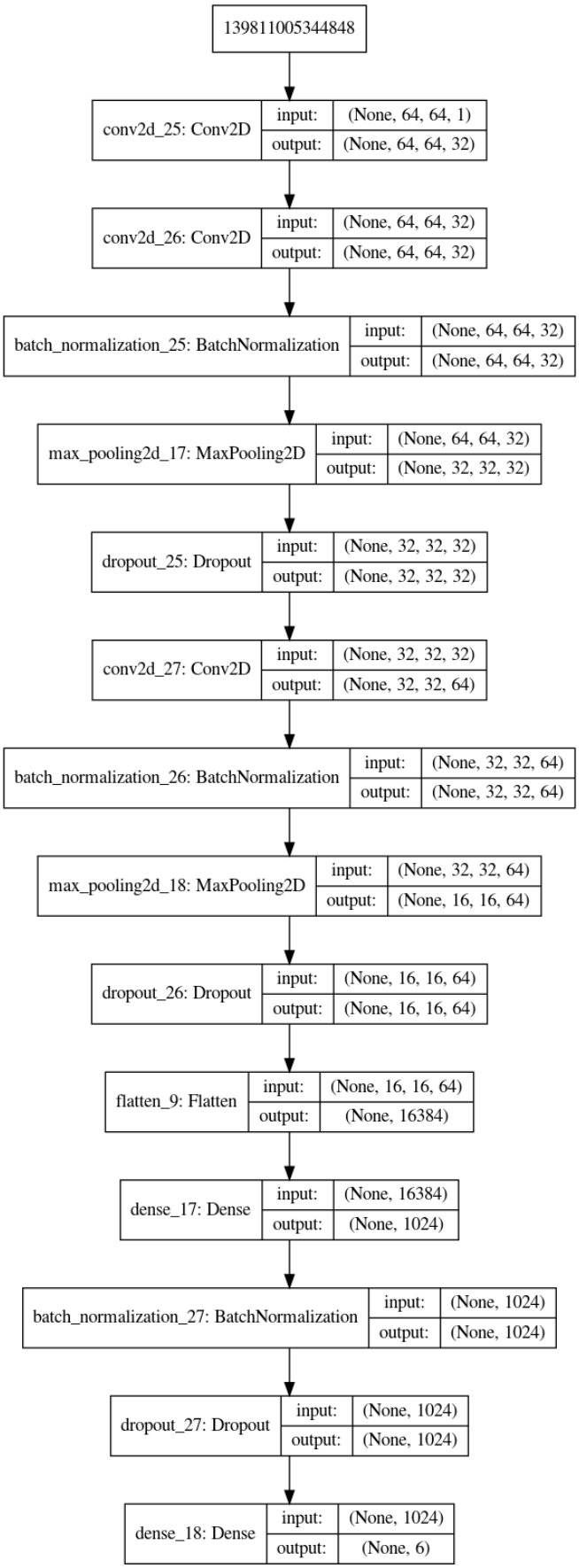


Figure 4. The structure of the proposed Deep Learning network for test with different preprocessed images.

4.4. Tools and Implementation

In this work, we used the following tools:

- SciKit-Image https://scikit-image.org/:
  All simple imag processing such as reading and writing images, Cropping, resizing and Filtering.
- SciKit-Learn https://scikit-learn.org/stable/: Data shuffling, Kmeans and Gaussian Mixtureclustering, Principal Component Analysis and performance metrics.
- Keras wth tensorflow backend https://keras.io/:
  VGG16, VGG19 and ResNet50 convolution network models with ImageNet weights.
- PyTorch https://pytorch.org/:
  GPU and CPU vision algorithms, optimization, scheduler, feature transfer, etc.

We used all these tools for development of this project. As an example, we used Keras following this process.

- For a given Training data set, Create Model (keras.layers);
- Configure Model (model.compile);
- Train model (model.fit);
- For given Evaluating and Testing data, evaluate the trained model (loss = model.evaluate);
- Get prediction (pred = model.predict).

4.5. Prediction Results with Different Networks

With the implemented DL structure, we used 300 images for the training step, 200 images for the validation step and finally 150 images for the testing step. We had to fix a great number of parameters such as dropout rates, optimization algorithms, regularization parameters, etc.

Figure 5 shows the evolution of the Loss function and the accuracy as a function of the epoch numbers for one of these different tests.

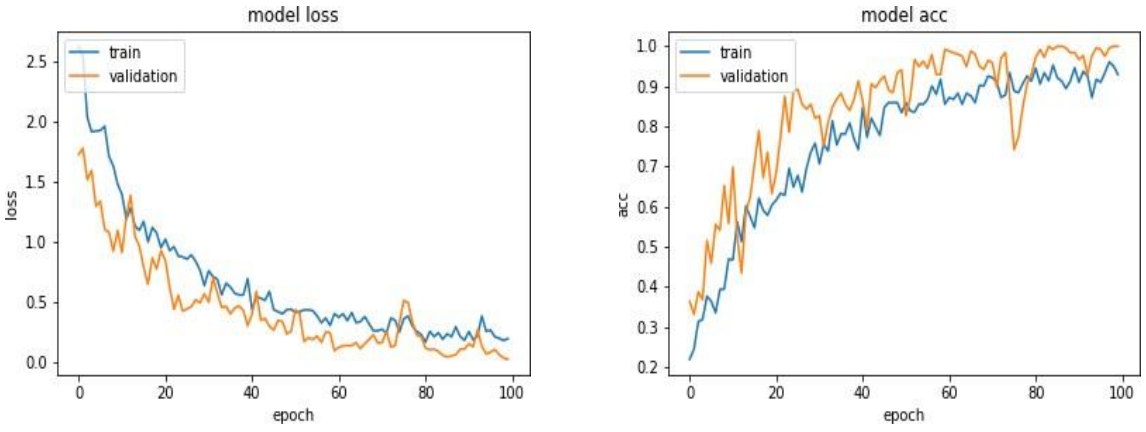


Figure 5. Evolution of the Loss function and the accuracy as a function of the epoch numbers.

Table 1 shows the prediction results obtained with different preprocessing of the data, both during the training and the testing

Table 1. Comparison of the classification accuracy for different preprocessing.

Class	CVS1	CVS2	CVS3	CVS4	CVS5	CVS6
Test0	0.910	0.820	0.654	0.617	0.863	0.939
Test1	0.890	0.734	0.661	0.580	0.849	0.924
Test2	0.890	0.734	0.661	0.580	0.849	0.924
Test3	0.932	0.824	0.644	0.561	0.839	0.849
Test4	0.880	0.890	0.568	0.602	0.861	0.932

5. Proposed Method, Structure and Metrics

As we mentioned, ResNet, DenseNet, InceptionResNet and Xception are among the most efficient models that have achieved state-of-the-art accuracy on large image datasets such as CIFAR, ImageNet and all other classical photography images.

However, in our case, we have a small database. Hopefully, these images are not too different from each other and the variability (inter and intra classes) are not as large as in the case of photography images.

So, we decided to make our specific DL network, specifically adapted for our application and train it directly with our pre-processed images.

What we learned in previous sections can be summarized as follows:

- Almost all the pre-trained networks with feature transfer did not really work better. This is due to the fact that these networks are trained with great databases with photographic images which are more diverse than the radiographic images we have.
- A cropping of images of the size 512 × 512 to the specific informative part of cervical vertebra and possibly resizing them to 256 × 256 for reducing the computational costs are enough. However, to improve more, we tried different pre-processing of the images: mean, median and entropic filter to the images.
- The proposed method is the same as in Figure 3 with slightly different parameters which are optimized for our application.

5.1. Entropic Filtering

In classical information theory, information entropy is the log-base-2 of the number of possible outcomes for a message. In Bayesian probability theory, the entropy of a probability distribution $p = \{p_1, \cdots, p_N\}$ is defined as

$$H(p) = -\sum_{n=1}^{N} p_n \ln p_n \tag{1}$$

From this definition, we see that, the highest entropy is obtained for a uniform distribution. We can use this property to measure the degree of uniformity of the distribution of the pixel values in an image. A local entropy of the distribution of a sliding window on an image can then be a measure of local uniformity of the image. Thus, for an image, local entropy is related to the complexity contained in a given neighborhood, typically defined by a structuring element. We use this property to pre-process the images as we are interested more on the contours than the gray levels of the images. Based on this, we can define an entropy filter which can detect subtle variations in the local gray level distribution.

In image processing, there are many other local filterings, such as mean or median. To show the importance of our choice for the local entropy compared to the classical mean or median filtering, we show here the results of these filters on a noisy image obtained from a simple mask image with two constant value regions on which we added uniform noise. Figure 6 shows the results.

As it can be seen from Figure 6, the mean and the median filters are less effective than the entropy filter to catch the structural information of the image.

Figure 7 show these filters on a sample of images on which we are interested in.

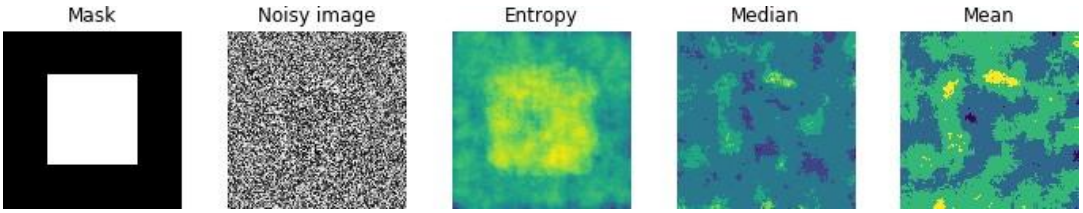


Figure 6. Comparison between three local filters: mean, median and entropy.

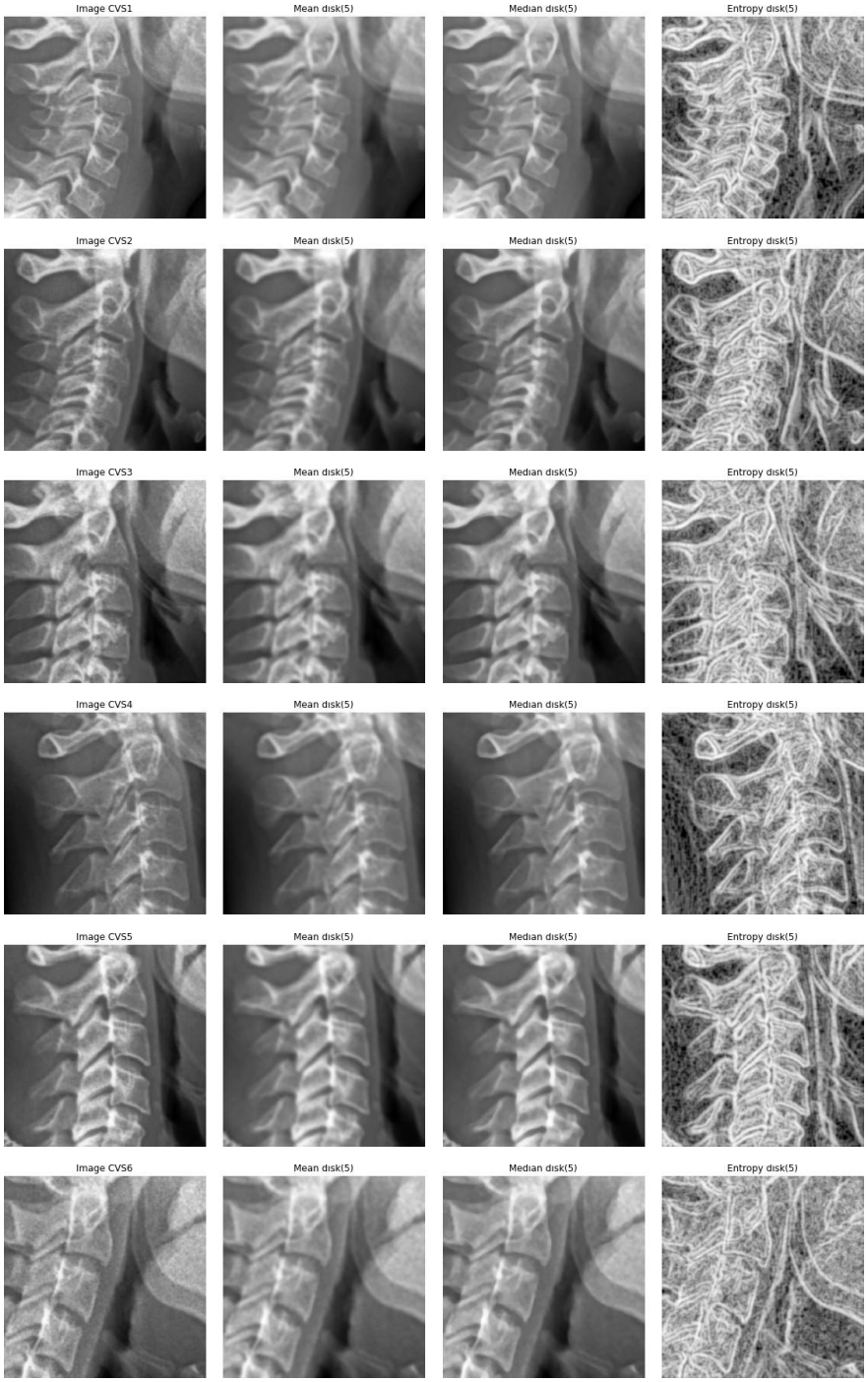


Figure 7. Comparison between three local filters: mean, median and entropy. Columns: from left to right: original, mean, median, entropy. Rows: from top to down CVS1, CVS2 to CVS6 cases.

5.2. Optimisation Criteria

As we are interested in the general classification problem with six classes, we use the categorial cross entropy as the optimisation criteria for learning.

Categorical cross-entropy is a loss function that is used for single label categorization. This is when only one category is applicable for each data point.

Categorical cross-entropy will compare the distribution of the predictions (the activation in the output layer, one for each class) with the true distribution, where the probability of the true class is set to 1 and 0 for the other classes. If we note by p_{ij} the probability that the sample i be in class j and by $\hat{p}_{ij}$ the probability of the predicted one, then cross-entropy is given by:

$$H(p_{ij}, \hat{p}_{ij}) = -\sum_i \sum_j p_{ij} \ln \frac{\hat{p}_{ij}}{p_{ij}} \quad (2)$$

When the true class is represented as a one-hot encoded vector, i.e., when the sample x_i has the label y_{ij} and the predicted label is noted $\hat{y}_{ij}$, the loss function associated to the cross-entropy becomes:

$$L(y_{ij}, \hat{y}_{ij}) = -\sum_i \sum_j y_{ij} \ln \hat{y}_{ij} \quad (3)$$

This is due to the fact that, y_{ij} can take only the values 0 and 1. Thus, the closer the model's outputs are to the training labels, the lower is this loss.

5.3. Optimisation Algorithms

We mainly used two standard algorithms: Stochastic Gradient Descent (SGD) and Adaptive Moment Estimation (Adam). We tried different variants of SGD (Batch and mini-batch without and with shuffling) with different learning rate and Adam with different learning rate and momentum. There is not really any easy way to fix these hyper parameters. We just tried different values to get the fastest convergence rate and a good generalization error. For a good review of these optimization algorithms, see [55] or [56].

5.4. Metrics Used to Measure the Quality of the Classification

Classical validation metrics used for two class classification (detection problem), such as True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN), Score or accuracy (S = TP/(TP + TN)), Precision (P = TP/(TP + FP)) and Recall (R = TP/(TP + FN)) are not enough for our multi class problem. The most complete metric is the confusion matrix.

5.5. Results

We proceeded many optimization of the parameters. For example, for the optimisations of the Learning Rate (LR) of the algorithm, we proceeded these steps:

- set a lower and an upper bounds for LR. For example: $(1 \times 10^{-1}; 1 \times 10^{1})$;
- start the training of the network;
- increase the LR in an exponential way after each batch update;
- for each LR, record the loss at each batch update;
- train for a few epoches and plot the loss as a function of the LR;
- examine the plot and identify the optimal LR;
- update the LR and train the network for the full set of training data.
- apply the trained network to the evaluation and test data and look for a possibly over fitting problem;
- If all are right, then between those images in the evaluation and test set data, move those which are classified correctly with a very high probability to the training set and update the evaluation and test sets with new images until the end of the data base.

In the following, we show some partial results obtained with different sizes of the data base. In these experiments, we used 300 images as the validation and 300 as the test. In the following, we changed the number of images in the training set.

5.6. Results with 360 Images

Figure 8 shows the loss and the accuracy functions during the learning and evaluation process when we used 360 images for training, 300 for evaluation and 300 for testing.

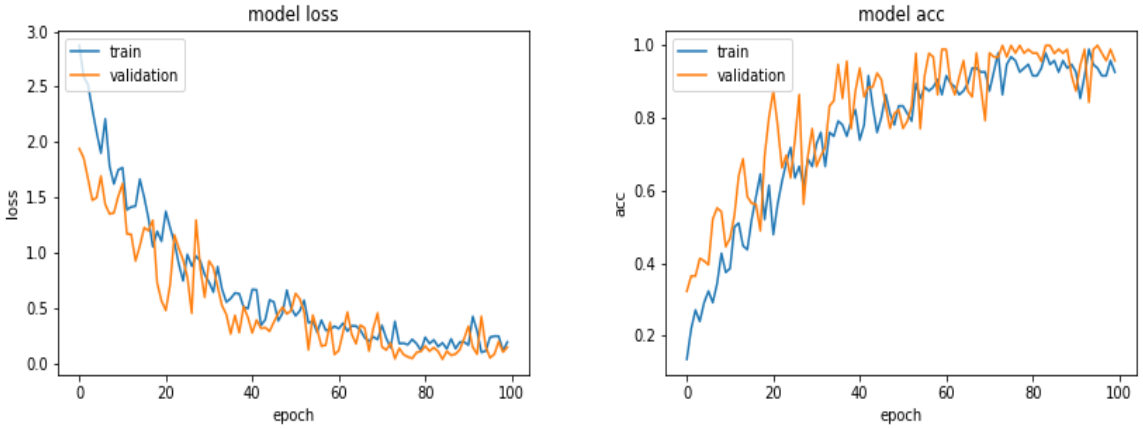


Figure 8. Evolution of the loss function and the accuracy as a function of the epoch numbers for the case of 360 images.

To illustrate the accuracy of the test results, we show in Figure 9 the probabilities of the prediction as an image. As we have 50 images in each of the 6 classes, the prediction probabilities can be showed as an image of 6 lines and 50 columns for each class.

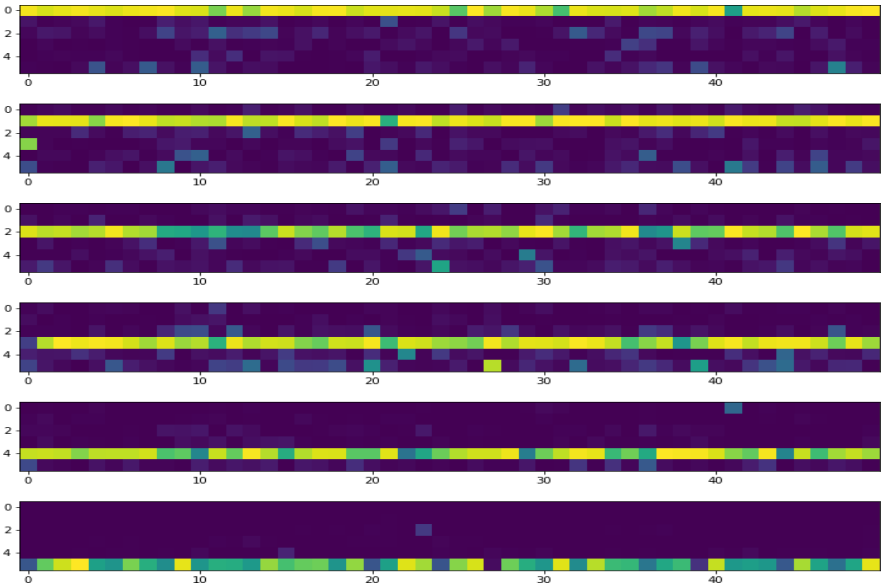


Figure 9. The prediction probabilities for each of the six classes are shown as an image. In each class, there are 50 images. The classes CVS1, ..., CVS6 are shown from top to down.

5.7. Results with 600 Images

Figure 10 shows the loss and the accuracy functions during the learning and evaluation process when we used 600 images for training, 300 for evaluation and 300 for testing.

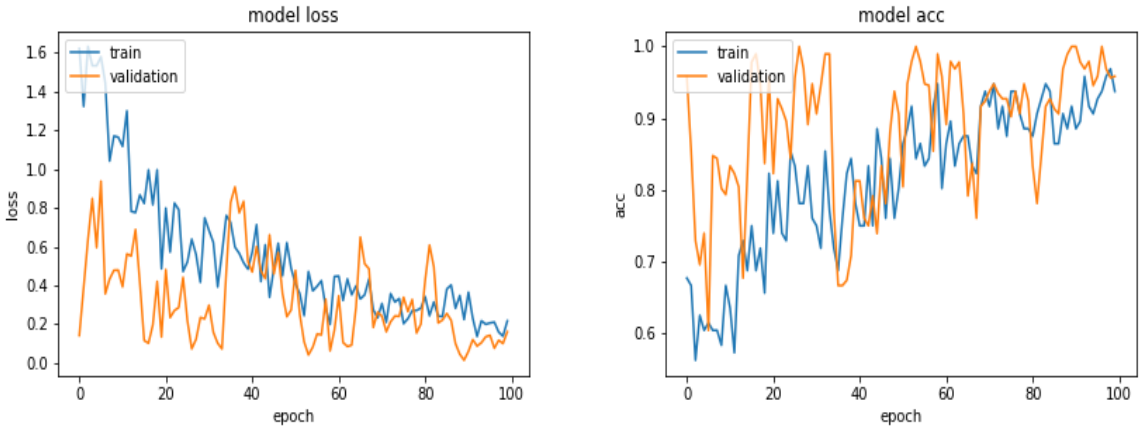


Figure 10. Evolution of the loss function and the accuracy as a function of the epoch numbers for thecase of 600 images.

5.8. Results with 900 Images

Figure 11 shows the loss and the accuracy functions during the learning and evaluation process when we used 900 images for training, 300 for evaluation and 300 for testing.

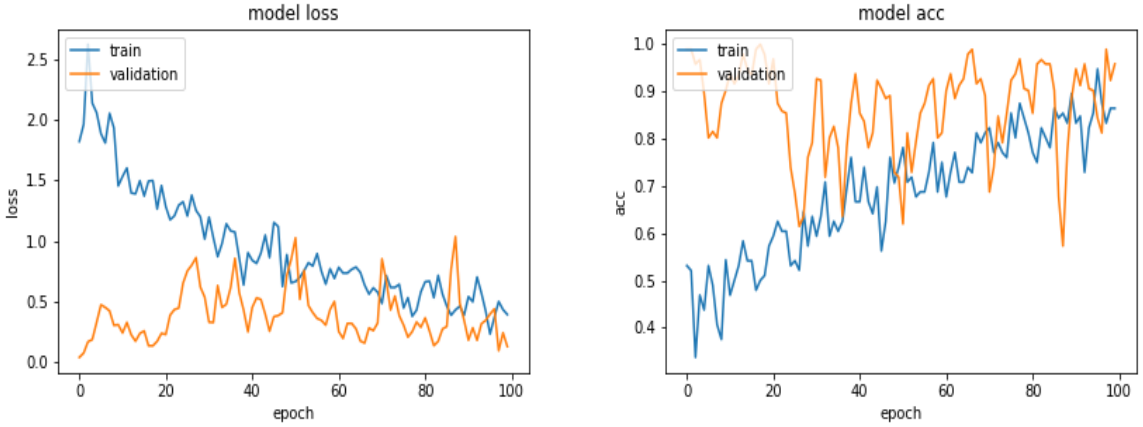


Figure 11. Evolution of the loss function and the accuracy as a function of the epoch numbers for thecase of 900 images.

5.9. Results with 1870 Images

Figure 12 shows the loss and the accuracy functions during the learning and evaluation process when we used all the available 1870 images for training, 300 for evaluation and 300 for testing.

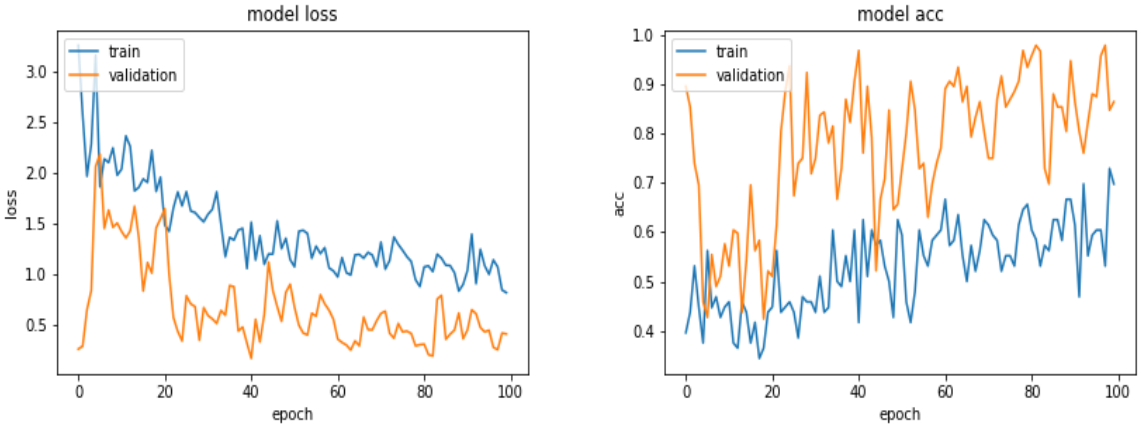


Figure 12. Evolution of the loss function and the accuracy as a function of the epoch numbers for thecase of 1870 images.

5.10. Comparison of Accuracy for Different Number of Training Images

Table 2 shows the obtained accuracy on 300 test images when the models have been trained with different number of images: 360, 600, 900 and 1870.

Table 2. Comparison of the classification accuracy for different number of training images.The optimization algorithm used for training was SGD.

No. of Images/Classes	CVS1	CVS2	CVS3	CVS4	CVS5	CVS6
360 images	0.917	0.977	0.970	0.980	0.870	0.850
600 images	1.000	0.997	1.000	0.993	1.000	0.997
900 images	0.980	0.890	0.993	0.963	0.933	0.927
1870 images	1.000	0.997	0.987	0.987	0.987	0.967

As we can see from this table that, in general, when the number of training images increases, the global accuracy improves. However, the improvement is not linear. Also, when the numbers of images in each class are the same, a better accuracy is obtained. This can be seen in the table when the numbers of images increases from 360 to 900. In those cases, the number of images in each class was equally distributed. In the last case, where we used all the images, this equilibrium is not, since we have 199, 184, 825, 300, 200 and 162 images, respectively, in CVS1, CVS2, CVS3, CVS4, CVS5 and CVS6.

The results in this table show that the accuracy is increased when going from 360 to 600, but then decreased when going to 900 or even to 1870 images. This is probably due to the great diversity of the images and the fact that the number of images in each class is not the same. We are still exploring this database.

5.11. Comparison of the Results for Different Optimization Algorithms

We tried different optimization algorithms. Between all, here we report on two classical ones: SGD and ADAM.

Figure 13 shows a comparison of the accuracy evolution during the optimization.

Table 3 shows a comparison of the accuracy obtained on the test data using Adam for different number of images during the training.

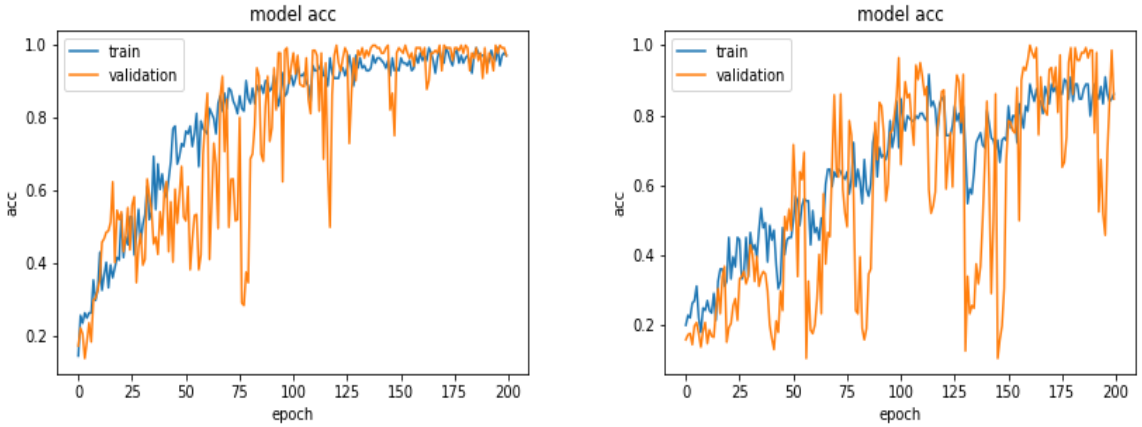


Figure 13. Evolution of the accuracy as a function of the epoch numbers for two different optimization algorithms SGD (**left**) and ADAM (**right**).

Table 3. Comparison of the classification accuracy for different number of training images.The optimization algorithm used for training was ADAM.

No. of Images / Classes	CVS1	CVS2	CVS3	CVS4	CVS5	CVS6
360 images	0.963	0.927	0.973	0.970	0.980	0.933
600 images	1.000	1.000	1.000	0.997	1.000	0.997
900 images	0.997	0.997	0.997	1.000	0.993	0.993
1870 images	0.990	0.970	0.953	0.980	0.977	0.990

5.12. Comparison of the Results for Different Preprocessing

We tried different preprocessing of the images before starting the training. Here, we report only the case of entropic filtering which was the most significant.

Figure 14 shows the evolution of the accuracy as a function of the epoch numbers for the two cases of without any filtering and with entropic filtering.

Table 4 shows the summary of the obtained accuracy during the different tests.

Table 4. Effect of preprocessing: Comparison between no preprocessing and entropy filter preprocessing.

No. of Images/Classes	CVS1	CVS2	CVS3	CVS4	CVS5	CVS6
300 images: No preprocessing	1.000	0.983	0.987	0.997	0.923	0.990
Entropy filter	1.000	1.000	1.000	1.000	1.000	1.000
600 images: No preprocessing	0.978	0.975	0.952	0.973	0.937	0.925
Entropy filter	0.995	0.997	0.990	0.997	0.995	0.995
900 images: No preprocessing	0.794	0.826	0.635	0.833	0.783	0.817
Entropy filter	0.984	0.984	0.988	0.996	0.993	0.996
1870 images: No preprocessing	0.879	0.880	0.651	0.833	0.885	0.911
Entropy filter	0.882	0.857	0.637	0.852	0.897	0.919

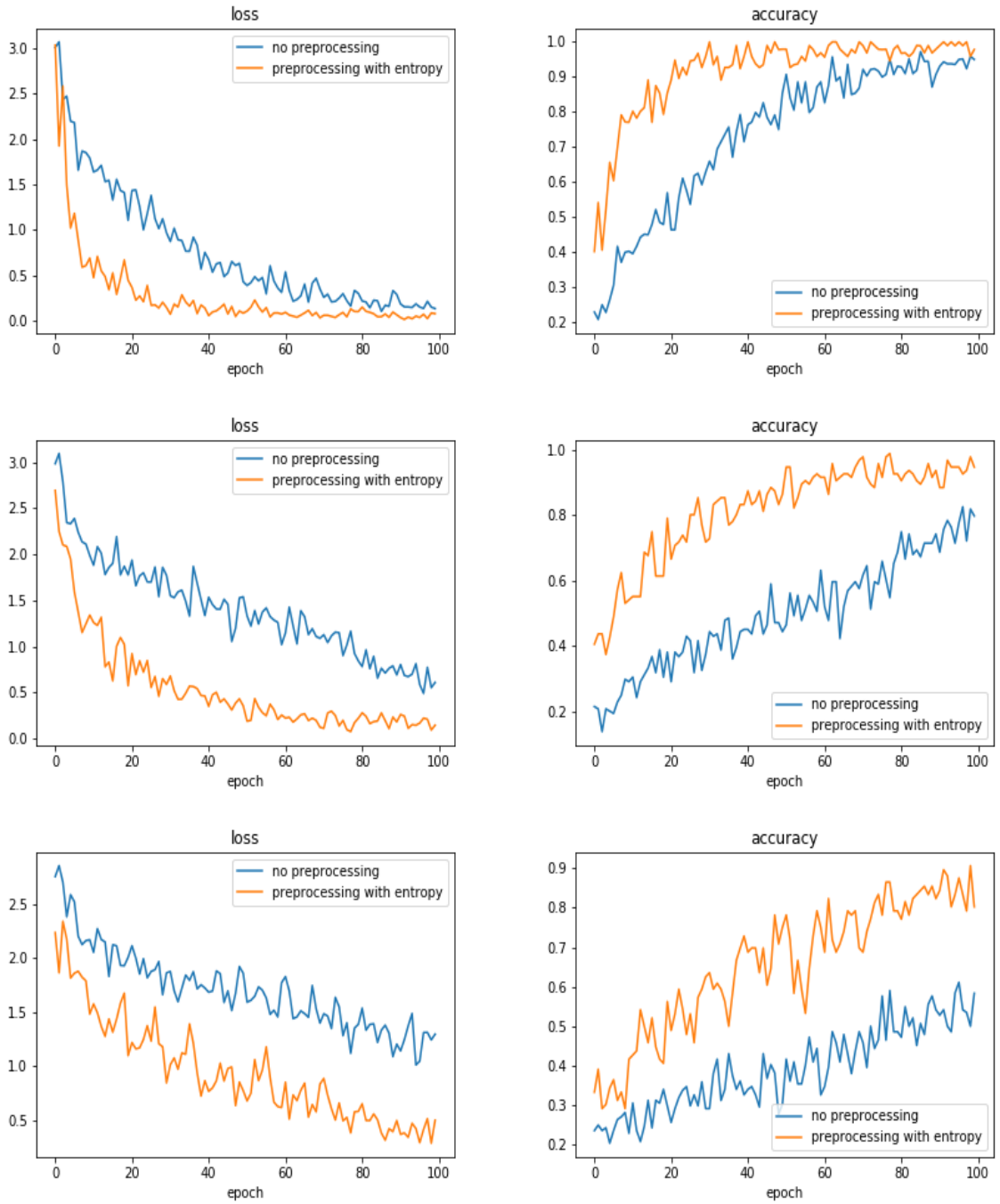


Figure 14. Evolution of loss and accuracy during the training as a function of the epoch numbers for two cases: without any filtering (**left**) and with entropic filtering (**right**). from top to botumn: The cases with 300, 600 and 900 training images.

5.13. Comparison of the Results as a Function of the Number of Layes

To see the performances of the proposed method as a function of the number of layers, we did a few experiments increasing the number of layers by one block of convolution, normalization, max-pooling and dropout layers. Figures 15 and 16 and Table 5 show the comparison between the proposed 6 layers and 6 + 1 layer. As we can see, increasing the number of layers did not resulted to better performances.

Table 5. Effect of the number of layers: Comparison between two cases of 6 and 7 layers. Increasing the number of layers decreases the score of the classification.

No. of Layers/Classes	CVS1	CVS2	CVS3	CVS4	CVS5	CVS6
300 images:						
6 layers:	0.913	0.850	0.802	0.912	0.895	0.913
7 layers:	0.892	0.805	0.890	0.793	0.885	0.897
900 images:						
6 layers:	0.998	0.998	1.000	0.988	0.994	0.991
7 layers:	0.930	0.939	0.952	0.924	0.966	0.969

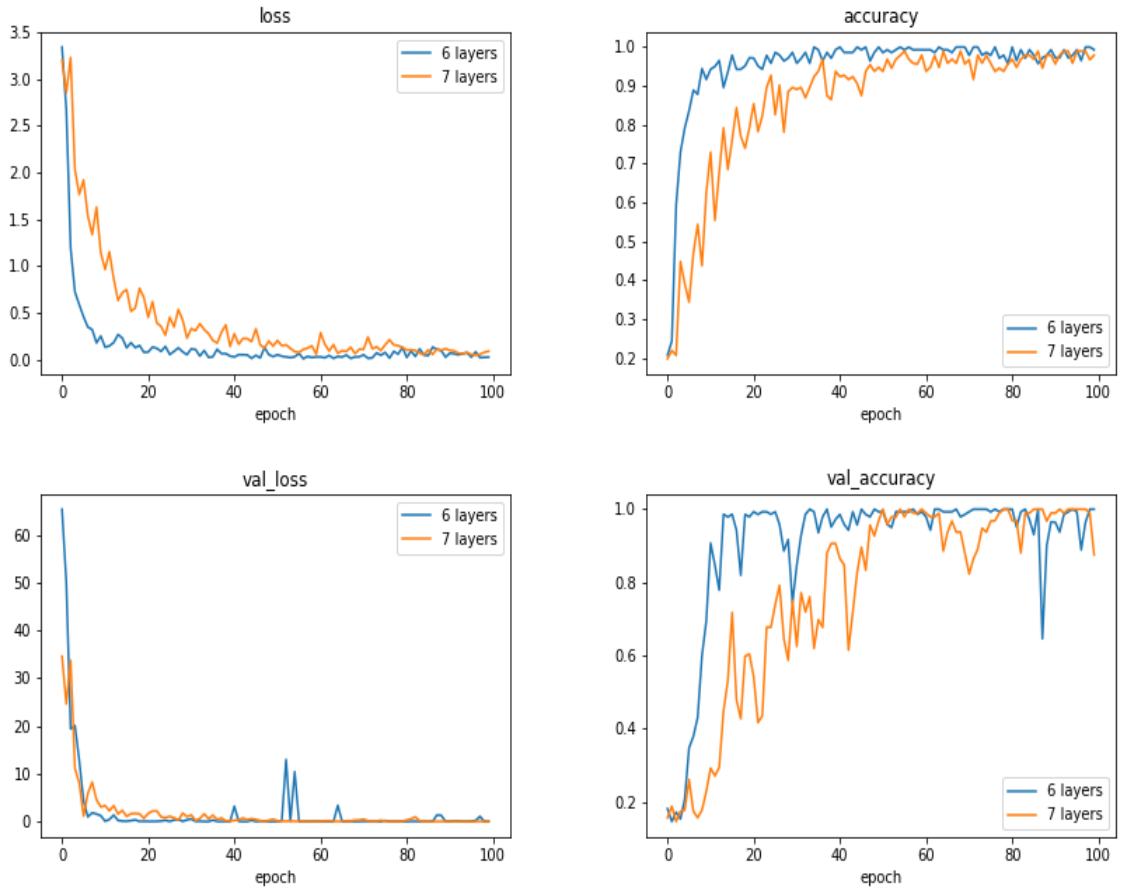


Figure 15. Evolution of loss and accuracy during the training (**upper row**) and during the validation (**lower row**) as a function of the epoch numbers for two cases: 6 layers and 7 layers networks. (The case with 300 images).

In the Table 6, we show the confusion matrix and the classification report for the two last cases: 900 images for training, 300 for validation and 900 for testing.

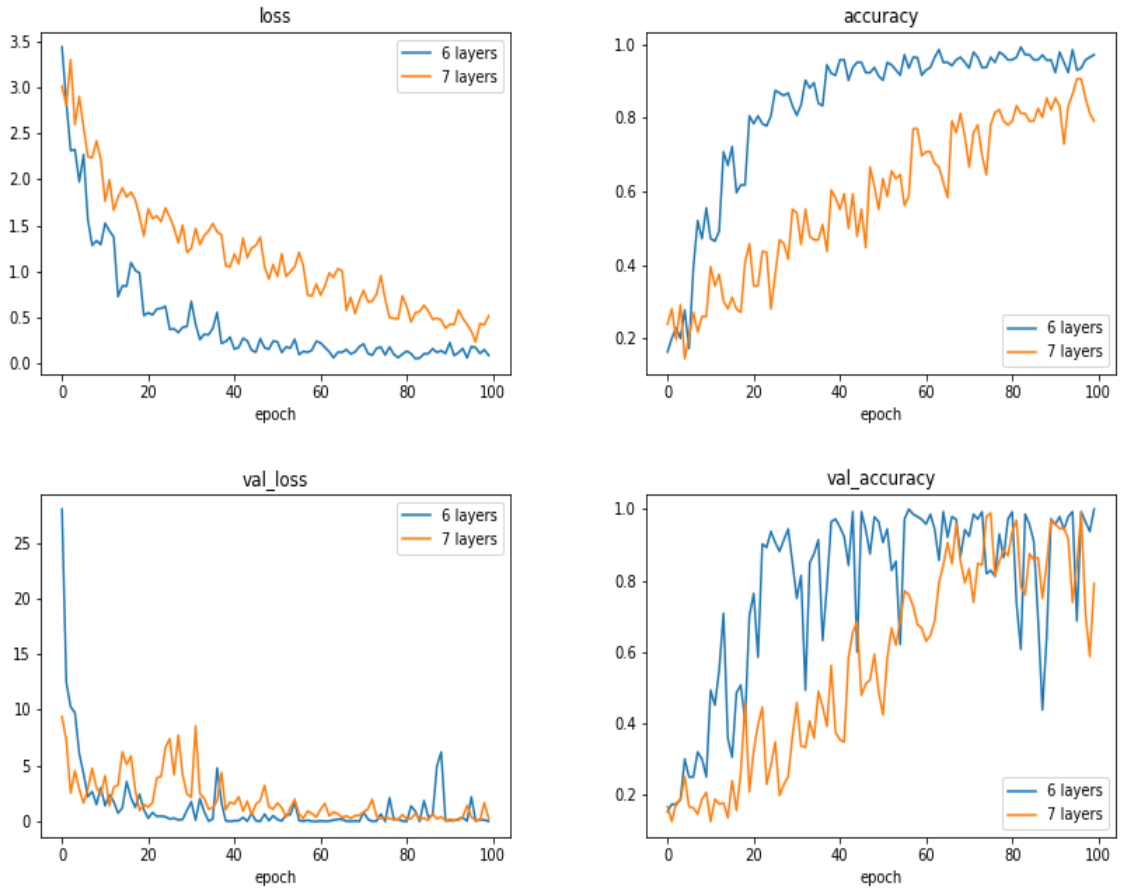


Figure 16. Evolution of loss and accuracy during the training (**upper row**) and during the validation (**lower row**) as a function of the epoch numbers for two cases: 6 layers and 7 layers networks. (The case with 900 images).

Table 6. Confusion matrix and classification report for the two cases of 6 layers and 7 layers.

Case with 6 Layers				Case with 7 Layers			
Confusion Matrix				Confusion Matrix			
[[101 6 4 25 12 2]				[[101 6 4 25 12 2]			
[0 110 3 30 5 2]				[0 110 3 30 5 2]			
[0 0 122 22 6 0]				[0 0 122 22 6 0]			
[0 0 0 148 2 0]				[0 0 0 148 2 0]			
[1 1 1 8 139 0]				[1 1 1 8 139 0]			
[0 0 0 16 2 132]]				[0 0 0 16 2 132]]			
Classification Report				Classification Report			
precision	recall	f1-score	support	precision	recall	f1-score	support
CVS1	0.99	0.67	0.80 150	CVS1	0.99	0.67	0.80 150
CVS2	0.94	0.73	0.82 150	CVS2	0.94	0.73	0.82 150
CVS3	0.94	0.81	0.87 150	CVS3	0.94	0.81	0.87 150
CVS4	0.59	0.99	0.74 150	CVS4	0.59	0.99	0.74 150
CVS5	0.84	0.93	0.88 150	CVS5	0.84	0.93	0.88 150
CVS6	0.97	0.88	0.92 150	CVS6	0.97	0.88	0.92 150

5.14. Discussions

As we can see from the experiments and the results obtained, we may make a few remarks:

- Many existing CNN and DL structures have been developed particularly for photography images. They could not be used as they are for our X-ray radiography images. First, in general, they are

constructed for colored images while the X-ray images are gray level. Using the pre-trained models cannot be used directly and even trying to fine tune them was not effective. Trying to train them directly was not successful, as in general, they need a very large number of images.

- We proposed a simple six-layer structure with five combination of convolution, normalization, max-pooling and dropout layers and one combination of dense, normalization and dropout layers. This structure was the most successful for our case.
- The choices of pre-processing, optimization criteria, optimization algorithms and optimization of hyper-parameters are very difficult. These need experience and deep knowledge of mathematics and techniques.
- We used categorical entropy as the optimization criteria, SGD or ADAM for optimization algorithm, and manually fixing the hyper-parameters using cross validation.
- Many performance criteria such as general accuracy, categorical accuracy, precision, recall and many other criteria, all can be derived from the confusion matrix and can be used to measure the performances of a classification method. In this paper, we mainly showed the accuracy.
- In our case, we show that when the training data are balanced, in general, the performances improve when the number of images increases. However, when the dataset is unbalanced, the situation is different.
- We did many experiments on the importance of the pre-processing of the images. Some pre-processings such as cropping on the interesting part of the image is almost always necessary, but other linear processing methods such as mean filtering or resizing did not have great impact. However, we discovered that an entropic filtering has improved the classification task.
- With small and moderate data, increasing the number layers does not improve the performances, and even, decrease them. A kind of balance between the number of parameters and the number of data has to be found. In our case,an approach of six compound layes (Convolution, Normalization, MaxPooling and DropOut) was the right one. When increased even one more layer, the performances decreased.

6. Conclusions

In this work, first, we described an image classification task which is important for the orthodontics community. More precisely, the objective was classifying the lateral radiographs of a great number of patients with the aim of determining the cervical vertebra maturation degree of bones, which is an important parameter for orthodontists. Then, we developed and presented a specifically designed classification method for this task.

In a first step, we tried different existing CNN and DL methods, models and algorithms to see if we can adapt and use them. The conclusion was that the existing models are, in general, for photography images. In our case, we had gray level X-ray radiography images. The adaptation of those methods could not be efficient. We then proposed a DL-CNN classification method which is particularly adapted for this task. The proposed model is a five-layer CNN with Batch Normalization, Max Pooling and Dropout layers. The optimization criteria is the Categorical Cross Entropy and the optimization algorithms we tested are SGD and Adam. To show the performances of the optimization algorithms during the training and evaluation, we used the training and validation loss and accuracy as function of number of epoches. Using the Confusion Matrix to compare the performances of the predictions, in this paper, we showed the obtained accuracy in each test.

To train the proposed model, in a first step, we used 360 first available labeled images for training,300 for validation and hyper parameter tuning and finally 300 for testing. With these data, even if during the training and validation, we could obtain an accuracy of more than 95%, the accracy for the testing images did not exceeded 90%. In a second step, we increased the number of the training dataset to 600, then to 900 and finally to all available 1870 images at the time of writing this paper. When going from 360 to 600 and 900 images equally distributed in the six classes, the permances improved and then stopped improving. For the case where we used the whole available 1870 images,

the improvement was not really significant due to the fact that the numbers of images in the six classes were not equally distributed and there was greater diversity.

The dataset is still going to grow and we try to have equally distributed samples. During this development, we continue to develop other network structures and test other optimization algorithms. A clinically usable package is also planned within our project. This paper is an extended version of a communication at MaxEnt 2019 international workshop [57].

Author Contributions: M.M., C.L. contributed more on orthodontics and data base preparation and A.M.-D. more on the CV, ML and classification method. All authors have read and agreed to the published version of the manuscript.

Funding: ISCT and Makaremi Orthodontics cabinet cooperaction contract.

Conflicts of Interest: No any conflicts of interest.

References

1. Baccetti, T.; Franchi, L.; McNamara, J.A. The Cervical Vertebral Maturation (CVM) Method for the Assessment of Optimal Treatment Timing in Dentofacial Orthopedics. *Semin. Orthod.* **2005**, *11*, 119–129. [CrossRef]
2. Patcas, R.; Herzog, G.; Peltomäki, T.; Markic, G. New perspectives on the relationship between mandibular and statural growth. *Eur. J. Orthod.* **2016**, *38*, 13–21. [CrossRef] [PubMed]
3. Moore, R.N.; Moyer, B.A.; DuBois, L.M. Skeletal maturation and craniofaciat growth. *Am. J. Orthod. Dentofac. Orthod.* **1990**, *98*, 33–40. [CrossRef]
4. Hunter, C.J. The corelation of facial growth with body height and skeletal maturation at adolescence.pdf. *Angle Orthod.* **1966**, *36*, 44–54.
5. Raberin, M.; Cozor, I.; Gobert-Jacquart, S. Les vertèbres cervicales: Indicateurs du dynamisme de la croissance mandibulaire? *Orthod. Fr.* **2012**, *83*, 45–58. [CrossRef]
6. Franchi, L.; Baccetti, T.; McNamara, J.A. Mandibular growth as related to cervical vertebral maturation and body height. *Am. J. Orthod. Dentofac. Orthop.* **2000**, *118*, 335–340. [CrossRef]
7. De Stefani, A.; Bruno, G.; Siviero, L.; Crivellin, G.; Mazzoleni, S.; Gracco, A. Évaluation radiologique de l'âge osseux avec la maturation de la phalange médiane du doigt majeur chez un patient orthodontique péripubertaire. *Int Orthod.* **2018**, *16*, 499–513. [CrossRef]
8. Krisztina, M.I.; Ogodescu, A.; Réka, G.; Zsuzsa, B. Evaluation of the Skeletal Maturation Using Lower First Premolar Mineralisation. *Acta Med. Marisiensis* **2013**, *59*, 289–292. [CrossRef]
9. Pyle, S.I.; Waterhouse, A.M.; Greulich, W.W. Attributes of the radiographic standard of reference for the National Health Examination Survey. *Am. J. Phys. Anthropol.* **1971**, *35*, 331–337. [CrossRef]
10. Loder, R.T. Applicability of the Greulich and Pyle Skeletal Age Standards to Black and White Children of Today. *Arch. Pediatr. Adolesc. Med.* **1993**, *147*, 1329. [CrossRef]
11. Bunch, P.M.; Altes, T.A.; McIlhenny, J.; Patrie, J.; Gaskin, C.M. Skeletal development of the hand and wrist: Digital bone age companion—A suitable alternative to the Greulich and Pyle atlas for bone age assessment? *Skelet. Radiol.* **2017**, *46*, 785–793. [CrossRef] [PubMed]
12. Srinivasan, B.; Padmanabhan, S.; Chitharanjan, A.B. Constancy of cervical vertebral maturation indicator in adults: A cross-sectional study. *Int. Orthod.* **2018**, *16*, 486–498. [CrossRef] [PubMed]
13. Shim, J.J.; Bogowicz, P.; Heo, G.; Lagravère, M.O. Interrelationship and limitations of conventional radiographic assessments of skeletal maturation. *Int. Orthod.* **2012**, *10*, 135–147. [CrossRef] [PubMed]
14. O'Reilly, M.T.; Yanniello, G.J. O'REILLY Mandibular growth changes and maturation of cervical vertebrae.pdf.
Angle Orthod. **1988**, *58*, 179–184. [PubMed]
15. Hassel, B.; Farman, A.G. Skeletal maturation evaluation using cervical vertebrae. *Am. J. Orthod. Dentofac. Orthop.* **1995**, *107*, 58–66. [CrossRef]
16. Elhaddaoui, R.; Benyahia, H.; Azaroual, F.; Zaoui, F. Intérêt de la méthode de maturation des vertèbres cervicales (CVM) en orthopédie dento-faciale: Mise au point. *Rev. Stomatol. Chir Maxillo-Faciale Chir Orale* **2014**, *115*, 293–300. [CrossRef]

17. Jaqueira, L.M.F.; Armond, M.C.; Pereira, L.J.; de Alcântara, C.E.P.; Marques, L.S. Determining skeletal maturation stage using cervical vertebrae: Evaluation of three diagnostic methods. *Braz. Oral Res.* **2010**, *24*, 433–437. [CrossRef]
18. Uysal, T.; Ramoglu, S.I.; Basciftci, F.A.; Sari, Z. Chronologic age and skeletal maturation of the cervical vertebrae and hand-wrist: Is there a relationship? *Am. J. Orthod. Dentofac. Orthop.* **2006**, *130*, 622–628. [CrossRef]
19. Hosni, S.; Burnside, G.; Watkinson, S.; Harrison, J.E. Comparison of statural height growth velocity at different cervical vertebral maturation stages. *Am. J. Orthod. Dentofac. Orthop.* **2018**, *154*, 545–553. [CrossRef]
20. Perinetti, G.; Contardo, L.; Castaldo, A.; McNamara, J.A.; Franchi, L. Diagnostic reliability of the cervical vertebral maturation method and standing height in the identification of the mandibular growth spurt. *Angle Orthod.* **2016**, *86*, 599–609. [CrossRef]
21. Mahajan, S. Evaluation of skeletal maturation by comparing the hand wrist radiograph and cervical vertebrae as seen in lateral cephalogram. *Indian J. Dent. Res.* **2011**, *22*, 309. [CrossRef] [PubMed]
22. Sachan, K.; Tandon, P.; Sharma, V. A correlative study of dental age and skeletal maturation. *Indian J. Dent. Res.* **2011**, *22*, 882. [CrossRef] [PubMed]
23. Danaei, S.M.; Karamifar, A.; Sardarian, A.; Shahidi, S.; Karamifar, H.; Alipour, A.; Ghodsi Boushehri, S. Measuring agreement between cervical vertebrae and hand-wrist maturation in determining skeletal age: Reassessing the theory in patients with short stature. *Am. J. Orthod. Dentofac. Orthop.* **2014**, *146*, 294–298. [CrossRef] [PubMed]
24. Ball, G.; Woodside, D.; Tompson, B.; Hunter, W.S.; Posluns, J. Relationship between cervical vertebral maturation and mandibular growth. *Am. J. Orthod. Dentofac. Orthop.* **2011**, *139*, e455–e461. [CrossRef]
25. Ballrick, J.W.; Fields, H.W.; Beck, F.M.; Sun, Z.; Germak, J. The cervical vertebrae staging method's reliability in detecting pre and post mandibular growth. *Orthod. Waves* **2013**, *72*, 105–111. [CrossRef]
26. Perinetti, G.; Caprioglio, A.; Contardo, L. Visual assessment of the cervical vertebral maturation stages: A study of diagnostic accuracy and repeatability. *Angle Orthod.* **2014**, *84*, 951–956. [CrossRef]
27. McNamara, J.A.; Franchi, L. The cervical vertebral maturation method: « A user's guide ». *Angle Orthod.* **2018**, *88*, 133–143. [CrossRef]
28. Santiago, R.C.; Cunha, A.R.; Júnior, G.C.; Fernandes, N.; Campos, M.J.S.; Costa, L.F.; Vitral, R.W.; Bolognese, A.M. New software for cervical vertebral geometry assessment and its relationship to skeletal maturation—A pilot study. *Dentomaxillofac. Radiol.* **2014**, *43*, 20130238. [CrossRef]
29. Padalino, S.; Sfondrini, M.F.; Chenuil, L.; Scudeller, L.; Gandini, P. Fiabilité de l'analyse de la maturité squelettique selon la méthode CVM des vertèbres cervicales faite par un logiciel dédié. *Int. Orthod.* **2014**, *12*, 483–493. [CrossRef]
30. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* **2012**, 1097–1105. [CrossRef]
31. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
32. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
33. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
34. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
35. Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of

the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.

36. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.
37. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.
38. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas,
39. Tajbakhsh, N.; Shin, J.Y.; Gurudu, S.R.; Hurst, R.T.; Kendall, C.B.; Gotway, M.B.; Liang, J. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *IEEE Trans. Med. Imaging* **2016**, *35*, 1299–1312. [CrossRef] [PubMed]
40. Kamnitsas, K.; Ledig, C.; Newcombe, V.F.J.; Simpson, J.P.; Kane, A.D.; Menon, D.K.; Rueckert, D.; Glocker, B. Efficient multi-scale 3d cnn with fully connected CRF for accurate brain lesion segmentation. *Med. Image Anal.* **2017**, *36*, 61–78. [CrossRef] [PubMed]
41. Kathirvel, C.T.R. Classifying Diabetic Retinopathy using Deep Learn-ing Architecture. *Int. J. Eng. Res.-Tech-Nol.* **2016**, *5*. 19–24. [CrossRef]
42. Lessmann, N.; Isgum, I.; Setio, A.; de Vos, B.D.; Ciompi, F.; de Jong, P.A.; Oudkerk, M.; Willem, P.T.M.; Viergever, M.A.; van Ginneken, B. Deep convolutional neural networks for automatic coronary calcium scoring in a screening study with low-dose chest CT. In *SPIE Medical Imaging*; International Society for Optics and Photonics: Bellingham, WA, USA, 2016; p. 978511.
43. Alansary, A.; Oktay, O.; Li, Y.; Le Folgoc, L.; Hou, B.; Vaillant, G.; Kamnitsas, K.; Vlontzos, A.; Glocker, B.; Kainz, B.; et al. Evaluating reinforcement learning agents for anatomical landmarkdetection. *Med. Image Anal.* **2019**, *53*, 156–164. [CrossRef]
44. Sahiner, B.; Pezeshk, A.; Hadjiiski, L.M.; Wang, X.; Drukker, K.; Cha, K.H.; Summers, R.M.; Giger, M.L. Deep learningin medical imaging and radiation therapy. *Med. Phys.* **2019**, *46*, e1–e36. [CrossRef]
45. Frid-Adar, M.; Klang, E.; Amitai, M.; Goldberger, J.; Greenspan, H. Synthetic data augmentation using GAN for improved liverlesion classification. In Proceedings of the 2018 IEEE 15th International Symposiumon Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 289–293.
46. Gao, F.; Wu, T.; Li, J.; Zheng, B.; Ruan, L.; Shang, D.; Patel, B. SD-CNN: A shallow-Deep CNN for Improved Breast Cancer diagnosis. *Comput. Med. Imaging Graph.* **2018**, *70*, 53–62.
47. Samala, R.K.; Chan, H.P.; Hadjiiski, L.M.; Helvie, M.A.; Richter, C.; Cha, K. Evolutionary pruning of transfer learned deep convolutional neural network for breast cancer diagnosis in digital breast tomosynthesis. *Phys. Med. Biol.* **2018**, *63*, 095005. [CrossRef]
48. Al, W.A.; Yun, I.D. Reinforcing Medical Image Classifier to Improve Generalization on Small Datasets. Available online: https://arxiv.org/abs/1909.05630v2 (accessed on 15 October 2019).
49. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. Available online: https://arxiv.org/abs/1512.03385 (accessed on 10 October 2019).
50. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. Available online: https://arxiv.org/abs/1608.06993 (accessed on 10 October 2019).
51. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. Available online: https://papers.nips.cc/paper/4824-imagenet-classification-with-deep- convolutional-neural-networks.pdf (accessed on 10 October 2019).
52. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. Available online: https://arxiv.org/pdf/1409.1556.pdf (accessed on 10 October 2019).
53. Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level Accuracy with 50× Fewer Parameters and <0.5 MB Model Size. Available online: https://arxiv.org/abs/ 1602.07360 (accessed on 10 October 2019).
54. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J. Rethinking the Inception Architecture for

Computer Vision. Available online: https://arxiv.org/pdf/1512.00567v1.pdf (accessed on 10 October 2019).
55. Ruder, S. An Overview of Gradient Descent Optimization Algorithms. Available online: http://ruder.io/ optimizing-gradient-descent (accessed on 10 October 2019).
56. Ruder, S. An Overview of Gradient Descent Optimization Algorithms. Available online: https://arxiv.org/ abs/1609.04747 (accessed on 10 October 2019).
57. Special Issue “MaxEnt 2019—The 39th International Workshop on Bayesian Inference and Maximum EntropyMethods in Science and Engineering”, accessed on line 29 September 2019.

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution(CC BY) license (http://creativecommons.org/licenses/by/4.0/).