



Sélection de caractéristiques stables pour la segmentation d'images histologiques par calcul haute performance

Clément Bouvier

► To cite this version:

Clément Bouvier. Sélection de caractéristiques stables pour la segmentation d'images histologiques par calcul haute performance. Bio-informatique [q-bio.QM]. Sorbonne Université, 2019. Français. NNT : 2019SORUS004 . tel-03115972

HAL Id: tel-03115972

<https://theses.hal.science/tel-03115972>

Submitted on 20 Jan 2021

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.



THÈSE DE DOCTORAT DE SORBONNE UNIVERSITÉ

Ecole Doctorale Informatique, Télécommunications et Électronique
(Paris)

Présentée par

Clément BOUVIER

Pour obtenir le grade de

DOCTEUR de SORBONNE UNIVERSITÉ

Sujet de la thèse :

Sélection de caractéristiques stables pour la segmentation d'images histologiques par calcul haute performance

soutenue le 18 janvier 2019

devant le jury composé de :

M. Vincent	FROUIN	Rapporteur
M. Aymeric	HISTACE	Rapporteur
M. Dominique	BÉRÉZIAT	Examineur
M. Christophe	DUCOTTET	Examineur
Mme. B. Helena	RODRIGUEZ	Examinatrice
M. Thierry	DELZESCAUX	Directeur de thèse
M. Cédric	CLOUCHOUX	Co-encadrant de thèse, invité au jury

Remerciements :

Cette période de thèse fut pleine de nombreuses rencontres et d'événements souvent trop acides ou trop sucrés. Il faut apprendre à casser et reconstruire, à déchanter et s'émerveiller, se remettre constamment en question pour espérer exercer. Mais j'ai eu énormément de chance durant ces dernières années, principalement celle d'avoir été très bien entouré. Je vais donc tenter de rendre hommage à ces personnes formidables auxquelles je dois énormément.

Car il faut bien commencer par quelque part, je tiens à remercier chaleureusement mon jury de thèse pour avoir accepté d'évaluer mon travail : M. Dominique Béréziat (président du jury), M. Vincent Frouin (rapporteur), M. Aymeric Histace (rapporteur), M. Christophe Ducottet (examinateur) et Mme. Helena Rodriguez (examinatrice). J'ai pris note de tous vos commentaires, soyez assurés qu'ils ne seront pas ignorés. Un merci plus spécifique pour mes rapporteurs qui, avec leurs remarques justes et constructives, m'ont permis d'améliorer ma soutenance.

Je te remercie, Thierry Delzescaux, de m'avoir accepté en tant que thésard. Tu m'as beaucoup apporté que ce soit sur le plan professionnel ou humain. J'ai parfois un comportement particulier et une capacité innée de me noyer dans un verre d'eau. Mais malgré cela, tu as toujours su me pousser et m'accompagner afin que je puisse donner le meilleur de moi-même (sic). Tu as aussi une manière très saine d'aborder la recherche, emprunte de pragmatisme et d'humanité. Le bien-être de tes collaborateurs est une de tes priorités, et j'ai pu l'apprécier durant ces trois dernières années. Je pense que je n'aurais pas pu espérer meilleur directeur de thèse.

C'est ton tour Cédric Clouchoux. Merci pour ce soutien indéfectible durant ces trois ans. Tu es un incroyable encadrant et collègue. Merci pour ces conversations autour de l'avenir et de la musique, j'espère avoir pu te partager mon amour des nanars auditifs. Sans ton abnégation, il n'y aurait pas eu d'article de conférence et de voyage à Montréal. Ce fut un plaisir de découvrir cette ville chère à ton cœur par tes yeux et tes (vieilles) habitudes. J'ai hâte de continuer les développements et la recherche à 3DNS sous ta direction éclairée. Je profite de ce paragraphe pour rendre hommage à CAPSUL, l'arlésienne parmi les arlésiennes, petit ange parti trop tôt.

L'heure est au sire Souedet, le deuxième encadrant mais malheureusement non-officiel (je suis d'accord avec Thierry, tu aurais dû t'asseoir à la table du jury pour me cuisiner). Merci Nicolas pour les conversations autour de la politique et de notre bien trop vieille 5ème république, tu m'as permis de mettre de l'eau dans un vin bien trop âpre. Tu m'as aussi beaucoup appris en développement informatique et j'aurais aimé être bien meilleur afin de t'éviter le surplus de travail que tu as eu lors de ces derniers mois. J'espère avoir l'occasion de me rattraper dans le futur et t'éviter des tracas.

Je souhaite aussi remercier toutes les personnes de MIRCen qui m'ont accompagné et aidé durant ces trois dernières années. Je tiens à remercier Marc, Anne-Sophie et Caroline pour leur aide sur les aspects biologiques de mon travail. Je remercie aussi tous les doctorants de MIRCen. Un big up tout particulier à Clément pour toutes ces soirées de zic et de philosophie. On réitère quand tu veux (j'anticipe aussi la prochaine Colonie de Vacances) !

Je veux aussi rendre hommage à l'équipe de 3DNS et à la société NEOXIA. J'ai passé d'excellents moments à vos côtés. Bonne chance à Clara et Sébastien pour vos thèses respectives, je suis persuadé que vous allez tout défoncer (vos origines géographiques sont au final vos seules limites (en même temps, vous auriez pu faire un effort)) ! Bon courage à mon futur encadré (oui, toi l'ardéchois nommé Jérémy) et à Naaïm, vous allez devenir de supers ingénieurs. Je te remercie Gilles pour ton implication dans le projet et de ton humanité avec tes partenaires, tu m'as permis de voir une autre facette du patronat. Helena, tu m'as énormément appris sur plein de sujets, tu es probablement la personne la plus intelligente que je connaisse. Un merci pêle-mêle à Flavie, Mohamed-Amine et Yann.

Je tiens à remercier toute ma famille pour leur support tout au long de mon existence. Je n'aurais jamais pu faire ce cursus sans votre aide et votre soutien, surtout lors des dernières semaines de mon doctorat ! Je vous remercie aussi les potes pour votre présence à mes côtés, les parties de jv sans aucune mauvaise foi, pour les concerts et les raves toujours préparés avec une rigueur militaire. Je crois pouvoir parler pour vous en disant qu'il n'y a jamais eu d'excès et que la modération fut, au final, notre maître mot. Du coup, je mets les mains en l'air pour vous : Armand, Ilan, Kevin, Nicolas et Sara.

Pour finir, je remercie les personnes que je n'ai pas citées et qui m'apprécient, les groupes qui peuplent mon écoute (spéciale dédicace à Incapacitants), l'équipe de R&M et Fréquence Moderne. La fin est là, la fin d'une suite d'années l'attendant en messie, la fin de trop nombreuses pâtes parfumées au pesto fade et brûlées au fond de la casserole en attendant la faim. Je tourne en rond mais pas dans le même plan.

Oui.

Sommaire :

Introduction	9
I Contexte	13
1. Système nerveux central (SNC)	13
1.1. Tissu nerveux	13
1.2. Anatomie et fonction du SNC	14
1.3. Maladie d'Alzheimer et recherche préclinique	15
2. Imagerie cérébrale	16
2.1. Imagerie <i>in-vivo</i>	16
2.1.1. TEP	16
2.1.2. IRM	17
2.2. Imagerie <i>post-mortem</i>	18
2.2.1. Protocoles de production des tissus	18
2.2.2. Imagerie de sections histologiques entières	20
2.3. Quantité de données numérisées	22
3. Problématique	22
II Etat de l'art	24
4. Quantification des marqueurs histologiques	24
4.1. Stratégie de quantification du tissu marqué	24
4.2. Méthodes de quantification des données d'histologie	25
4.2.1. Méthodes de quantification manuelles	25
4.2.2. Méthodes de quantification semi-automatiques	25
4.2.3. Méthodes de quantification automatiques	26
5. Méthodes d'apprentissage automatique supervisée	28
5.1. Evaluation et validation d'une segmentation	28
5.2. Méthodes d'apprentissage automatique classiques en histologie	30
5.2.1. Machine à support de vecteur	30
5.2.2. Modèle de mélange de gaussiennes	31
5.2.3. Forêt aléatoire	32
5.2.4. Apprentissage profond (<i>Deep Learning</i>)	33
5.3. Définition des caractéristiques	35
5.3.1. Espaces colorimétriques	35
5.3.2. Espaces décrivant la texture	39
5.3.3. Complexité algorithmique	43
6. Réduction de dimensionnalité	44
6.1. Réduction de dimensionnalité et problème de stabilité	44
6.2. Familles de méthodes de sélection de caractéristiques	45
6.2.1. Méthodes à filtres	45
6.2.2. Méthodes incluses	46
6.2.3. Méthodes intégrées	46
6.3. Objectifs et difficultés de la méthode de sélection souhaitée	48

III Extraction et sélection de caractéristiques	49
7. Bases de données biologiques	49
7.1. Cadre éthique de l'étude	49
7.2. Étude du marqueur NeuN	49
7.2.1. Production des données histologiques NeuN	49
7.2.2. Segmentations manuelles de NeuN	50
7.3. Étude du marqueur anti-Phox2B	51
7.3.1. Production des données histologiques anti-Phox2B	51
7.3.2. Segmentations manuelles d'anti-Phox2B	53
7.4. Vecteur initial de caractéristiques utilisé pour NeuN et anti-Phox2B	54
8. Développements logiciel réalisés	55
8.1. Matériel informatique utilisé	55
8.2. Environnement logiciel BrainVISA	56
8.3. Algorithmes de segmentation	57
8.3.1. BioVision	57
8.3.2. Multi-seuillage	57
8.3.3. Forêt aléatoire	58
8.4. Extraction de caractéristiques	58
9. Exploration des espaces de caractéristiques par Force Brute	60
9.1. Principe et objectifs de l'exploration par Force Brute	60
9.2. Méthode de sélection des familles de caractéristiques	63
9.2.1. Définition d'une famille de caractéristiques	63
9.2.2. Critères de sélection des familles	64
9.2.3. Sélection des meilleures familles de caractéristiques	65
10. Sélection de caractéristiques par Force Brute	66
10.1. Principe et objectifs de la sélection par Force Brute	66
10.2. Méthode de sélection des caractéristiques	67
10.2.1. Algorithme de la méthode proposée	67
10.2.2. Critères de sélection de caractéristiques	68
IV Mise en œuvre et validation de la méthode	70
11. Comparaison des méthodes de segmentation	70
11.1. Résultats obtenus avec un marqueur spécifique : NeuN	70
11.2. Résultats obtenus avec un marqueur non-spécifique : anti-Phox2B	72
12. Sélection de caractéristiques pour le marqueur spécifique NeuN	74
12.1. Rappel de l'objectif de la segmentation	74
12.2. Sélection basée sur la Force Brute	75
12.2.1. Familles sélectionnées pour NeuN	75
12.2.2. Vecteur de caractéristiques stables	80
12.3. Gain computationnel	87
13. Sélection de caractéristiques pour un marqueur non-spécifique : anti-Phox2B	88
13.1. Rappel de l'objectif de la segmentation	88
13.2. Sélection basée sur la Force Brute	88
13.2.1. Familles sélectionnées pour anti-Phox2B	88
13.2.2. Vecteur de caractéristiques stables	92

13.3. Gain computationnel	99
Discussion et Perspectives	100
Annexes :	
Annexe 1 : Régions sélectionnées pour la base NeuN	108
Annexe 2 : Tableaux des familles triées par les VMeP et VMoP	113
Annexe 3 : Publications	118
Références bibliographiques	119
Résumé et Abstract	131

Introduction

Les maladies neurodégénératives sont caractérisées par la perte progressive des neurones, cellules spécialisées dans la transmission de l'influx nerveux. La maladie d'Alzheimer a la plus grande prévalence parmi les autres maladies neurodégénératives. Elle est caractérisée par deux lésions provoquant une atrophie cérébrale qui se traduit par la perte de mémoire, l'apathie¹, l'apraxie², l'aphasie³ et enfin la mort. Malheureusement, il n'existe à ce jour aucune thérapie permettant de stopper l'évolution de cette maladie.

Au cours du développement de nouvelles thérapies contre la maladie d'Alzheimer, la recherche translationnelle est un paradigme de recherche utilisé pour limiter les risques de ses nouveaux médicaments et d'améliorer la santé des malades lors des recherches cliniques. La recherche préclinique étudie l'effet de ces nouvelles thérapies sur les cellules et sur les modèles animaux qui sont soumis à un état physique ou psychique similaire à un état humain. Ces thérapies peuvent être évaluées grâce à différents outils, notamment l'imagerie biomédicale est principalement utilisée pour quantifier des biomarqueurs d'intérêt, molécules se fixant spécifiquement sur des molécules d'intérêt. D'autre part, l'histologie est aussi un outil à la disposition des chercheurs. Elle correspond à la science de l'étude des tissus biologiques *ex vivo* ou *post mortem*. Les tissus étudiés par cette technique peuvent être colorés par des marqueurs, révélant des objets biologiques invisibles à l'œil nu comme les neurones ou des agrégats pathologiques dans le cas de la maladie d'Alzheimer. Elle repose principalement sur l'étude qualitative ou quantitative des tissus observés par des microscopes optiques [Gartner, 2015][Vandenberghe *et al.*, 2016] et qui peuvent numériser les tissus à une résolution sub-microscopique⁴. Cela permet d'observer les cellules les constituant mais aussi augmente drastiquement le nombre de pixels générés, la mémoire physique nécessaire à la conservation des images ainsi que les ressources de calcul nécessaires pour les analyser.

L'étude quantitative de ces images peut être réalisée manuellement par des opérateurs humains. Par exemple, la stéréologie et la segmentation manuelle requièrent beaucoup de temps et sont fastidieuses ce qui les rend difficilement applicables à de nombreux organes entiers. Ces méthodes sont cependant les plus utilisées à ce jour par les biologistes. Néanmoins, on constate que les segmentations manuelles sont de plus en plus remplacées par des algorithmes permettant un traitement automatique rapide et efficace des données. Les algorithmes d'apprentissage automatique supervisé sont de plus en plus utilisés dans ce contexte car ils intègrent les segmentations manuelles réalisées par un expert afin de proposer des segmentations semblables. Leur but est d'attribuer une classe à un pixel, cette classe pouvant être par exemple la présence de biomarqueurs [Doyle *et al.*, 2008] [Vandenberghe *et al.*, 2016]. Cependant les différentes étapes de l'algorithme nécessitent le calcul de caractéristiques, qui sont des informations supplémentaires extraites ou calculées à partir de la donnée brute. Le nombre de caractéristiques est généralement compris entre quelques dizaines et quelques milliers, multipliant d'autant la quantité de données. Le temps de traitement est alors

¹ Apathie : Insensibilité, indifférence, absence de réaction aux stimuli habituels, acceptée sans inquiétude par un patient dont les intérêts intellectuels sont amoindris à des degrés variables. [Dictionnaire de l'Académie de médecine, 2018].

² Apraxie : Trouble acquis de la réalisation intentionnelle des gestes finalisés, non explicable par une atteinte sensorimotrice élémentaire, une anomalie de la coordination, une altération de l'attention ou de la vigilance, un affaiblissement intellectuel important, ni par une perturbation de la compréhension secondaire à une aphasie. [Dictionnaire de l'Académie de médecine, 2018].

³ Aphasie : Trouble acquis du langage secondaire à une lésion cérébrale habituellement de l'hémisphère gauche. [Dictionnaire de l'Académie de médecine, 2018].

⁴ Résolution inférieure au micromètre.

fortement accru par l'extraction de ces nombreuses caractéristiques. Ce point limite fortement l'intérêt de l'emploi de des méthodes automatiques pour les études exhaustives de données massives (incluant plusieurs dizaines de Téraoctets) que ce soit sur des stations de travail classiques ou même sur des infrastructures de calcul haute performance (CHP).

Pour pallier ce problème, une solution est de réduire le nombre de caractéristiques, en utilisant des algorithmes de réduction de dimensionnalité, comme les **méthodes de sélection de caractéristiques** [Guyon, 2003]. Ces méthodes permettent d'alléger le modèle de données en estimant la pertinence d'une caractéristique ou d'un jeu de caractéristiques pour le procédé d'apprentissage automatique. Les étapes d'extraction et de classification s'opèrent donc avec un nombre restreint d'éléments, ce qui réduit le coût calculatoire. Cependant la sélection de caractéristiques a des effets secondaires comme le phénomène de pointe, pouvant ponctuellement améliorer la qualité de segmentation mais en sacrifiant la stabilité des ensembles sélectionnés [Kuncheva *et al.*, 2018]. La stabilité est devenue une propriété cruciale à prendre en compte afin de garantir un certain degré de reproductibilité et de généralité de ces méthodes. Plusieurs travaux ont proposés diverses méthodes en prenant en compte la stabilité sans toutefois en faire un critère de sélection [Meinshausen *et al.*, 2010][Awada *et al.*, 2012][Shah *et al.*, 2012][Kuncheva *et al.*, 2018].

L'objectif de mon travail de recherche est de développer une méthode permettant de sélectionner des ensembles de caractéristiques minimaux et stables pour la segmentation de biomarqueurs sur des images histologiques de grandes dimensions. Cette méthodologie doit aussi permettre une réduction du temps de traitement des images. Ce travail a été réalisé dans le cadre d'une convention CIFRE entre la société NEOXIA et le laboratoire des maladies neurodégénératives CNRS/CEA UMR-9199 de MIRcen (Molecular Imaging Research Center). MIRcen et NEOXIA sont impliqués dans le Projet d'Investissement d'Avenir 3DNS « 3DNeuroSecure » (contrat numéro P112331-3422142). Ce travail a pu bénéficier des ressources de ces deux entités. MIRcen m'a donné l'accès à BrainVISA (la plateforme logicielle collaborative), aux plateformes d'histologie et de traitement d'image. De plus, MIRcen a pu me faire bénéficier de sa grande expertise dans la recherche préclinique par le biais de discussions régulières avec plusieurs scientifiques de différents domaines. Quant à NEOXIA, la société m'a permis d'avoir des accès privilégiés aux formations sur les outils *Google* et aux infrastructures virtuelles de calculs hautes performances (*Cloud Computing*). Grâce à la société, j'ai pu acquérir une plus grande rigueur de développement et cela m'a permis d'industrialiser mes procédés. Les développements logiciels résultants de ce travail seront diffusés sur la plateforme BrainVISA et sont en cours d'industrialisation pour être portés sur la plateforme de micro-services 3DNeuroSecure.

Ce manuscrit est divisé en quatre chapitres.

Le chapitre I est dédié à la description du contexte biologique de ma thèse. Il décrit brièvement le tissu nerveux et détaille les principales techniques d'imagerie utilisées à MIRcen, notamment le protocole de production des tissus et leur numérisation. Ce chapitre expose aussi les problématiques et les différentes données traitées ainsi que leur utilité pour la recherche préclinique.

Le chapitre II présente l'état de l'art du problème de quantification en histologie. Il se concentre plus particulièrement sur la segmentation automatique avec des techniques d'apprentissage automatique et sur les méthodes de réduction de dimensionnalité. Les

problèmes liés aux méthodes de sélection de caractéristiques y sont présentés. Les objectifs de la méthode proposée sont aussi présentés.

Le chapitre III détaille les bases des données biologiques traitées et la méthodologie proposée. Plusieurs propositions ont résulté de mon travail:

- Une méthodologie d'explorations des espaces de caractéristiques
- Des critères pour analyser et trier les caractéristiques
- La définition d'une échelle intermédiaire entre les espaces et ses caractéristiques
- Un algorithme original de sélection de caractéristiques

Le chapitre IV compare différents algorithmes d'apprentissage automatique et analyse les résultats produits avec notre méthodologie sur deux marqueurs différents le Neuronal Nuclei (NeuN) et l'anticorps anti-Phox2B. Ces marqueurs sont utilisés sur différents organes du système nerveux central de macaque : l'encéphale pour NeuN et le tronc cérébral pour anti-Phox2B.

Le manuscrit se conclut sur une discussion des différents résultats et de la méthodologie proposée ainsi que sur les perspectives possibles à mon travail. Nous nous concentrerons sur les traitements de post-segmentation et sur leur intérêt pour d'autres domaines que la neurologie.

I / Contexte

1. Système nerveux central et la maladie d'Alzheimer

La biologie est la science de l'étude du vivant. Bien que le concept de vie soit discuté par les biologistes épistémologiques [Jacob, 1970], leur objet d'étude s'étend des molécules aux écosystèmes en passant par la cellule. Le système nerveux central (SNC), composé de l'encéphale et de la moelle spinale, ainsi que la maladie d'Alzheimer seront au cœur du présent chapitre.

1.1. Tissu nerveux

Le tissu nerveux est commun aux organismes vivants du sous-règne *bilateria*. Il s'organise en plusieurs systèmes et intègrent deux familles cellulaires : les neurones et les cellules gliales.

Les neurones assurent la transmission de l'influx nerveux. Ils sont composés d'un corps cellulaire, le *soma* ou *péricaryon*, et d'une structure arborescente servant aux connexions nerveuses dont les prolongements cytoplasmiques sont appelés *neurites*. Le soma peut prendre une grande diversité de forme et de taille. Les *neurites* se décomposent en deux catégories : 1) l'axone, qui peut atteindre un mètre de longueur et voie de sortie de l'influx nerveux, est souvent entouré d'une gaine de myéline qui accélère la transmission de l'influx nerveux ; 2) l'*arbre dendritique*, voie d'entrée de l'influx nerveux, il comporte des *dendrites* et permet les connexions avec de nombreux axones issus d'autres neurones via les *synapses*.

Les *synapses* peuvent être tripartites voire quadripartites [Schafer *et al.*, 2013][Rudy *et al.*, 2015]. Dans ces cas, les cellules gliales interstitielles participent à l'activité synaptique en la régulant. Les neurones et les cellules gliales forment ainsi un réseau discret permettant la transmission et le traitement de l'information. Les neurones ne sont cependant pas les seules cellules composant le tissu nerveux du système nerveux central, les cellules gliales représentent quant à elles plus de la moitié des cellules du cerveau humain [Azevedo *et al.*, 2009][Squire *et al.*, 2012].

Il est estimé que le corps cellulaire d'un neurone a un rayon compris entre 2,5 et 15 μm [Andersen *et al.*, 2016]. Ce sont les neurones qui vont nous intéresser dans ce manuscrit.

Les astrocytes sont des cellules étoilées étroitement connectées aux neurones et qui forment l'essentiel du tissu extraneuronal. Elles participent au maintien de la barrière hémato-encéphalique, régulent l'apport en nutriments aux neurones, contribuent à la communication interneurones et maintiennent l'homéostasie de l'encéphale. Elles jouent également un rôle immunitaire et réparateur en cas d'atteinte des neurones, participant à la création des cicatrices gliales.

Les cellules microgliales ont une fonction similaire à celle des macrophages dans le reste du corps. Une fois activées, elles forment la principale défense immunitaire du système nerveux et ont pour rôle de détecter et phagocyter les débris neuronaux et les agents infectieux présents dans le cerveau. Non activées, ces cellules participent aussi à la maturation des synapses au sein des synapses quadripartites [Schafer *et al.*, 2013].

Les oligodendrocytes forment les gaines de myéline qui entourent les axones dans le SNC (ce rôle est joué par les cellules de Schwann dans le système nerveux périphérique [Squire *et al.*, 2012]). Ces gaines accélèrent la vitesse de transmission de l'influx nerveux dans les axones.

1.2. Anatomie et fonction du système nerveux central

Le tissu nerveux est organisé en deux structures principales pour les organismes vivants de l'embranchement des chordés : *l'encéphale* et la *moelle spinale*. L'encéphale est organisé en plusieurs sous-structures appelées régions. Six régions principales sont définies sur la base du développement nerveux à partir du tube neural, le système nerveux primitif des embryons des chordés [Kandel *et al.*, 2000] (Figure 1). Ces régions sont elles-mêmes décomposées en sous-régions, présentant un certain degré de corrélation anatomo-fonctionnelle. Néanmoins, nous n'irons pas à ce niveau de détails dans ce manuscrit, ces corrélations étant encore débattues [Rakic, 1988][Zilles *et al.*, 1997].

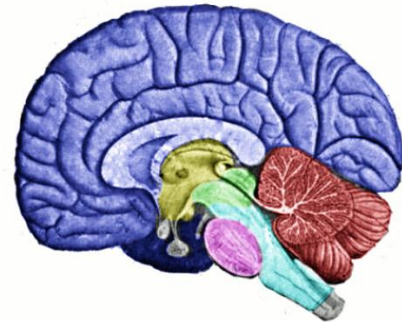
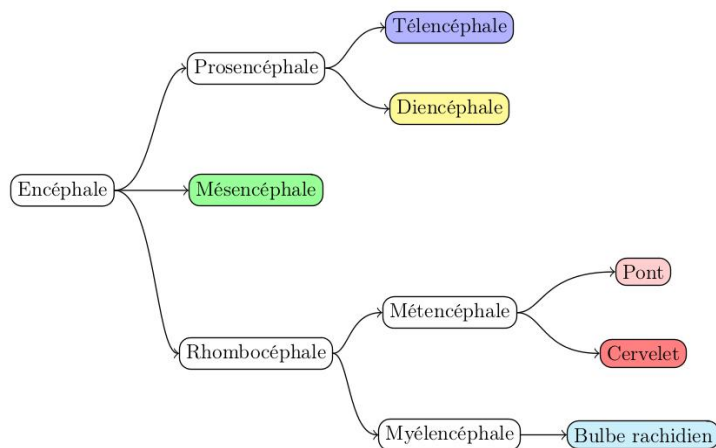


Figure 1 : Division anatomique de l'encéphale (d'après la Figure 1.2 de la thèse de Yaël Balbastre) [Balbastre, 2016].

Le télencéphale est la structure nerveuse la plus développée chez les mammifères. Il regroupe l'intégralité du cortex cérébral qui est lui-même divisé en zones anatomiques comme le *sulci* ou le *gyri* ou en zones fonctionnelles assurant une fonction précise. Les fonctions principales sont les fonctions sensorielles, motrices et associatives [Dictionnaire de l'Académie de médecine, 2018].

Le diencéphale est la structure cérébrale contribuant à la régulation de l'activité du système nerveux autonome (système nerveux responsable des fonctions non soumises au contrôle volontaire), du système hormonal et du sommeil. Le diencéphale est lui-même composé du thalamus, de l'hypothalamus, de l'épithalamus et du sous-thalamus [Dictionnaire de l'Académie de médecine, 2018].

Le cervelet contribue à la coordination et la synchronisation des gestes et à la précision des mouvements. Il régule aussi l'inhibition comportementale en société [Jaber, 2017][Dictionnaire de l'Académie de médecine, 2018].

Le mésencéphale régit les fonctions élémentaires comme l'attention, le sommeil, le réveil et les mouvements de la tête et du cou [Dictionnaire de l'Académie de médecine, 2018].

Le pont relie le cervelet au mésencéphale. Il est impliqué dans les processus de vigilance, d'éveil et de sommeil [Dictionnaire de l'Académie de médecine, 2018].

Enfin, **le myélocéphale** (appelé *bulbe rachidien*) est la région reliant l'encéphale et la moelle spinale. Il régule les fonctions vitales végétatives comme le rythme cardiaque, la respiration et la pression artérielle [Llewellyn-Smith *et al.*, 2011].

1.3. Maladie d'Alzheimer et recherche préclinique

La maladie d'Alzheimer, décrite par Alois Alzheimer en 1906 [Berrios, 1990], est la maladie neurodégénérative la plus répandue dans le monde avec une prévalence de 5% à 8% pour les plus de 60 ans en 2017 [Witthaus *et al.*, 1999][World Health Organisation, 2017]. Elle est aussi une des maladies les plus coûteuses pour la société dans les pays développés [Hay *et al.*, 1987].

Les symptômes de la maladie évoluent au cours du temps [Förstl *et al.*, 1999]. Les stades précoces impliquent des difficultés à entreprendre des tâches complexes et la perte de la mémoire à court terme. Avec le temps, les troubles de la mémoire s'aggravent et les patients deviennent de plus en plus grabataires. L'aphasie (impossibilité de parler), l'apraxie (impossibilité de bouger) et l'agnosie (perte de la reconnaissance des stimuli sensoriels) s'ajoutent à la liste des symptômes. Dans le dernier stade de la maladie, le patient ne peut plus quitter son lit sans aide et se trouve dans un état d'apathie et de fatigue constante. Bien que la maladie d'Alzheimer soit terminale, la cause de la mort est souvent externe (pneumonie, infection liée aux escarres) [Förstl *et al.*, 1999].

Cette maladie est caractérisée à l'échelle du tissu par deux lésions :

Les plaques amyloïdes sont des accumulations extracellulaires de la β -amyloïde, forme clivée de la protéine précurseur de l'amyloïde (APP). Ces accumulations entraînent un dysfonctionnement des neurones environnants, puis la mort neuronale par apoptose⁵ ou par nécrose⁶ [Yankner *et al.*, 1990][Turner *et al.*, 2003][Priller *et al.*, 2006]. Bien que les causes de la formation de ces plaques restent inconnues, une hypothèse étudiée est celle de la propagation-agrégation, mécanisme similaire à celui de la maladie de Creutzfeldt-Jakob [Laurén *et al.*, 2009]. Une des hypothèses les plus récentes est la théorie de la *clearance*, expliquant le surplus de β -amyloïde par un manque de sommeil entraînant un mauvais renouvellement du liquide cébro-spinal, ce procédé étant indispensable au nettoyage des milieux extracellulaires [Xie *et al.*, 2013].

La dégénérescence neuro-fibrillaire : la protéine Tau est une macromolécule essentielle à la stabilité de la *tubuline*, protéine constituant majoritairement l'assemblage des *microtubules* qui forment le cytosquelette des axones. Les protéines tau se positionnent perpendiculairement à l'axone et assurent la rigidité des microtubules et le bon transport axonal. La dégénérescence neuro-fibrillaire est le détachement par hyperphosphorylation de ces

⁵ *Apoptosis*: Processus de mort cellulaire « naturelle », sans réaction inflammatoire, caractérisée par une fragmentation de l'ADN et une condensation de la chromatine. [Dictionnaire de l'Académie de médecine, 2018].

⁶ *Necrosis*: Mortification de cellules ou de tissus privés de leurs apports nutritionnels et vitaux. Mort cellulaire non programmée, conséquence d'un phénomène pathologique entraînant des réactions inflammatoires et la libération d'hydrolases lysosomiales. [Dictionnaire de l'Académie de médecine, 2018].

protéines Tau de la tubuline vers le milieu intracellulaire. L'activité axonale s'en trouve extrêmement dégradée, empêchant le fonctionnement normal du neurone [Hernández et Avila, 2007].

Ces deux lésions provoquent au niveau de l'encéphale une atrophie et une inflammation du tissu nerveux. C'est cette atrophie qui induit la majorité des symptômes de la maladie [Yankner *et al.*, 1990][Turner *et al.*, 2003][Priller *et al.*, 2006][Hernández & Avila, 2007].

Dans la recherche translationnelle, la **recherche préclinique** se positionne entre la recherche biologique fondamentale à l'échelle des cellules et la recherche clinique consacrée au diagnostic et à l'étude de nouveaux traitements pour l'homme [Jucker, 2010]. Les objets d'étude principaux en recherche préclinique sont les modèles animaux. Un **modèle animal** est un animal non humain ayant développé, par modification génétique, par provocation sous la forme de lésions ou naturellement, un état similaire à un état pathologique humain. L'enjeu scientifique de l'expérimentation sur les modèles animaux est de comprendre les mécanismes en jeu et de développer des thérapies pour améliorer leur condition. Une fois validée sur les modèles animaux, une thérapie est ensuite testée en recherche clinique. Les rongeurs et les lapins font partie des espèces les plus utilisées en tant que modèle animal à hauteur de 80% [Report from the commission to the council and the european parliament, 2010]. À MIRCen, les modèles animaux utilisés sont les rongeurs (souris et rats) et les primates non-humains (microcèbes et macaques). Toutes les expérimentations sont soumises aux normes et législations européennes visant la réduction du nombre d'animaux sacrifiés à but de recherche (Directive 2010/63/EU).

2. Imagerie cérébrale

*L'imagerie regroupe les moyens d'acquisition et de reconstruction d'images d'organismes à partir de phénomènes physiques. Le principe est d'observer une information biologique complexe de manière intelligible. Certaines informations peuvent être acquises uniquement pendant la vie de l'organisme, on la qualifie alors d'information in-vivo, réalisée le plus souvent avec des animaux sous anesthésie. Si l'acquisition a lieu après la mort de l'animal (par histologie par exemple), l'information est alors dite acquise post-mortem. Des méthodes intermédiaires permettant la microscopie in vivo existent, comme le NIRF [Gakuba *et al.*, 2018] ou le FMT-XCT [Ale *et al.*, 2012], mais elles ne seront pas abordées dans ce manuscrit. Après une description de quelques techniques d'imagerie in-vivo, le chapitre se concentrera sur les techniques de microscopie post-mortem qui sont au centre de mon travail de recherche.*

2.1. Imagerie *in-vivo*

2.1.1. Tomographie par Émission de Positons

La tomographie par émission de positons (TEP) est une technique d'imagerie utilisant des radiotraceurs nucléaires. Elle consiste à injecter un traceur ayant une cible moléculaire qui va lui conférer une spécificité de fixation et de biodistribution (par exemple : le métabolisme du glucose et la perfusion). Ce traceur est marqué par un atome radioactif qui émet des positons dont l'annihilation avec un électron produit deux photons émis à 180°. La détection en coïncidence de ces photons permet la localisation du lieu d'émission et donc de la position du traceur (Figure 2).

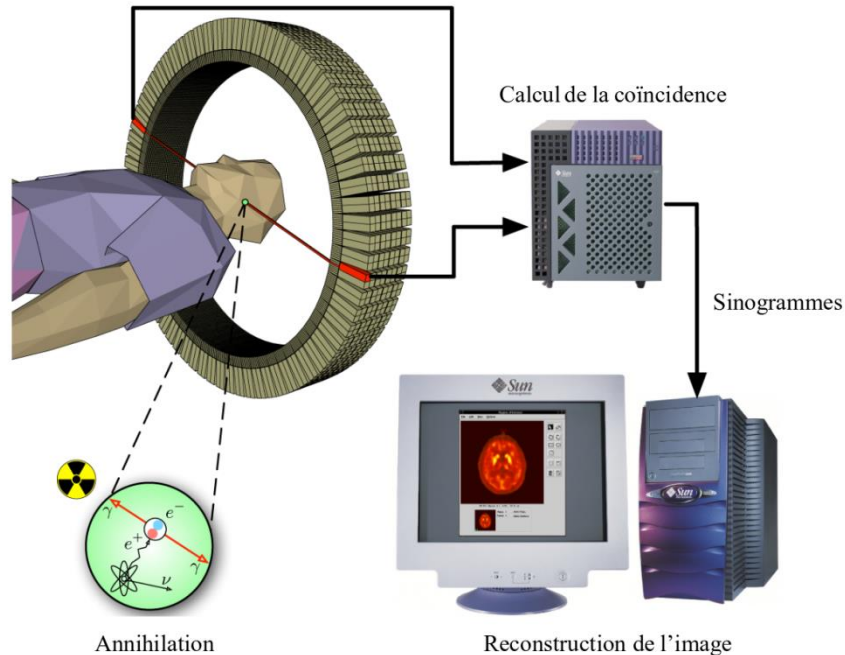


Figure 2 : Principe de reconstruction de l'image TEP. Une fois la coïncidence détectée, elle est interprétée sous forme d'images projetées, des sinogrammes, puis reconstruites en 3D (image traduite et tirée de Wikipedia).

La TEP a été développée dès le début des années 50 pour la détection des tumeurs cérébrales [W. H. Sweet, 1951]. La multiplication des radiotraceurs dans les années 70 a popularisé cette technique d'imagerie. La TEP est souvent utilisée pour quantifier une activité métabolique et, donc, visualiser une activité fonctionnelle. Ainsi, la TEP permet d'étudier l'activité du cerveau lors d'activités spécifiques, comme le sommeil [Braun *et al.*, 1997], de mesurer l'efficacité d'une thérapie [Wohne *et al.*, 2003] ou encore l'évolution d'une maladie [Young *et al.*, 1985]. Le temps d'acquisition est de l'ordre de la dizaine de minutes et les images acquises ont une résolution spatiale de l'ordre du millimètre. La TEP n'est pas invasive (c'est à dire ne nécessitant pas une opération chirurgicale) mais irradie le sujet.

2.1.2. Imagerie par Résonance Magnétique

L'imagerie par résonance magnétique (IRM) utilise le plus souvent la résonance magnétique des atomes d'hydrogène présents dans les molécules d'eau composant le corps humain. Un aimant avec une force magnétique comprise entre 0,2 et 17 Teslas va d'abord aligner les spins des molécules paramagnétiques. Puis grâce à une antenne radiofréquence, une impulsion électromagnétique va exciter les atomes d'hydrogène passant alors dans un état de haute énergie. Le temps d'excitation représente le temps de changement d'état de spin de l'atome d'hydrogène, le spin s'alignant avec le champ électromagnétique de l'aimant. Quand les impulsions sont stoppées, le spin retourne à sa position d'équilibre en émettant une onde électromagnétique. Des antennes captent ces ondes ainsi que leurs caractéristiques. Le temps de retour à l'équilibre du spin de l'atome d'hydrogène est appelé temps de relaxation. Ces temps d'excitation et de relaxation sont spécifiques au type de tissu, et les différences locales de ces ondes électromagnétiques captées permettent de caractériser localement les tissus. Une étape de reconstruction permet de restituer une image tridimensionnelle. Une image utilisant le contraste issu du temps d'excitation est appelé une image en pondération T1, celle utilisant le contraste issu du temps de relaxation est appelé une image en pondération T2 (Figure 3). Les

pondérations T1 et T2 sont des exemples parmi plusieurs séquences IRM existantes. Le choix de la séquence est motivé par l'objet biologique à observer et le contexte de l'observation [Knight *et al.*, 2016].

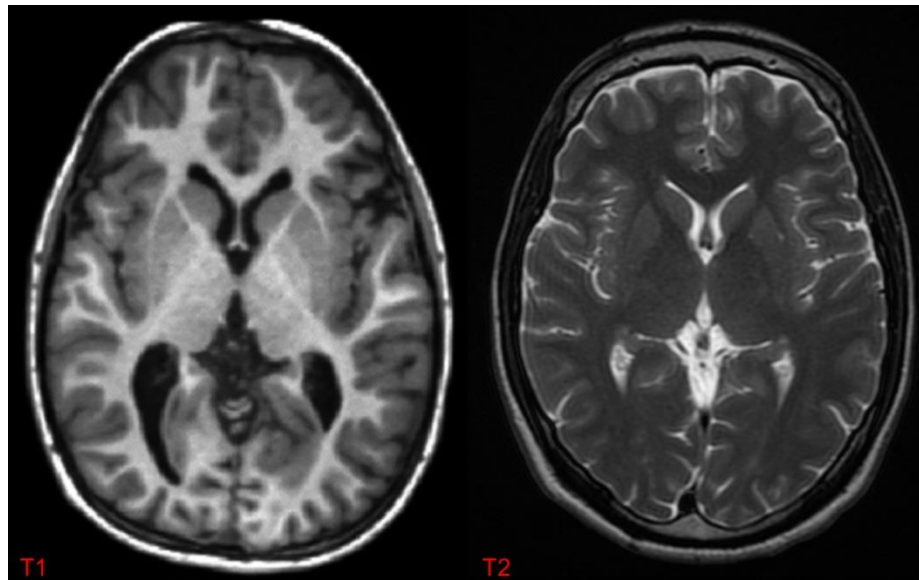


Figure 3 : Différence entre une image produite en pondération T1 et en pondération T2. Les deux images représentent un plan axial d'un cerveau humain.

L'origine de l'IRM dans l'imagerie biomédicale remonte à 1969 avec l'étude de Damadian pour détecter des tumeurs implantées dans des modèles murins (rongeurs) [Damadian, 1971]. Depuis, plusieurs études ont permis la visualisation, entre autre, de l'oxygénation du sang par effet BOLD (permettant de mesurer les différence d'oxygénation du sang) [Ogawa *et al.*, 1990], des fibres de matière blanche [Le Bihan *et al.*, 2001] ou des marqueurs pathologiques de la maladie d'Alzheimer [Jack *et al.*, 2004]. Cet examen n'est pas invasif et les images acquises ont habituellement une résolution spatiale isotrope comprise entre quelques dizaines de microns et le millimètre, la résolution dépendant de plusieurs paramètres (type d'examen, appareillage, gradients, temps d'acquisition ou encore séquence utilisée).

2.2. Imagerie *post-mortem*

2.2.1. Protocoles de production des tissus

L'histologie est la science de l'étude des tissus biologiques afin de comprendre leur organisation, leur fonctionnement et leur dysfonctionnement (histopathologie). Elle repose sur un ensemble de techniques *post-mortem* permettant de visualiser et de quantifier différents types d'objets biologiques. Pour préparer les tissus, plusieurs techniques ont été développées afin de rendre les observations plus proches de ce qu'il se passe dans le vivant. Il existe quatre grandes étapes traditionnellement pratiquées pour réaliser cette tâche [Gartner, 2015] :

- **Le prélèvement** consiste à récupérer la portion de tissu à étudier, soit par biopsie, soit par dissection du sujet, soit par frottis.
- **La fixation** par traitement chimique ou par congélation des tissus permet de retarder ses altérations dues à la mort cellulaire et de figer ses constituantes cellulaires dans un état stable. Les fixateurs les plus courants en histologie sont

la formaline, la paraformaldéhyde (PFA) et le liquide de Bouin. Ces deux produits précipitent et polymérisent les protéines et bloquent les réactions enzymatiques.

- **La coupe** est réalisée par des machines pouvant effectuer des sections de quelques micromètres à quelques dizaines de micromètres, comme un microtome, un vibratome ou un cryostat. Pour la réalisation de coupes très fines (quelques micromètres), il peut être nécessaire d'inclure au préalable des échantillons dans de la paraffine. Les coupes sont ensuite montées sur des lames de verre.
- La plupart des tissus de l'organisme ne sont pas naturellement colorés, ce qui rend leur observation difficile. **La coloration** permet donc de marquer spécifiquement différents objets biologiques. Il existe trois grandes familles de marquages : l'histochimie (par précipitation chimique), l'historadiographie (par réaction nucléaire) et l'immunohistochimie (par réaction immunitaire). Ces marqueurs sont couplés avec des colorants (chromogènes) ou des fluorophores pour être visibles soit à la lumière visible (**bright field** ou en français **champ clair**) soit par excitation des fluorophores (fluorescence).

À partir d'un échantillon biologique, une grande quantité de coupes histologiques peut être produite. Cela permet de réaliser plusieurs marquages en prélevant des séries entrelacées. La colocalisation de marqueurs physiologiques et pathologiques est ainsi rendue possible. À MIRCen, le nombre de séries marquées est très variable. Il est considéré qu'il y a colocalisation entre ces coupes car elles sont très proches. L'intérêt de la production sériée peut être de trouver des corrélations de localisation entre des marqueurs. Ces corrélations permettent de mieux comprendre les phénomènes biologiques associés.

Un grand nombre de marqueurs sont accessibles en histologie, le plus connu étant probablement l'*Hématoxyline et Éosine* (H&E) largement utilisé en oncologie [Ross *et al.*, 2006]. L'hématoxyline, spécifique aux acides nucléiques, colore les noyaux en violet. L'éosine, spécifique aux fibres de collagènes, colore le cytoplasme en rose. Ce marqueur n'est donc pas spécifique à un type cellulaire particulier. Une très grande quantité d'autres marqueurs histologiques a été développée pour ces cas particuliers. Pour le système nerveux central, un des plus utilisés est le marqueur *Neuronal Nuclei* (NeuN). NeuN est un anticorps spécifique aux noyaux des neurones. La Figure 4 illustre une section de cerveau de macaque marquée avec l'anticorps NeuN et numérisée à haute résolution (grossissement x20).

Au contraire de NeuN, l'anticorps anti-Phox2B marque des neurones spécifiques dans le système nerveux central. Ces neurones contrôlent les fonctions végétatives et sont responsables de la respiration automatique. NeuN comme anti-Phox2B sont des **marqueurs physiologiques**. L'anticorps BAM-10, quant à lui, est un marqueur de la β -amyloïde permettant de mettre en évidence les plaques amyloïdes une lésion caractéristique de la maladie d'Alzheimer. L'objet biologique marqué n'est donc pas une cellule mais un agrégat de protéines. Ce type de marqueur est nommé **marqueur pathologique**. Les trois marqueurs présentés ci-dessus sont des marqueurs immunohistochimiques et sont quelques exemples d'une partie des marqueurs utilisés à MIRCen. Néanmoins, ces quelques exemples ne constituent en aucun cas un échantillonnage représentatif des marqueurs existants ni même de ceux utilisés à MIRCen.

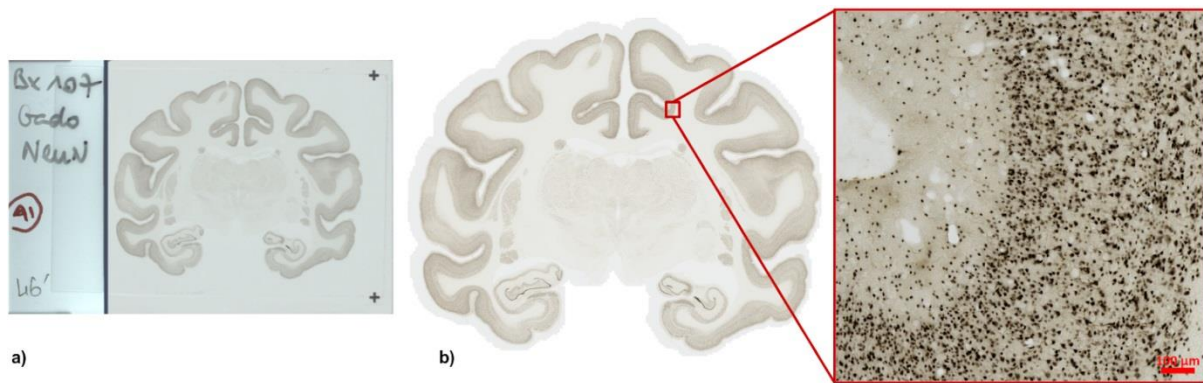


Figure 4 : **a)** Lame de verre portant une section de cerveau de macaque marquée au DAB-Ni NeuN avant numérisation. **b)** Section numérisée à haute résolution avec le scanner de lames Axio Scan Z.1 (grossissement x20). Le grossissement à droite montre la diversité de forme et de distribution des neurones marqués dans le cortex cérébral du macaque.

Récemment, plusieurs protocoles ont permis de produire et d'analyser des organes *post mortem* directement en trois dimensions sans avoir recours à la coupe. On peut notamment citer l'**acquisition par rayons X** [Vågberg *et al.*, 2018] ou les techniques de **clarification** qui ont suscités un regain d'intérêt ces dernières années avec la mise au point par exemple de CLARITY [Chung *et al.*, 2013] ou de iDISCO [Renier *et al.*, 2014]. Les procédés de clarification emprisonnent le tissu dans une matrice d'hydrogel perméable aux lipides, molécules hydrophobe opaques à la lumière visible. La suite du traitement consiste à se débarrasser uniquement des lipides par électrolyse, rendant l'organe transparent à la lumière visible. Ces organes peuvent éventuellement être marqués par les mêmes anticorps que ceux utilisés en histologie classique, créant une véritable continuité avec la technique décrite précédemment. Cependant, le nombre de marqueurs fonctionnels validés avec ces techniques reste encore très limité. Contrairement à l'histologie classique, les marqueurs ont du mal à pénétrer et à se distribuer uniformément dans le tissu [Chung *et al.*, 2013]. Ceci ne facilite pas l'étude de certaines structures internes comme l'hippocampe, structures pourtant essentielles dans l'étude de la maladie d'Alzheimer. De plus, il est impossible de clarifier des organes de grande taille, la taille maximale traitable étant de l'ordre du cerveau de la souris adulte (soit environ 1 cm³).

2.2.2. Imagerie de sections histologiques entières

Les coupes produites par les protocoles d'histologie classiques sont ensuite visualisées avec un microscope optique ou numérisées afin de pouvoir être stockées et analysées, les tissus histologiques subissant une péremption au fil des décades (contamination bactérienne, mycoses, perte d'intégrité du tissu). Pour numériser des sections entières, MIRCen est équipé d'un scanner de lames virtuelles Axio Scan Z.1 (<https://www.zeiss.com/>) capable d'acquérir ces sections entières avec une résolution planaire inférieure au micromètre. À haute-résolution (grossissement x20, 0,22x0,22 µm²), la numérisation peut prendre de plusieurs heures à quelques jours pour 50 sections de cerveau de macaque ou 800 sections de cerveau de souris (ce qui correspond à un demi cerveau de macaque et à trois cerveaux de souris pour un marquage).

Dans le cas des protocoles de clarification, un **microscope à feuille de lumière** numérise le volume en trois dimensions en quelques minutes jusqu'à plusieurs heures en fonction des réglages. Le traitement et la préparation des tissus prend au moins deux semaines

(de 8 à 18 jours [Renier *et al.*, 2014]), ce temps variant en fonction de la taille de l'organe à clarifier et des marqueurs utilisés.

Dans le cas des études *post-mortem*, la visualisation ou la quantification d'un volume en trois dimensions (3D) permet d'obtenir des informations plus précises quant à la localisation spécifique d'un marqueur. Dans le cas de l'imagerie cérébrale, la reconstruction *post-mortem* 3D offre la possibilité d'étudier le cerveau dans son intégralité de les comparer avec les volumes produits *in vivo*. Cette approche peut être utilisée pour valider de nouvelles techniques d'imagerie (instrumentation, séquences IRM, produits de contraste) ou encore quantifier précisément la distribution d'un marqueur voire de corréler plusieurs marqueurs d'intérêts sur une région ou l'ensemble du cerveau.

Pour traiter en trois dimensions des organes de grandes tailles comme un cerveau de macaque (5x6x8 cm), des séries de sections histologiques doivent être empilées et reconstruites en trois dimensions. Cette étape peut être réalisée pour un ou plusieurs marqueurs rendant possibles une colocalisation des marqueurs sur un volume entier à haute résolution. Des protocoles spécifiques ont été développés et sont fréquemment utilisés à MIRCen pour mener diverses études biologiques [Dauguet *et al.*, 2007][Dubois *et al.*, 2010][Vandenberghe *et al.*, 2016].

Le **recalage** est l'étape de traitement de l'image centrale permettant de réaliser la reconstruction 3D tout en compensant la déformation des tissus due aux procédés de production d'une section histologique. Pour mener à bien cette tâche, il est possible d'aligner les coupes vers une référence géométrique soit acquise *in vivo* (comme une IRM [Malandain *et al.*, 2004]) soit acquise juste avant la coupe des sections (comme la reconstruction d'une référence photographique **blockface**) [Dauguet *et al.*, 2007]. Les déformations peuvent être dues à des contractions, à la coloration, à la fixation, à des distorsions mécaniques des tissus dues au prélèvement ou à la coupe. Le recalage à l'aide des photographies prises pré-coupe est la technique la plus utilisée à MIRCen du fait de sa très grande efficacité et précision. Les photographies sont prises avant la coupe de chaque section de tissu. En fin d'expérience, les photographies sont empilées produisant ainsi un volume 3D cohérent qui sera utilisé pour reconstruire la série de coupes histologiques qui seront produites dans un second temps.

La méthode de correspondance par bloc (*Block-Matching*) est utilisée pour recaler les sections à aligner avec leurs photographies correspondantes [Ourselin *et al.*, 2001][Dauguet *et al.*, 2007]. Des blocs de taille fixée sont définis dans les deux images puis une corrélation est calculée pour chaque paire de blocs entre les deux images, ce qui génère une carte de corrélation et un champ de déformation.

Une transformation rigide optimale ayant trois paramètres (un paramètre de rotation et deux paramètres de translation) est ensuite calculée entre la photographie de référence et sa section histologique. Cette transformation optimale maximise la fonction de coût de la corrélation entre la section transformée et la photographie correspondante.

À partir du champ de déformation, une transformation affine est ensuite estimée en utilisant l'estimation rigide comme initialisation. Une transformation affine a cinq paramètres, un paramètre de rotation, deux de translation, un de dilatation et un de cisaillement. La transformation affine optimale est la transformation minimisant la perte quadratique entre deux blocs corrélés.

Ce protocole parvient, dans la majorité des cas, à corriger les principales déformations des sections histologiques et à reconstruire l'organe en trois dimensions. Les seules déformations qui ne sont pas correctement corrigées sont les déformations non-linéaires dues à, par exemple, des déchirures ou des repliements.

2.3. Quantité de données numérisées

Dans la suite de ce manuscrit, nous parlerons de quantité de données en termes de nombre d'éléments. Pour une image bidimensionnelle, nous parlerons de pixels. En effet, en fonction des besoins ou de l'étape de traitement, un pixel peut représenter un entier naturel, un entier relatif ou un nombre à virgule flottante. Ces représentations vont prendre plus ou moins de place en termes de mémoire physique, ce qui ne permet pas de représenter la donnée de façon unique. La représentation sous la forme du nombre d'éléments permet de pallier le problème.

La Figure 5 présente les méthodes d'imagerie évoquées précédemment et les met en parallèle en fonction de la résolution spatiale tridimensionnelle en m^3 et de l'ordre de grandeur en nombre de pixels correspondant à un cerveau de macaque et de souris. Un cerveau de souris mesurant moins de 1 cm^3 et celle d'un cerveau de macaque environ 1 dm^3 , chaque puissance de dix multiplie ou divise la quantité de pixels par mille. À $1\text{ }\mu m^3$ isotrope de résolution, un cerveau entier de souris numérisé est une image tridimensionnelle comportant **plusieurs Téra-pixels** (Tp) quant à un cerveau entier de macaque, l'image comporte **plusieurs Péta-pixels** (Pp).

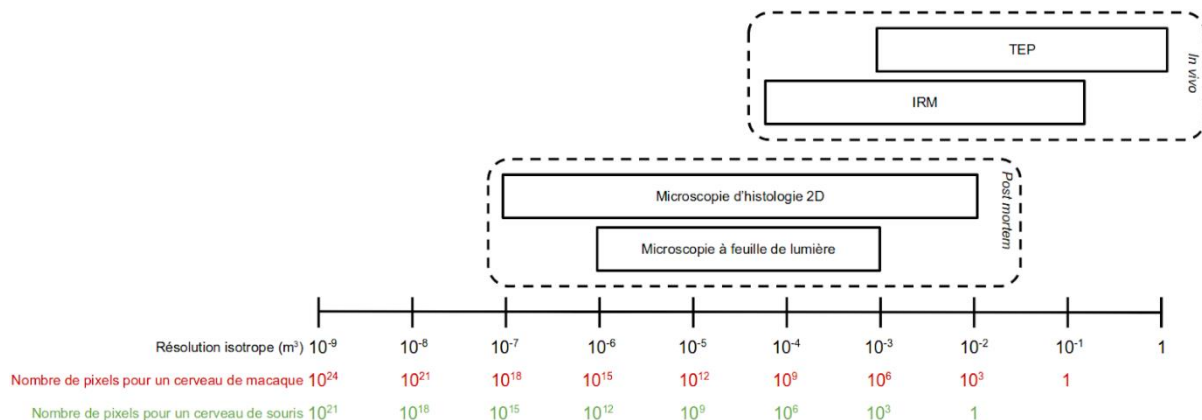


Figure 5 : Synthèse des méthodes d'imagerie présentées en fonction de la résolution isotrope d'imagerie et du nombre de pixels pour un cerveau numérisé à cette résolution.

3. Problématique

L'histologie est généralement considérée comme la modalité d'imagerie de référence et à ce titre est largement utilisée pour comprendre l'organisation *post-mortem* des tissus et le développement des maladies au sein d'un organisme. Dans le cadre de la recherche préclinique sur la maladie d'Alzheimer, l'histologie 2D et 3D sert, par le biais de la quantification des marqueurs pathologiques, à **quantifier l'avancement d'une pathologie, à en comprendre les mécanismes et enfin à évaluer l'efficacité de nouvelles approches thérapeutiques**.

Malheureusement les techniques de quantification pour analyser ce type de données sont limitées (Section 4.2). Les techniques manuelles comme la stéréologie permettent une quantification sur un petit nombre de régions du cerveau mais s'avèrent longues et fastidieuses. Quant aux quantifications automatiques, elles sont coûteuses en temps de calcul ou en quantité de données nécessaires à leur fonctionnement. Dans ce contexte, les segmentations par apprentissage automatique présentent de nombreux avantages par rapport aux autres méthodes de quantification (Section 4.2.3). Malheureusement, le nombre important d'informations intermédiaires nécessaires (caractéristiques) rend ces méthodes difficiles voire

impossibles à utiliser sur des données massives comme des images de cerveaux entiers de modèles animaux, acquises à l'échelle microscopique.

Une solution consiste à trouver un moyen de réduire le nombre de caractéristiques à utiliser pour faire fonctionner les algorithmes d'apprentissage automatique sur des images massives le tout en préservant une bonne qualité de segmentation et donc de quantification. Le pixel pouvant être décrit par un vecteur pouvant comporter jusqu'à plusieurs centaines de caractéristiques (colorimétrie, texture), la **réduction de la dimensionnalité** du problème devient une étape nécessaire à la réalisation d'une segmentation rapide et proche de la segmentation manuelle d'un expert [Kalousis *et al.*, 2005][Chandrashekar *et al.*, 2014]. Les algorithmes de sélection de caractéristiques (Section 6) permettent de réduire un vecteur de caractéristiques initial vers un vecteur de caractéristiques optimal [Kudo *et al.*, 2000][Chandrashekar *et al.*, 2014]. Cependant, la réduction du nombre de caractéristiques a des effets secondaires comme le phénomène de pointe [Kuncheva *et al.*, 2018]. Ce phénomène permet ponctuellement de surpasser la qualité de segmentation initiale avec un vecteur de caractéristiques réduit. Ce phénomène, aussi appelé paradoxe de Hughes, relève au final d'un vecteur optimal localement dans l'espace des caractéristiques [Sima et Dougherty, 2008]. **Cependant, les vecteurs sélectionnés par ces algorithmes sont souvent instables car peu généralisables**, c'est-à-dire que les vecteurs sont trop dépendant de la base d'apprentissage utilisée pour la sélection d'un sous-ensemble de caractéristiques [Kalousis *et al.*, 2005][Kuncheva *et al.*, 2018].

Mon travail de recherche est basé sur des méthodes que j'ai développé afin d'extraire et sélectionner des caractéristiques stables, segmenter une image à l'échelle du pixel et paralléliser ces procédés. Mes principaux objectifs consistaient à :

1. **Déterminer le meilleur algorithme** dans notre cas d'étude afin de segmenter des images histologiques à l'échelle du pixel.
2. **Proposer une méthodologie stable pour réduire le nombre de caractéristiques** utilisées lors de l'étape de classification dans les méthodes d'apprentissage automatique.
3. **Analyser plusieurs marqueurs histologiques** existants et utilisés en recherche préclinique afin de proposer pour chacun d'eux un **vecteur de caractéristiques de taille minimale et produisant une segmentation reproductible et de qualité**.
4. **Réduire les temps de traitement ainsi que la quantité de mémoire** utilisée afin de permettre la réalisation d'une quantification de marqueurs histologiques sur une quantité toujours plus importante de données.

La méthodologie proposée dans mon travail a été appliquée à deux marqueurs histologiques différents. Le marqueur *Neuronal Nuclei* (NeuN) et anti-Phox2B qui sont utilisés pour localiser des neurones spécifiques respectivement dans le cerveau (encéphale) et le tronc cérébral de macaque, un modèle animal proche de l'homme.

II / Etat de l'art

4. Quantification des marqueurs histologiques

L'analyse des données histologiques repose le plus souvent sur la quantification du tissu marqué. Les marqueurs histologiques introduisent des contrastes entre le tissu marqué ou non, mais contrairement à l'imagerie TEP qui permet une mesure quantitative, le contraste observé en histologie dépend quant à lui de la concentration de marqueur utilisée donnant une information qualitative. Des étalonnages sont généralement réalisés au préalable ou reproduits à partir d'autres études pour fournir un contraste satisfaisant. Les images sont alors analysées par un certain nombre de méthodes que nous présenterons dans ce chapitre.

4.1. Stratégies de quantification du tissu marqué

Les marqueurs histologiques révèlent des objets biologiques de différentes natures. En fonction de cette nature et de l'objectif de l'étude, différentes stratégies de quantification peuvent être adoptées.

La **quantification au niveau du pixel** est utilisée pour quantifier des fractions d'occupation pour un marqueur particulier. Cette stratégie est notamment utilisée lorsque le marqueur ne révèle pas des objets biologiques individualisables. Par exemple dans le cadre de l'étude de la maladie d'Alzheimer, l'anticorps BAM-10 révèle des agrégats amyloïdes pathologiques sous la forme d'un marquage dépendant du chromogène utilisé. La fraction d'occupation des plaques est souvent une information suffisante et pertinente pour la quantifier. Bien que la quantification au niveau du pixel puisse être utilisée pour d'autres marqueurs, les **marqueurs pathologiques** sont souvent ciblés par ce type de quantification.

La **quantification des objets** sous la forme d'un dénombrement est une information intéressante à exploiter pour quantifier la perte ou le développement neuronal, par exemple. Les pixels sont alors accumulés sous forme de structures individualisées et dénombrable, ce qui nécessite des informations et méthodologies supplémentaires afin d'individualiser des objets qui se chevauchent ou se touchent [Avenal *et al.*, 2013][You *et al.*, 2016]. Cette stratégie est souvent utilisée pour des **marqueurs physiologiques cellulaires**.

De plus, à l'issue de la phase de quantification, le résultat n'est pas nécessairement un nombre d'objets ou une fraction d'occupation. En recherche clinique, la quantité de tissu marqué n'est pas forcément une information pertinente aux vues de l'évaluation d'un traitement. En oncologie par exemple, une **échelle catégorielle** nommée classification *Tumour-Node-Metastasis* (TNM) [Denoix 1944], sera préférée pour décrire l'avancement de la maladie [Gospodarowicz *et al.*, 2017]. Cette échelle est un score histologique basé sur les quantifications effectuées sur les images visualisées. Dans le domaine des neuropathologies, l'échelle de Braak définit l'avancement de la maladie d'Alzheimer et de Parkinson à partir des données cliniques [Braak et Braak, 1991]. Ces deux échelles bénéficient d'une reconnaissance et d'une standardisation internationale.

Cependant dans le contexte de la recherche préclinique, des **échelles continues ou discrètes** seront préférées car elles dénotent une quantité permettant des comparaisons plus nuancées au sein d'une même étude. Contrairement aux échelles catégorielles, les quantifications d'objets ou de pixels peuvent être directement utilisées.

Aux stratégies de quantification s'ajoutent de nombreuses méthodes qui permettent de réaliser ces quantifications. Dans la Section suivante les principales méthodes utilisées dans les études histologiques seront présentées brièvement.

4.2. Méthodes de quantification de données d'histologie

4.2.1. Méthodes de quantification manuelles

La **segmentation manuelle** est opérée par un expert qui compte ou délimite tous les objets biologiques d'intérêt. Cette méthode est fastidieuse et coûteuse en temps et en énergie (plusieurs jours sont nécessaires pour segmenter une base d'une centaine d'images de 512x512 pixels). Traiter de grands volumes de données ainsi est très difficile voire impossible. Néanmoins, cette approche peut servir à créer des bases d'apprentissage automatique qui seront utilisées dans le cadre de l'apprentissage supervisé (voir Section 3.2.3).

Une méthode très largement utilisée par les biologistes est la **stéréologie**. À l'instar de la segmentation manuelle, un opérateur va compter ou délimiter les objets biologiques d'intérêt mais seulement à certains endroits choisis avec un échantillonnage adapté [West *et al.*, 1991]. Cette quantification va permettre une estimation sur toute la région étudiée sur plusieurs sections. Bien que la stéréologie garantisse que l'estimée (la quantité estimée par la stéréologie) est non-biaisée, elle ne garantit pas que le nombre estimé (la quantité estimée par un opérateur humain) soit correcte.

Les principaux défauts de ces techniques sont qu'elles sont très consommatrices en temps humain, une segmentation manuelle ou une étude par stéréologie pouvant prendre plusieurs jours voire plusieurs semaines. De plus, les résultats de segmentation par plusieurs opérateurs humains peuvent être très différents [Kermany *et al.*, 2018].

4.2.2. Méthodes de quantification semi-automatiques

Les méthodes semi-automatiques sont des méthodes requérant partiellement l'intervention d'un opérateur humain pour fonctionner. Cette intervention est généralement l'estimation empirique d'un ou plusieurs paramètres ou une supervision du processus, permettant de traiter de plus grandes quantités de données par rapport aux méthodes manuelles. La quantification par **densité optique** en est un bon exemple [Bacus *et al.*, 1989]. Cette méthode consiste à mesurer l'absorption des couleurs primaires de la lumière par les marqueurs histologiques utilisés sur la section ou la région d'intérêt (RDI). Pour un marqueur standardisé et une région donnée, il est possible de mesurer une baisse ou une augmentation de la proportion du marqueur.

$$DO_{CP} = -\log_{10}\left(\frac{I_{CP}}{I_{0,CP}}\right) \quad (1)$$

La densité optique DO pour une couleur primaire CP est définie par le logarithme en base 10 du rapport des intensités lumineuses entrante $I_{0,CP}$ et sortante I_{CP} . Pour calculer la quantité de marqueurs, la densité optique $DO = C * M$ doit être déconvoluée avec le vecteur des densités optiques (C) pour chaque marqueur et la matrice de transition (M) calculée à partir de l'absorption des couleurs primaires pour chaque marqueur isolé. DO et M étant connues, la déconvolution consiste à inverser M et à trouver les valeurs d'intérêts C . Les calculs matriciels étant facilement automatisables, ce sont les mesures de DO et de M qui sont réalisées par un

opérateur humain. Cette méthode a été utilisée par exemple en 2017 pour évaluer le taux de marqueur NeuN sur une souris transgénique modèle d'amyloïdose [Brureau *et al.*, 2017].

La méthode la plus utilisée pour quantifier un marqueur en histologie est la segmentation par **seuillage**. Cette méthode s'applique sur une image en niveaux de gris $I(x)$ avec x étant un pixel donné. L'image seuillée I_{se} comporte deux classes :

$$I_{se}(x) = \begin{cases} 1 & \text{si } I(x) \geq S \\ 0 & \text{sinon} \end{cases} \quad (2)$$

S est un seuil choisi arbitrairement par un opérateur humain, il correspond à une valeur comprise entre le niveau de gris minimal et maximal de $I(x)$. Une fois le seuil fixé, l'intégralité de du marquage de la section à segmenter est binarisée. L'opération de seuillage est rapide et simple mais la qualité de segmentation résultante est très dépendante du choix du seuil. Il est efficace pour des marqueurs très spécifiques et contrastés. En revanche, si les sections acquises par microscopie présentent des disparités au niveau de l'illumination d'acquisition en fond clair (**bright field**) ou du procédé de coloration, le seuillage sera biaisé et le résultat de la quantification très éloigné de la réalité. Un deuxième désavantage de cette méthode est que la classe doit être définissable uniquement via son intensité en niveau de gris. L'intégration d'une information multidimensionnelle intégrant colorimétrie et texture n'est pas possible alors qu'elle peut être importante pour l'identification automatique du degré TMN d'une tumeur par exemple [Doyle *et al.*, 2008][Tambasco *et al.*, 2008][Basavanhally *et al.*, 2013]. Enfin, une variabilité inter- voire intra-opérateur non négligeable peut être facilement introduite.

4.2.3. Méthodes de quantification automatiques

Le seuillage peut aussi être réalisé de **manière automatique**. De multiples algorithmes de **seuillage adaptatif** existent. L'algorithme d'Otsu est souvent utilisé en segmentation en optimisant le seuil à appliquer sur un histogramme de manière à minimiser les variances intra-classes [Otsu, 1975]. D'autres algorithmes optimisent leur seuil par d'autres approches, par l'algorithme d'Estimation-Maximisation (EM [Dempster *et al.*, 1977]) par exemple pour la segmentation de plaques amyloïdes [Feki *et al.*, 2007]. Le seuillage adaptatif possède les mêmes défauts que le seuillage classique. Bien qu'il atténue les effets de disparité entre images à traiter le seuil est optimisé pour une image et généralisé aux autres ou alors pour chaque image. Le temps de traitement est fonction de la taille des images traitées : plus l'image a un nombre important de pixels plus l'histogramme sera long à construire et plus l'optimisation prendra du temps. Si l'image est découpée en plusieurs images plus petites afin de réduire le temps de traitement, des discontinuités peuvent apparaître aux frontières entre les petites images car les différents seuils peuvent être très différents. Ces méthodes ne sont généralement pas utilisées dans le cas d'images de grande taille, c'est à dire avec un grand nombre de pixels.

Dans ce contexte, les méthodes d'apprentissage automatique constituent une voie de plus en plus considérée pour réaliser des quantifications de marqueurs en histologie. Ces méthodes présentent plusieurs avantages comparés au seuillage. Tout d'abord, ces algorithmes intègrent des informations colorimétriques et texturales nommées **caractéristiques** (voir Section 5.3. pour plus de détails) et ils sont applicables sans limitation à **plus de deux classes**.

Les méthodes d'apprentissage automatique suivent trois étapes majeures :

1. l'**extraction de caractéristiques**,
2. l'**étape d'apprentissage** ou d'entraînement adaptant les paramètres de l'algorithme aux données suivant le type d'apprentissage choisi,
3. l'étape de **classification** ou de **généralisation** qui utilise le modèle d'apprentissage généré précédemment pour prédire la classe d'un élément donné.

Selon leur manière d'adapter l'algorithme à la donnée lors des étapes décrites ci-dessus, les méthodes d'apprentissage automatique peuvent être classifiées en trois familles :

L'**apprentissage par renforcement** est surtout utilisé dans le domaine de la robotique. La machine, aussi appelée dans ce contexte **agent**, effectue une action au sein d'un environnement spécifique (par exemple le mouvement d'un robot [Bagnell et Schneider, 2001], l'envoi de mail promotionnels [Abe *et al.*, 2004], jouer un tour de backgammon [Tesauro *et al.*, 1992]) dont les effets sont jugés par un ensemble de règles définies par les chercheurs. En fonction de ces règles, une récompense positive ou négative est donnée à l'agent, modifiant son état. L'objectif de l'agent est de recevoir un maximum de récompenses positives. Cette famille d'apprentissage est inspirée des procédés de décisions Markoviens [Bellman, 1957] mais aussi des procédés de conditionnement béhavioriste classiques en psychologie (aussi appelé conditionnement Pavlovien [Pavlov, 1927]).

Lorsque qu'aucun modèle, qu'aucune règle spécifique ou qu'aucune base de référence n'est connue pour un problème donné, les méthodes d'**apprentissage non-supervisé** sont la catégorie de méthodes à utiliser. Elles regroupent les éléments selon leurs caractéristiques communes et reposent donc sur une métrique permettant de comparer les caractéristiques des éléments à classer entre elles. Ces méthodes sont souvent utilisées en génomique pour regrouper des séquences d'ARN ou d'ADN ensemble, permettant, par exemple, la construction d'arbres phylogénétiques [Nguyen Quang *et al.*, 2018].

Enfin, les méthodes d'**apprentissage supervisé** sont les plus populaires dans la classification d'images car elles reposent sur de larges bases segmentées ou classifiées manuellement par des experts. Ces bases sont appelées **base d'apprentissage** et l'algorithme va exploiter les données décrites dans ces bases pour créer un modèle qui lui permettra de réaliser la tâche de classification. Dans le cas de l'histologie, une base d'apprentissage est une collection d'images segmentées manuellement. De cette façon, l'algorithme intègre la connaissance des experts à travers l'étape d'apprentissage. Les principales méthodes d'apprentissage supervisé utilisées en histologie sont la machine à support de vecteur (MSV) [Cruz-Roa *et al.*, 2011], les modèles de mélange de gaussiennes (MMG) [Chubb *et al.*, 2006], les forêts aléatoires (FA) [DiFranco *et al.*, 2011] et l'apprentissage profond (DL) [Ciresan *et al.*, 2013]. Etant donné que nous favorisons l'utilisation de cette dernière catégorie, les méthodes d'apprentissage supervisé seront développées en détail dans la Section suivante.

5. Méthodes d'apprentissage automatique supervisée

La segmentation automatique permet de réaliser des quantifications fines et exhaustives des tissus marqués [Schmitt et al., 2008][Roullier et al., 2013]. Ce chapitre est consacré aux techniques d'apprentissage automatique supervisé et à l'extraction des caractéristiques à partir d'images histologiques.

5.1. Evaluation et validation d'une segmentation

La segmentation pixel à pixel permet, comme expliqué dans la Section 4.1., de quantifier précisément la présence d'un marqueur dans une image histologique. La Figure 6 présente la chaîne de traitement globale de la segmentation pixel à pixel utilisant les méthodes d'apprentissage automatique. Chaque pixel est décrit par un vecteur de caractéristiques qui va permettre au classifieur d'inférer sa classe.

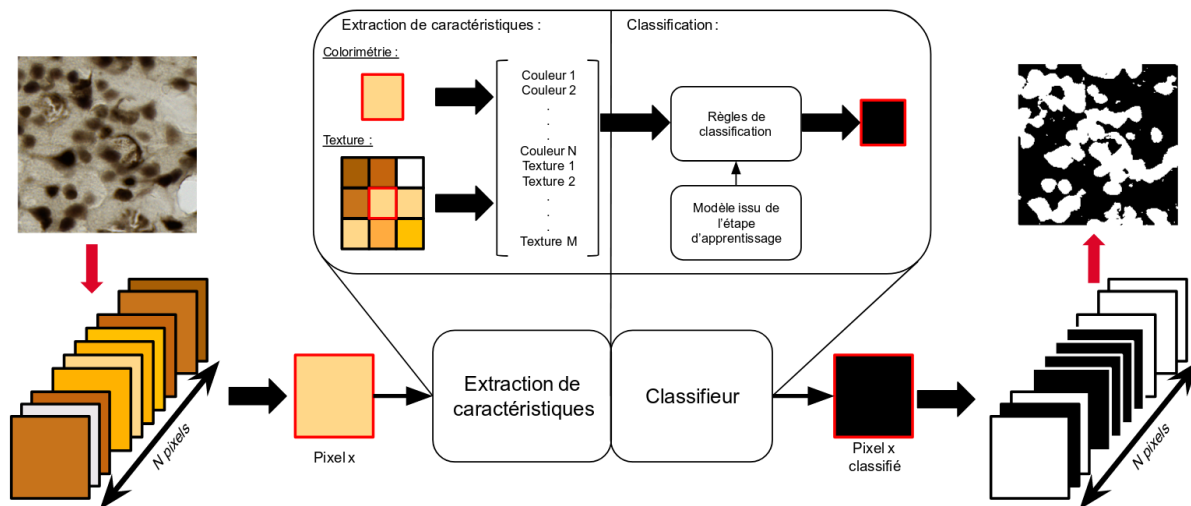


Figure 6 : Procédé de segmentation par les méthodes d'apprentissage automatique pixel à pixel. L'image d'entrée est un exemple d'un tissu numérisé à haute-résolution et marqué avec l'anticorps NeuN, les pixels bruns représentent les pixels de neurones marqués. L'image de sortie est un exemple de segmentation binaire où les pixels blancs sont les pixels détectés comme des pixels de neurones par le procédé de segmentation pixel à pixel.

Afin de mesurer la qualité d'une segmentation, plusieurs métriques ont été proposées pour comparer une segmentation automatique à la segmentation manuelle de référence. La base utilisée pour faire cette évaluation est appelé **base de test**. Plus la segmentation automatique sera proche de la segmentation manuelle, plus la segmentation automatique sera considérée comme de bonne qualité. Selon Taha et Hanbury, parmi les métriques développées pour évaluer les segmentations d'images biomédicales, le F-Score est la métrique la plus appropriée pour une segmentation ayant les propriétés suivantes [Taha et Hanbury, 2015] :

- L'évaluation est basée sur la superposition de la segmentation testée et de la segmentation de référence.
- Les segmentations ne sont pas supposées être parfaites, des valeurs aberrantes (*outliers*) sont tolérées.

La superposition entre les segmentations automatiques et manuelles est la manière la plus naturelle d'évaluer une segmentation automatique. De plus elle permet de conserver la spatialisation du marqueur, la forme et l'organisation des tissus pouvant être des informations importantes pour la suite de l'analyse de ces images [You *et al.*, 2016]. De plus, une segmentation parfaite est impossible à réaliser sur une image de plusieurs Téra-pixels.

Le **F-Score** (aussi appelé critère de Dice ou F1 Score) repose sur le calcul de la **précision et de la spécificité**, elles-mêmes reposant sur **les erreurs statistiques de type I et de type II** [Powers, 2011]. Lors d'une évaluation statistique, une prédiction peut se trouver dans quatre cas différents :

- La prédiction est conforme à la classe d'intérêt et la prédiction est vraie selon la référence. Dans le cas d'une segmentation du tissu marqué d'une section histologique, la classe "tissu marqué" a été prédite correctement. La prédiction est donc juste et ces éléments sont des **vrais positifs**.
- La prédiction est conforme à la classe d'intérêt et la prédiction est fausse selon la référence. Dans le cas d'une segmentation du tissu marqué d'une section histologique, la classe "tissu marqué" a été prédite alors que la classe "tissu non marqué" était attendue. La prédiction présente une erreur, cette erreur s'appelle **l'erreur de type I** et ces éléments sont des **faux positifs**.
- La prédiction n'est pas conforme à la classe d'intérêt et la prédiction est fausse selon la référence. Dans le cas d'une segmentation du tissu marqué d'une section histologique, la classe "tissu non-marqué" a été prédite alors que la classe "tissu marqué" était attendue. La prédiction présente une erreur, cette erreur s'appelle **erreur de type II** et ces éléments sont des **faux négatifs**.
- La prédiction n'est pas conforme à la classe d'intérêt et la prédiction est vraie selon la référence. Dans le cas d'une segmentation du tissu marqué d'une section histologique, la classe "tissu non marqué" a été prédite correctement. La prédiction est donc juste et ces éléments sont des **vrais négatifs**.

L'objectif est de minimiser les erreurs de prédiction afin de maximiser la précision (la proportion de faux positifs dans les prédictions) et la spécificité (la proportion de faux négatifs dans les éléments à prédire comme positifs). Pour mesurer si ce test est précis et spécifique, la **moyenne harmonique** de la précision et de la spécificité est calculée. Cette valeur est appelé F-Score.

$$F - Score = \frac{2*Précision*Spécificité}{Précision+Spécificité} \quad (3)$$

Le F-Score **évalue** quantitativement une segmentation par rapport à une référence, ici une segmentation manuelle. Un F-Score de 0 signifie qu'aucun des pixels n'est classifié dans la bonne classe et un F-Score de 1 signifie que la segmentation automatique est identique à la manuelle. On considère généralement qu'une segmentation automatique avec F-Score supérieur à 0,7 est convenable.

D'autre part, afin de **valider** une segmentation, il est important de montrer que sa qualité n'est pas spécifique à la constitution de la base d'apprentissage. La validation croisée est une méthodologie permettant de savoir si le modèle prédictif calculé est généralisable. Pour cela une **base de test** est définie, les images contenues dans la base de test n'étant pas utilisées dans

l'étape d'apprentissage. Deux grandes familles de validation croisée existent, tout d'abord les validations croisées exhaustives testant toutes les combinaisons possibles d'éléments tirés aléatoirement de la base d'apprentissage et de la base de test, et les validations croisées non exhaustives limitant le nombre de combinaisons testées.

En biologie, il est difficile de créer un ensemble base d'apprentissage - base de test représentant tous les cas possibles. La quantité de données à explorer est souvent très importante et complexe (Figure 5). De ce fait, les bases générées sont biaisées, car non exhaustives. Dans le cadre de ce travail, les bases ont été conçues pour être équilibrées en termes de quantité et d'intensité de marqueur. À l'échelle des pixels, les représentations des différentes classes sont comparables entre les deux bases (voir Section 7). Cependant, à l'échelle des images, les bases peuvent être déséquilibrées et chaque image avoir une proportion différente de tissu marqué (voir Section 7). Pour ces raisons, une validation croisée non exhaustive a été choisie. Ainsi la validation de la segmentation ne dépend pas des images constituant les bases mais de la proportion de tissu marqué. Sur la validation croisée réalisée en deux étapes, la première étape est nommée **validation directe** et la deuxième, inversant la base d'apprentissage et de test, la **validation croisée**.

5.2. Méthodes d'apprentissage automatique classiques en histologie

5.2.1. Machine à support de vecteurs

Les machines à support de vecteur (MSV) ont été introduites par Boser, Guyon et Vapnik en 1992 à partir de la théorie de Vapnik-Chervonenkis [Boser *et al.*, 1992]. Cette méthode est un classifieur linéaire reposant sur deux idées fondamentales :

- Dans une classification à deux classes, la marge est la distance entre la frontière de séparation des classes et de l'échantillon le plus proche. La frontière idéale selon la théorie de Vapnik-Chervonenkis est une frontière **maximisant la marge**.
- Lorsque le problème n'est pas linéairement séparable, c'est à dire ne tolérant pas d'hyperplan en tant que frontière, une idée est de transformer l'espace de représentation (l'espace de caractéristiques) en **un espace de plus grande dimension**.

Soit $s = h(x) = w^T x + w_0$ avec $s \in \{-1, 1\}$ la classe prédite, h la fonction de prédiction, $x = (c_1, c_2, \dots, c_k)$ le vecteur de k caractéristiques et $w = (w_1, w_2, \dots, w_k)$ le vecteur de poids, il est alors décidé que $h(x) = 0$ est l'hyperplan séparateur de l'ensemble (X, Y) avec X l'ensemble des vecteurs de caractéristiques de la base d'apprentissage et Y l'ensemble des classes associées. Si le problème est linéairement séparable, la frontière maximisant la marge est l'ensemble (w_0, w) maximisant la fonction de coût suivante :

$$\frac{1}{\|w\|} \min_i (w^T x_i + w_0) \text{ avec } 0 \leq i \leq k \quad (4)$$

Dans le cas d'un problème qui n'est pas linéairement séparable, une solution consiste à transformer l'espace de caractéristiques en un espace de plus grandes dimensions, éventuellement de dimension infinie, en utilisant une transformation non-linéaire. Cette transformation est généralement déterminée arbitrairement.

La machine à support de vecteurs a été utilisée avec succès pour segmenter des tissus sur des sections histologiques numérisées [Naik *et al.*, 2008][Cruz-Roa *et al.*, 2011][Mutu Rama

Krishnan *et al.*, 2011][Fatima *et al.*, 2017]. Dans ces articles, la taille du vecteur de caractéristiques dépasse la centaine d'éléments. La grande dimension permet d'augmenter la probabilité d'avoir une classification linéairement séparable. Dans notre objectif de minimisation de la taille du vecteur de caractéristique, **la MSV n'est pas un classifieur adapté à notre problématique.**

5.2.2. Modèle de mélange de gaussiennes

Le modèle de mélange de gaussiennes (MMG) est un **modèle probabiliste**. La prédiction n'est pas une classe mais un vecteur de probabilités. Ce vecteur représentant la probabilité, selon le modèle prédictif, d'un élément d'appartenir à une classe donnée. Un exemple de l'utilisation de cette méthode pour la segmentation des tissus histologiques est l'algorithme **BioVision** [Chubb *et al.*, 2006]. BioVision estime la densité de probabilité pour chaque classe $j \in [1, \dots, s]$:

$$f_x(x|y = j) = \sum_{i=1}^k \frac{\rho_i}{(2\pi)^2 \sqrt{|\Sigma_i|}} e^{-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i)} \quad (5)$$

avec, pour chaque caractéristique i du vecteur de caractéristiques x , ρ_i le coefficient de mélange, μ_i la moyenne et Σ_i la matrice de covariance de la distribution gaussienne. Les paramètres sont déterminés par l'étape d'apprentissage. Dans le cas de BioVision, l'algorithme de recuit simulé est utilisé [Geman et Geman, 1984] mais d'autres algorithmes comme l'algorithme d'Estimation-Maximisation (EM) peuvent être utilisés [Moore, 1999]. Lors de l'étape de classification, la probabilité pour un pixel p_i d'appartenir à la classe j est obtenue par le théorème de Bayes :

$$P(y_i = j|p_i) = \frac{P(y_i = j)P(p_i|y=j)}{\sum_{a=1}^S P(x_i|y=a)} \quad (6)$$

Avec $P(y_i = j)$ la proportion des pixels classifiés j dans la base d'apprentissage. La classe finale est la classe qui maximise $P(y_i = j|p_i)$.

Plusieurs paramètres sont disponibles pour affiner l'apprentissage :

- Les nombres maximum et minimum de modèles gaussiens sont utilisés pour modéliser une classe. L'algorithme d'apprentissage ajoute les modèles gaussiens un par un jusqu'à ce que la distribution d'une classe par caractéristique soit approchée avec le mélange de gaussiennes.
- La température de recuit minimale définit la condition d'arrêt de l'apprentissage. La sortie est un modèle enregistré en format texte détaillant l'intégralité des paramètres gaussiens trouvés pour chaque caractéristique et chaque classe.
- Il est aussi possible d'affecter une probabilité *a priori* pour chaque classe, dans le but de réduire le déséquilibre de présence d'une classe dans la base d'apprentissage.

La principale hypothèse de ces méthodes est que les caractéristiques sont modélisables sous la forme d'un mélange de distributions gaussiennes dans l'espace des caractéristiques. Cette hypothèse n'est pas vérifiée avec toutes les caractéristiques notamment avec les caractéristiques de textures mais obtient de bons résultats avec les caractéristiques colorimétriques [Vandenberghe, thèse, 2016]. **Nous allons inclure BioVision dans nos futurs tests.**

5.2.3. Forêt aléatoire

L'algorithme des forêts aléatoires, introduit par Tim Kam Ho et Murray Hill en 1995, est un algorithme d'apprentissage automatique statistique multiclass. Il est l'amélioration de l'algorithme des arbres de décision [Wu *et al.*, 1975][Ho et Hill, 1995].

Soit (x, y) base d'apprentissage avec x un élément et y sa classe associée. L'ensemble d'apprentissage peut être exprimé en fonction des caractéristiques c_i comme $(x, y) = (c_1, c_2, \dots, c_k, y)$. Pour construire un arbre de décision (Figure 7), nous devons préalablement décider du noeud racine (le tout premier embranchement). Il s'agit donc de déterminer x tel que l'entropie de Shannon de la caractéristique c_i soit minimale pour la classe y . Puis, le procédé est réitéré jusqu'à l'épuisement des caractéristiques ou si la profondeur d'arbre qui a été fixée est atteinte (le nombre maximal d'embranchements).

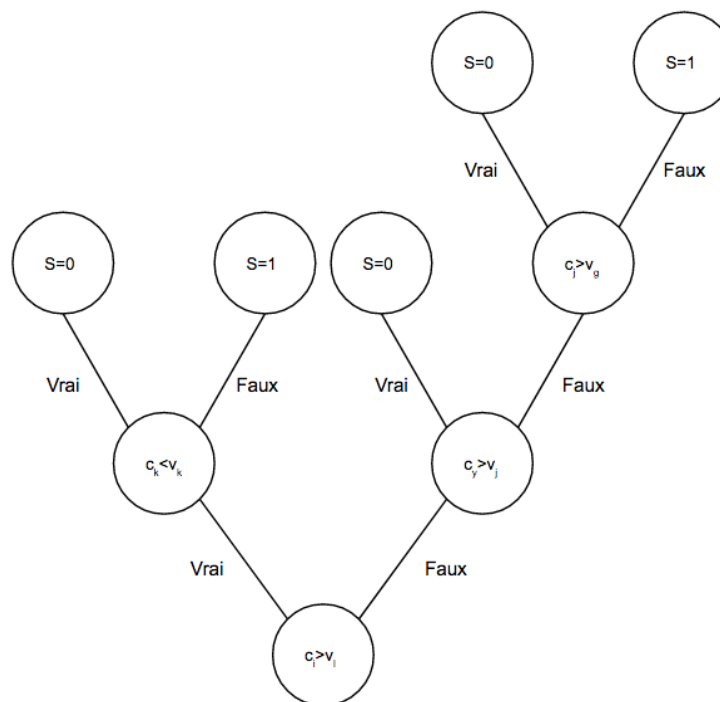


Figure 7 : Exemple d'arbre de décision avec deux classes de sortie. Des tests successifs ont été réalisés jusqu'à arriver à une décision, c'est-à-dire une classe S déterminée. Les valeurs v sont les valeurs déterminées par la minimisation de l'entropie de Shannon.

Une forêt aléatoire est un ensemble d'arbres de décisions entraînés sur une amorce agrégante (*bagging*) de l'ensemble d'apprentissage initial [Breiman, 2001]. Soit un ensemble de données d'apprentissage $X = \{x_1, x_2, \dots, x_n\}$ et l'ensemble de labels associés $Y = \{y_1, y_2, \dots, y_n\}$. Le *bagging* est la sélection aléatoire avec remise de b données d'apprentissage avec leurs labels associés X_b et Y_b . Ce tirage aléatoire est répété b fois et à chaque tirage, un arbre de décision est entraîné sur q caractéristiques tirées aléatoirement sur l'ensemble (X_b, Y_b) . La décision finale S est la décision majoritaire parmi les b arbres de décisions.

Il y a deux principaux paramètres utilisés dans l'apprentissage des forêts aléatoires. Le nombre d'arbres de décision de la forêt aléatoire permet d'augmenter la stabilité de la prédiction. En effet, plus le nombre d'arbres est important plus le résultat du vote majoritaire sera fiable, le nombre de prédictions étant plus grand. La profondeur d'arbre définit la profondeur

d'embranchement de chaque arbre de décision. Plus ce nombre est grand, plus l'espace des caractéristiques sera découpé de manière précise, augmentant généralement le F-Score.

Les forêts aléatoires possèdent la capacité de trier les caractéristiques par importance pour l'apprentissage de la base de données. Cette importance est la moyenne sur tous les arbres de l'importance des caractéristiques pour chaque arbre. L'importance d'une caractéristique pour un arbre donné est proportionnelle à sa position dans l'arbre. Plus l'embranchement la concernant est près de la racine, plus son importance sera grande. L'importance d'une caractéristique sera appelée **facteur d'importance** dans la suite de ce manuscrit. Néanmoins, le principal défaut de cette méthode est qu'elle est sensible à la proportion des classes sur lesquelles les arbres de décision sont entraînés. **Les forêts aléatoires ont tendance à être biaisées et à donner plus d'importance aux classes majoritaires.** Pour pallier ce problème, un système de pondération a été mis en place pour renforcer l'importance des classes minoritaires [Chen *et al.*, 2004]. Cette méthode s'appelle la **forêt aléatoire pondérée**.

Cette méthode a été utilisée pour résoudre un grand nombre de problèmes de segmentation de tissus marqués [DiFranco *et al.*, 2011][Vandenberghe *et al.*, 2015][Kather *et al.*, 2016][Fatima *et al.*, 2017] ou de sélection de caractéristiques [Kooi *et al.*, 2016][Liu, 2017][Yaqub *et al.*, 2018]. **Cet algorithme a été utilisé dans le cadre de ce travail** d'autant plus qu'il avait déjà été testé au sein de notre laboratoire avec des résultats de très bonne qualité.

5.2.4. Apprentissage profond (*Deep Learning*)

Dans le domaine de la reconnaissance de forme et de l'intelligence artificielle, les méthodes d'apprentissage profond sont devenues incontournables, suite à l'obtention dans plusieurs domaines de résultats impressionnants dont, par exemple, le jeu de go [Silver *et al.*, 2016], la génération automatique d'images et l'algèbre de concepts [Radford *et al.*, 2015], la détection de lésion vasculaire [Leenhardt *et al.*, 2018] ou la segmentation de tissu marqué [Ciresan *et al.*, 2013][Van Eycke *et al.*, 2018]. Bien que ces méthodes puissent paraître comme étant en dehors de la définition classique des méthodes d'apprentissage automatique, elles suivent exactement le même schéma procédural décrit Section 5.1 Figure 6. Les caractéristiques utilisées sont cependant souvent induites par la méthode elle-même [LeCun *et al.*, 2015]. Un algorithme d'apprentissage profond concatène plusieurs étapes de traitements composées de **neurones artificiels** (Figure 8) et parfois de pré- et post-traitements [Barat et Ducottet, 2016]. Les neurones artificiels, introduits en informatique par Rosenblatt en 1957 [Rosenblatt, 1957], contiennent des fonctions mathématiques pouvant aller de la convolution à la rectification linéaire dans le cas des réseaux convolutifs, particulièrement utilisés en segmentation d'images [LeCun *et al.*, 2015].

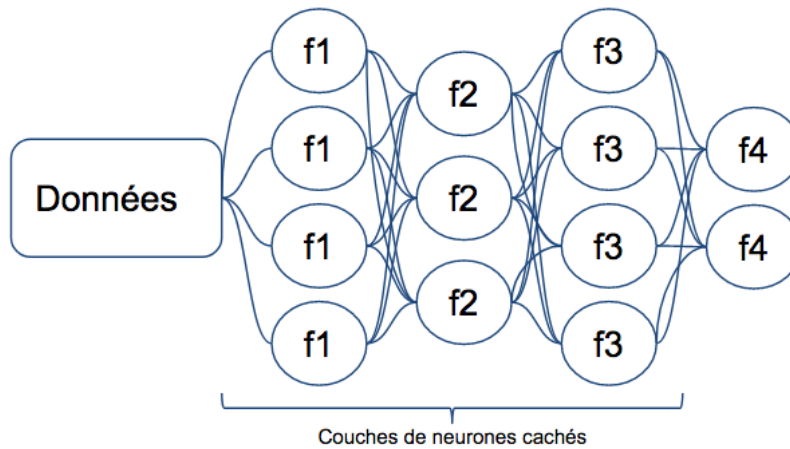


Figure 8 : Schéma d'une architecture classique d'apprentissage profond. Chaque couche de neurones cachés effectue une opération spécifique, la fonction nommée f4 est l'opération de classification (ici deux classes en sortie).

L'étape d'apprentissage est généralement réalisée par la rétropropagation stochastique du gradient [LeCun *et al.*, 1988]. Cette méthode, illustrée Figure 9, consiste à calculer les matrices Jacobiennes des fonctions utilisées dans les couches de neurones cachés. Néanmoins, **l'apprentissage profond non supervisé n'est pas possible pour le moment**. Une piste intéressante pour pallier ce problème est l'utilisation des réseaux de Boltzmann, malheureusement ces réseaux demandent un temps presque infini pour être entraînés sans compromis [Salakhutdinov et Larochelle, 2010].

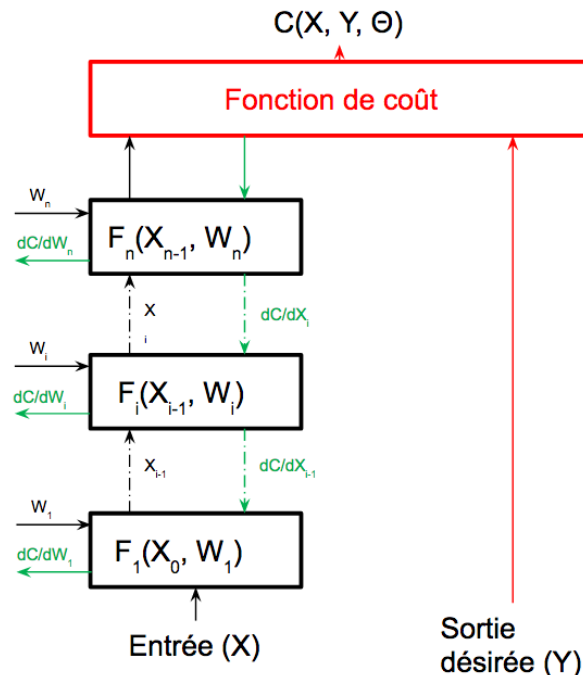


Figure 9 : Pour une base d'apprentissage (X, Y) , les matrices Jacobiennes sont calculées en fonction des poids W_i et des entrées X_{i-1} . Ces matrices optimisent de proche en proche les poids et les paramètres des fonctions F_i .

Parmi les principaux défauts de l'apprentissage profond, les temps d'apprentissage et de généralisation ne permettent pas leur utilisation sur de grandes images à l'échelle du pixel [Cruz-Roa *et al.*, 2018]. De plus, l'apprentissage profond est une famille de méthodes très sensibles à la base d'apprentissage utilisée, une erreur sur 4,02% des données de cette base pouvant amener une erreur de classification de 97% [Papernot *et al.*, 2015]. Dans le cadre de mon travail, cela signifierait que si 4,02% des objets (dans notre cas, les pixels) ne sont pas affectés à la bonne classe dans la base segmentée manuellement, la méthode d'apprentissage profond pourrait se tromper d'affectation de classe dans 97% des pixels de la base de test. Dans le travail de comparaison entre la classification d'une base de 200 000 images biomédicale entre des opérateurs humains et des algorithmes d'apprentissage profonds de Kermany *et al.*, l'erreur médiane des opérateurs humains est de 4,40% [Kermany *et al.*, 2018]. Le nombre de pixels à segmenter manuellement dans nos études est de plusieurs dizaines de millions, l'erreur humaine de segmentation est d'autant plus probable. Pour toutes ces raisons, **les méthodes d'apprentissage profond ne seront pas utilisées dans le cadre de ce travail.**

5.3. Définition des caractéristiques

Une caractéristique est un chiffre entier ou flottant extrait à partir de la donnée brute. À l'échelle du pixel, il existe deux familles d'espaces de caractéristiques : les espaces colorimétriques qui caractérisent la couleur du pixel et les espaces texturaux qui décrivent l'organisation des pixels avoisinant le pixel étudié. Ces caractéristiques permettent de décrire chaque pixel et de favoriser la prise de décision des méthodes d'apprentissage automatique. La multiplicité des informations facilite la séparation des différentes classes constituant de l'image.

5.3.1. Espaces colorimétriques

La couleur est la perception d'une ou d'un mélange d'ondes électromagnétiques entre 380 et 780 nanomètres (nm) par l'œil humain (Figure 10).

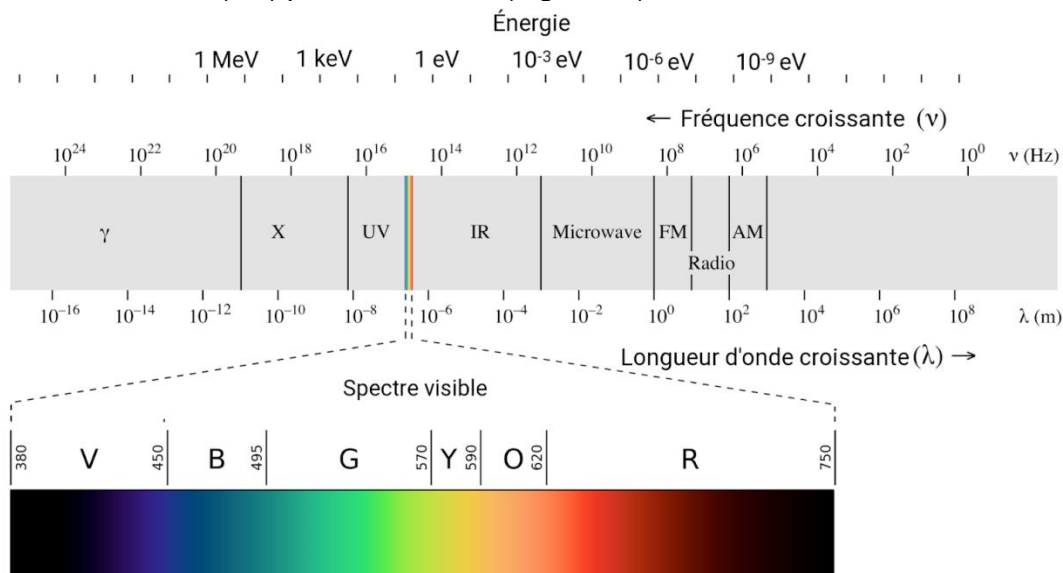


Figure 10 : Positionnement du spectre visible dans le spectre électromagnétique (image tirée de Wikipédia).

Les ondes électromagnétiques seules sont dites monochromatiques et dans cette suite de **couleurs pures** (rouge, orange, jaune, vert, bleu violet), certaines comme le brun ou le rose ne sont pas représentées. Ces couleurs seront appelées **couleurs composites** dans la suite de ce manuscrit car étant issues d'un mélange d'ondes monochromatiques [Newton, 1704]. Pour nommer et différencier les couleurs, des notions caractérisant des propriétés psychophysiques ont été définies, inspirées de la théorie de la couleur de Von Goethe [Von Goethe, 2015]. Ainsi, une couleur peut être caractérisée par trois propriétés principales (Figure 11) :

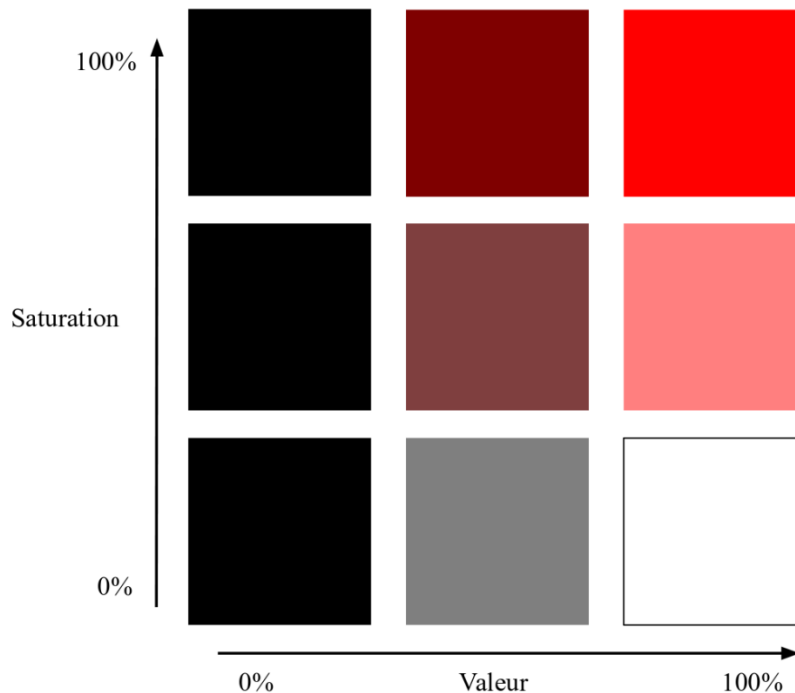


Figure 11 : Exemple de variation de la saturation et de la valeur pour le ton rouge.

- Sa valeur représentant sa luminosité entre le noir et le blanc. Une faible valeur décrit une couleur sombre et une valeur élevée décrit une couleur claire.
- Sa vivacité (ou son niveau de saturation) représentant son éloignement du gris. Une faible valeur décrit une couleur terne et une forte valeur une couleur vive.
- Son ton ou teinte décrivant la couleur en elle-même (rouge, bleu, jaune sont des tons).

Les teintes élémentaires sont définies comme étant les teintes qui ne partagent pas de champ chromatique. Par exemple, il n'existe pas de vert-rouge mais il existe du bleu-violet (indigo), du jaune-rouge (orange). Les deux couples de teintes élémentaires sont le vert-rouge et le jaune-bleu.

Les **caractéristiques colorimétriques** sont les éléments d'un vecteur d'un espace colorimétrique. Les espaces colorimétriques sont des espaces vectoriels de représentation de la couleur visible et différenciable par l'homme. Au cours des derniers siècles, plusieurs systèmes de représentation ont vu le jour afin d'approcher la couleur encodée par notre propre système de perception biologique.

Espace Rouge Vert Bleu :

Le premier espace colorimétrique a été imaginé par Hermann Grassmann en 1853, le père de l'algèbre linéaire. Cet espace, nommé Loi de Grassmann, est basé sur le travail de Hermann Von Helmholtz et décrit la couleur comme étant le mélange additif de trois couleurs primaires : le rouge, le vert et le bleu [Grassmann, 1853]. Il décrit aussi la manière d'additionner deux couleurs (addition vectorielle) et d'en diminuer la valeur (multiplication d'un vecteur par un réel). C'est au début des années 1930 que la Commission Internationale de l'Eclairage (CIE) définit de manière précise l'espace **Rouge Vert Bleu** (RVB) obéissant à la Loi de Grassmann. Le CIE a repris les données expérimentales de William David Wright [WD Wright, 1929] et de John Guild [Guild, 1932] issues d'expériences psychologiques qui consistaient à demander à plusieurs observateurs de reconstituer une couleur monochromatique à l'aide des trois sources de couleurs modulables en intensité. Ces sources de couleurs sont le rouge à 700 nm, le vert à 546,1 nm et le bleu à 435,8 nm. Ces données déterminent les fonctions colorimétriques $R(\lambda)$, $V(\lambda)$ et $B(\lambda)$ avec λ la longueur d'onde (Figure 12).

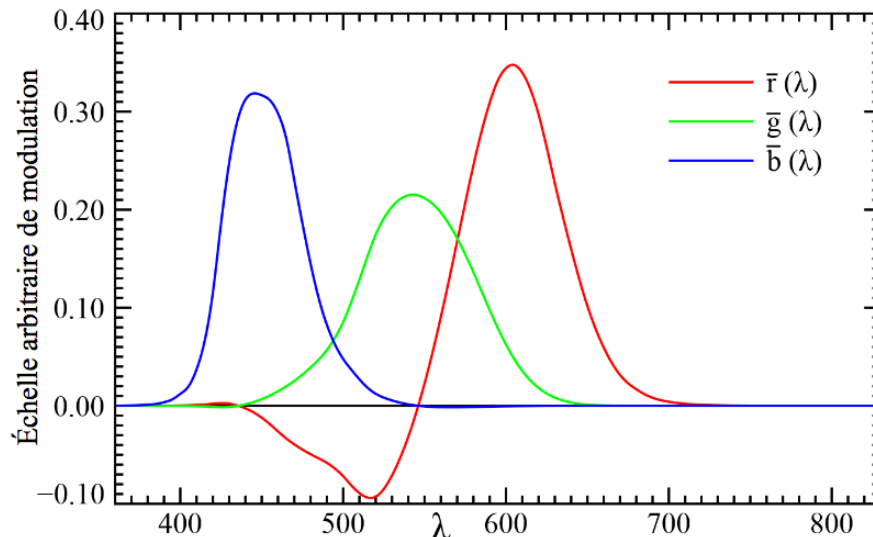


Figure 12 : Fonctions colorimétriques de l'espace CIE RVB. Chaque couleur issue d'une fréquence est exprimée comme la somme des trois fonctions. L'axe des ordonnées est une échelle arbitraire représentant la modulation des différentes sources de couleurs primaires.

Espace CIE XYZ :

À partir de la décomposition de RVB, la CIE décida de construire un nouvel espace colorimétrique nommé **CIE XYZ** plus proche de la vision humaine. Cet espace est une transformation linéaire de RVB et répond à cinq critères :

- Toutes les fonctions colorimétriques $R(\lambda)$, $V(\lambda)$ et $B(\lambda)$ sont positives
- Y est exactement égale à la l'efficacité lumineuse spectrale
- Le point blanc est égal à $X = Y = Z = \frac{1}{3}$
- Le gamut (l'ensemble des couleurs représentées) de CIE XYZ doit inclure celui de RVB
- La fonction Z est égale à zéro au-dessus de 650 nm

Soit R , V et B les caractéristiques de l'espace RVB. La formule permettant le calcul de X , Y et Z se trouve ci-dessous.

$$X = \frac{1}{0,17697} (0,49 * R + 0,31 * V + 0,2 * B) \quad (7)$$

$$Y = \frac{1}{0,17697} (0,17697 * R + 0,8124 * V + 0,01063 * B) \quad (8)$$

$$Z = \frac{1}{0,17697} (0,01 * V + 0,99 * B) \quad (9)$$

Les espaces RVB et CIE XYZ sont linéaires. Ces espaces sont stables selon les principales propriétés de la lumière (teinte, saturation, valeur). Dans ce contexte, la stabilité traduit la proportionnalité de variation entre les *efférents* et les *afférents*, une petite modification d'une propriété de la couleur à transcrire implique une petite modification dans la couleur décrite par l'espace colorimétrique. Par contre, les propriétés de la lumière sont corrélées entre elles, un déplacement dans l'espace colorimétrique implique un déplacement dans toutes les propriétés de la lumière [Cheng *et al.*, 2001]. Dans le contexte de la vision par ordinateur, cette corrélation est problématique pour séparer des objets avec des tons différents par exemple.

Espace CIE L*a*b* :

Pour décorrélérer les propriétés de la lumière, des espaces séparant teintes et luminance ont été inventés. Le premier est l'espace de Hunter proposé en 1946 et standardisé en espace **CIE L*a*b*** en 1976. Cet espace propose de séparer la luminance L* des couples de teintes élémentaires rouge-vert a* et bleu-jaune b*. Les composantes de CIE L*a*b* sont obtenues par une transformation non-linéaire de CIE XYZ. Les formules de conversion se trouvent ci-dessous.

$$L^* = 116 * f\left(\frac{Y}{Y_n}\right) - 16 \quad (10)$$

$$a^* = 500 * \left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right) \quad (11)$$

$$b^* = 200 * \left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right) \quad (12)$$

$$\text{avec } f(t) = \begin{cases} \sqrt[3]{t} & \text{si } t > \delta^3 \\ \frac{t}{3 * \delta^2} & \text{sinon} \end{cases} \text{ avec } \delta = \frac{6}{29}$$

X_n, Y_n et Z_n sont les valeurs de référence du point blanc de l'espace colorimétrique CIE XYZ présenté plus haut. Dans la norme D65, standardisant le point blanc à la lumière du jour, les valeurs sont : $X_n = 95,047$, $Y_n = 100,00$ et $Z_n = 108,883$.

Espace Hue Saturation Value :

Hue Saturation Value (HSV) est un espace colorimétrique non-linéaire construit à partir de RVB [Joblove et Greenberg, 1978]. Cet espace sépare le ton des autres propriétés de la couleur : la saturation et la valeur. Soit R, V, B les caractéristiques normalisées de l'espace RVB, trois paramètres sont calculés : $C_{max} = \max(R, V, B)$, $C_{min} = \min(R, V, B)$ et $\Delta = C_{max} - C_{min}$. Ces paramètres sont utilisés pour calculer H, S et V .

$$H = \begin{cases} 0^\circ \text{ si } \Delta = 0 \\ 60^\circ * \left(\frac{V-B}{\Delta} \text{ modulo } 6 \right) \text{ si } C_{max} = R \\ 60^\circ * \left(\frac{B-R}{\Delta} + 2 \right) \text{ si } C_{max} = V \\ 60^\circ * \left(\frac{R-V}{\Delta} + 4 \right) \text{ si } C_{max} = B \end{cases} \quad (13)$$

$$S = \begin{cases} 0 \text{ si } C_{max} = 0 \\ \frac{\Delta}{C_{max}} \text{ sinon} \end{cases} \quad (14)$$

$$V = C_{max} \quad (15)$$

Les espaces non linéaires CIE L*a*b* et HSV sont instables à des bas niveaux de saturation [Cheng *et al.*, 2001]. Par contre, ils sont adaptés pour séparer des éléments en fonction de leur teinte, de leur saturation ou de leur valeur. Le Tableau 1 résume les espaces colorimétriques présentés (RVB, CIE XYZ, CIE L*a*b* et HSV) et utilisés dans le cadre de ce travail.

	Eloigné de la vision humaine	Proche de la vision humaine
Linéaire	RVB	CIE XYZ
Non linéaire	HSV	CIE L*a*b*

Tableau 1 : Espaces colorimétriques utilisés dans nos études. Tous les espaces colorimétriques peuvent être calculés à partir de RVB, HSV est une transformation non linéaire de RVB, CIE L*a*b* de CIE XYZ, enfin CIE XYZ est une transformation linéaire de RVB.

5.3.2. Espaces décrivant la texture

La texture d'une image est définie par Haralick comme l'organisation statistique des intensités des pixels [Haralick *et al.*, 1973]. La texture est une information liée à l'échelle d'analyse et de calcul de l'image comme représenté Figure 13. À l'instar des caractéristiques colorimétriques, **les caractéristiques texturales** sont des éléments d'un vecteur d'**espace descriptif de texture**, se différenciant par leurs paradigmes d'extraction. Ce paradigme peut être d'origine statistique, d'origine fréquentielle, d'origine descriptive ou même psychologique. Cette partie s'attache à décrire les différents espaces descriptifs de la texture utilisés dans ce manuscrit.

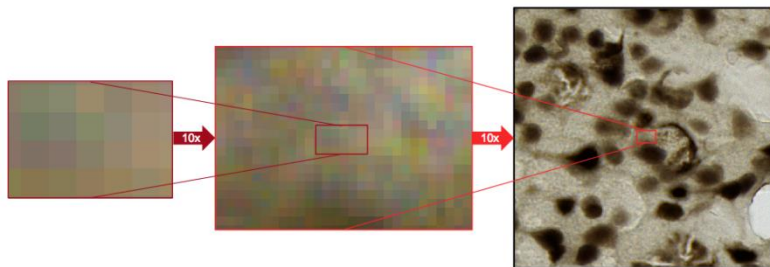


Figure 13 : Exemple de l'importance de l'échelle pour l'analyse de la texture dans une image. Les images successives représentent deux changements d'échelles successifs réalisés à partir de l'image originale. Les informations interprétables sur les images produites successivement sont très différentes.

Espace descriptif d'Haralick :

Le premier espace descriptif de texture intègre les caractéristiques statistiques des matrices de cooccurrences de niveaux de gris (MCNG) [Haralick *et al.*, 1973]. Les matrices de cooccurrences de niveaux de gris sont des matrices symétriques définies positives décrivant les suites d'occurrences entre deux niveaux de gris selon un angle donné comme expliqué dans la Figure 14. Généralement quatre angles sont utilisés afin de décrire les directions principales en connexité-1 : 0° , 45° , 90° et 135° .

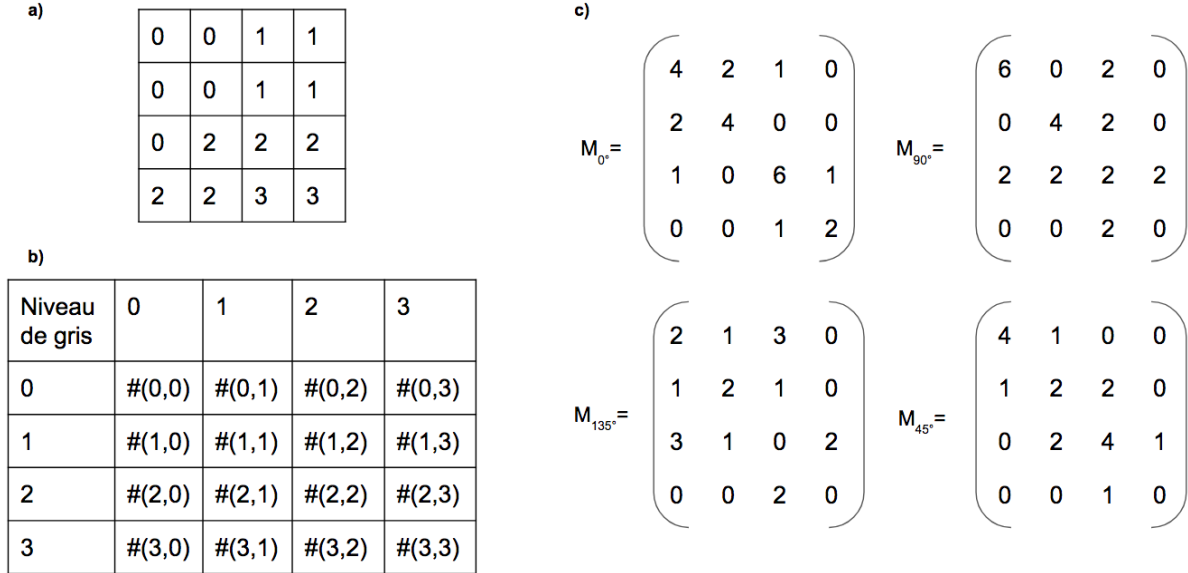


Figure 14 : Construction des matrices de cooccurrences des niveaux de gris. **a)** Exemple d'une image 4x4 pixels avec 4 niveaux de gris. **b)** Forme générale d'une MCNG, $\#(i,j)$ représente le nombre d'occurrences de deux niveaux de gris d'intensité i et j voisins. **c)** MCNG de l'image **a)** dans quatre directions différentes (0° , 45° , 90° , 135°).

Pour chaque MCNG, quatorze caractéristiques texturales statistiques sont extraites. Pour quatre MCNG, cinquante-six caractéristiques sont calculées. Seules les quatre premières caractéristiques sur les quatorze sont extraites pour la segmentation d'une image histologique, les autres caractéristiques étant redondantes avec les quatre premières [Doyle *et al.*, 2007][Sertel *et al.*, 2009][Kather *et al.*, 2016]. On peut aussi citer l'utilisation d'autres caractéristiques d'Haralick en radiomique pour caractériser des tumeurs sur une IRM [Goya-Outi *et al.*, 2018].

Soit $p(i,j)$ la valeur normalisée en position (i,j) de la MCNG, $p_x(i)$ et $p_y(j)$ probabilité marginales en x et en y , N_g le nombre de niveaux de gris dans l'image, $p_{x+y}(k) = \sum_{i+j=k} p(i,j)$ avec $k = 2, 3, \dots, 2N_g$ la probabilité marginale cumulée et $p_{x-y}(k) = \sum_{|i-j|=k} p(i,j)$ avec $k = 0, 1, \dots, N_g - 1$ la probabilité marginale différentielle. Les quatorze caractéristiques décrites par Haralick sont détaillées dans le Tableau 2.

1	Deuxième moment angulaire	$f_1 = \sum_i \sum_j p(i,j)^2$
2	Contraste	$f_2 = \sum_{n=0}^{N_g-1} n^2 \left(\sum_{ i-j =n} p(i,j) \right)$
3	Corrélation	$f_3 = \frac{\sum_i \sum_j ijp(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$ avec μ_x, μ_y, σ_x et σ_y moyennes et écarts-types des probabilités marginales p_x et p_y
4	Variance	$f_4 = \sum_i \sum_j (i - \mu)^2 p(i,j)$
5	Moment différentiel inverse	$f_5 = \sum_i \sum_j \frac{1}{1 + (i-j)^2} p(i,j)$
6	Moyenne de la probabilité marginale cumulée	$f_6 = \sum_{i=2}^{2N_g} ip_{x+y}(i)$
7	Variance de la probabilité marginale cumulée	$f_7 = \sum_{i=2}^{2N_g} (i - f_6)^2 p_{x+y}(i)$
8	Entropie de la probabilité marginale cumulée	$f_8 = - \sum_{i=2}^{2N_g} p_{x+y}(i) \log(p_{x+y}(i))$
9	Entropie	$f_9 = - \sum_i \sum_j p(i,j) \log(p(i,j))$
10	Variance de la probabilité marginale différentielle	$f_{10} = \sum_{i=0}^{N_g-1} (i - \mu_{x-y})^2 p_{x-y}(i)$ avec μ_{x-y} moyenne de la probabilité marginale différentielle
11	Entropie de la probabilité marginale différentielle	$f_{11} = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \log(p_{x-y}(i))$
12	Information mutuelle de la corrélation	$f_{12} = \frac{f_9 - HXY1}{\max(HX, HY)}$ avec HX et HY entropies des probabilités marginales p_x et p_y , $HXY1 = - \sum_i \sum_j p(i,j) \log(p_x(i)p_y(j))$
13	Information mutuelle de la corrélation	$f_{13} = (1 - \exp(-2(HXY2 - f_9)))^{\frac{1}{2}}$ avec $HXY2 = - \sum_i \sum_j p_x(i)p_y(j) \log(p_x(i)p_y(j))$
14	Coefficient de corrélation maximale	$f_{14} = (\text{Deuxième plus grande valeur propre de } Q)^{\frac{1}{2}}$ avec $Q(i,j) = \sum_k \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)}$

Tableau 2 : Caractéristiques statistiques calculées sur chacune des MCNG. Les quatre premières caractéristiques sont les plus utilisées dans le cadre de la segmentation d'images histologiques.

Les espaces d'Haralick possèdent un grand nombre de caractéristiques qui sont coûteuses en temps de calcul et en mémoire utilisée, les quatre MCNG nécessitent au moins un parcours entier de l'image à analyser. De plus, les caractéristiques en elles-mêmes nécessitent le parcours de chacune des MCNG. Cet espace de caractéristique est le plus coûteux des espaces étudiés dans ce manuscrit.

Filtres de Gabor :

Un filtre de Gabor est un filtre linéaire convolutif utilisant la décomposition fréquentielle pour analyser les propriétés texturales d'une image [Gabor, 1946][Fogel et Sagi, 1989]. Ce filtre bidimensionnel minimise le produit des aires effectives occupées dans le domaine spatial et fréquentiel de l'image [Daugman, 1987]. En ce sens, les filtres de Gabor permettent une description compacte de l'espace image-fréquence. Mathématiquement, un filtre de Gabor est une gaussienne multipliée par une sinusoïde. Ce filtre est paramétré par cinq éléments :

- λ longueur d'onde de la sinusoïde
- θ l'orientation du filtre de Gabor par rapport à l'image
- ψ la phase de la sinusoïde
- σ l'écart-type de la gaussienne
- γ le rapport spatial définissant le support elliptique du filtre

Un filtre de Gabor est défini dans le plan complexe, il possède donc une partie réelle et une partie imaginaire ayant leurs propres propriétés. La partie imaginaire est un détecteur de contours (de changement de texture) et la partie réelle un détecteur de région de texture uniforme [Mehrotra *et al.*, 1992]. Pour (x, y) les coordonnées d'un pixel dans une image :

Partie réelle :

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'^2}{\lambda} + \psi\right) \quad (16)$$

Partie imaginaire :

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \sin\left(2\pi \frac{x'^2}{\lambda} + \psi\right) \quad (17)$$

avec $x' = x * \cos(\theta) + y * \sin(\theta)$ et $y' = -x * \sin(\theta) + y * \cos(\theta)$

Un espace de Gabor est un espace rassemblant les réponses de plusieurs filtres de Gabor se différenciant par leurs hyperparamètres. En fonction des choix des hyperparamètres, le nombre de caractéristiques N_{Gabor} de cet espace va varier. $N_{Gabor} = 2card(\lambda, \theta, \psi, \sigma, \gamma)$ avec $card$ fonction calculant le nombre des hyperparamètres choisis. Ainsi, le nombre de caractéristiques calculées se compte en dizaines voire en centaines [Doyle *et al.*, 2008][Ding *et al.*, 2016][Kather *et al.*, 2016].

Motifs Binaires Locaux (Local Binary Pattern) :

Les Motifs Binaires Locaux (MBL) sont des caractéristiques de textures basées sur la description du motif autour d'un point [Ojala *et al.*, 1994]. Les intensités des niveaux de gris autour d'un pixel spécifique sont seuillées par l'intensité du pixel central, ce qui forme un vecteur binaire. La traduction de ce vecteur en entier naturel est le MBL du pixel central. Le MBL fut amélioré en le rendant invariant par rotation et en incluant la notion de motifs uniformes (Figure 15) [Ojala *et al.*, 2002]. L'invariance par rotation est obtenue en minimisant le vecteur binaire. Un motif uniforme est défini comme un vecteur binaire ayant **aucune ou une seule série de 1**. Il existe donc exactement $P + 1$ motifs uniformes, P étant le nombre de pixels sur le cercle autour du pixel central. Le MBL est défini comme le nombre de '1' du vecteur binaire si celui-ci correspond à un motif uniforme. Sinon, le MBL est égal à $P + 1$, cette classe correspondant à tous les motifs non uniformes.

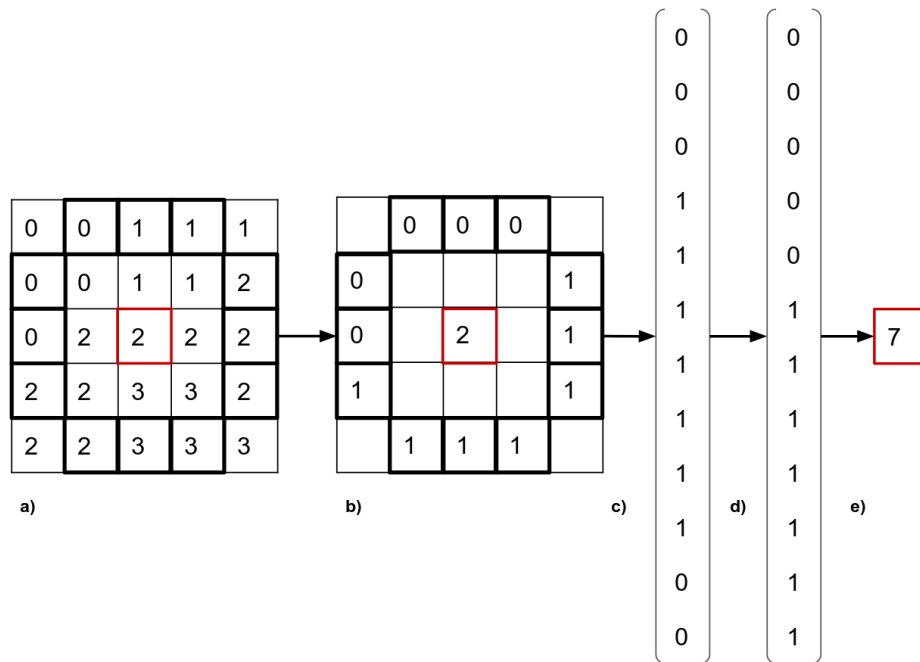


Figure 15 : Exemple du calcul du MBL invariant par rotation uniforme de rayon 2 sur une image de 5x5 pixels avec quatre niveaux de gris. **a)** Image d'origine, le pixel rouge est le pixel dont le MBL sera calculé et les pixels en noirs correspondent au cercle de rayon 2 autour du pixel rouge. **b)** Les pixels environnants sont seuillés et forme le vecteur **c)**. **d)** le vecteur minimisé et enfin, le MBL est calculé, le vecteur minimisé ayant sept '1' en **e)**.

Le MBL est composé d'une seule caractéristique de texture, ce qui est un avantage face aux espaces de Gabor et de Haralick. Néanmoins, le calcul du MBL nécessite le parcours d'un nombre fixe de pixels et d'une minimisation, il est plus coûteux à calculer par rapport à l'espace de Gabor mais reste moins coûteux par rapport à l'espace d'Haralick.

Pour extraire les caractéristiques décrits ci-dessus, la complexité algorithmique est une notion importante surtout lorsque l'on souhaite travailler sur des données massives. En effet, elle est un indicateur pertinent afin de déterminer le coût computationnel de l'extraction d'une caractéristique sur tous les éléments de la donnée en question. La Section suivante adresse cette notion

5.3.3. Complexité algorithmique

La complexité d'un algorithme est le nombre d'opérations élémentaires effectuées par un algorithme. Il est souvent exprimé comme étant borné par une combinaison de polynômes, logarithmes ou exponentielles en fonction de la taille de l'image n ($O(f(n))$). Un algorithme est considéré comme rapide lorsqu'il est sub-linéaire, c'est à dire proportionnel à $O(\log(n))$. Un algorithme est considéré comme lent lorsqu'il est polynomial et exponentiel.

Les complexités algorithmiques ont été estimées aussi grâce à k la taille en pixels du rayon de l'élément structurant sur lequel est réalisée l'extraction des caractéristiques du pixel. La complexité algorithmique est donc $O(f(n, k))$. Dans tous les espaces considérés, la complexité est linéaire en fonction du nombre de pixels, l'extraction de caractéristiques étant indépendante de la nature du pixel. Les caractéristiques colorimétriques n'ayant pas d'élément structurant, la complexité est donc constante vis-à-vis de ce paramètre. Les caractéristiques

d'Haralick nécessitent au moins un parcours de la fenêtre d'analyse et un parcours de chacune des MCNG. Le parcours de la fenêtre d'analyse est donc proportionnel à $O(n k^2)$ et le parcours des MCNG, bien que n'étant pas négligeable en terme de temps, ne dépend pas de k et est fixe ($O(n)$). Les filtres de Gabor ayant un élément structurant de taille limité par l'écart-type de la Gaussienne, paramètre qui est fixé dans le cadre de ce travail, sa complexité est constante par rapport à k . Enfin le MBL intègre un parcours du périmètre du cercle faisant office d'élément structurant et une minimisation par dichotomie du vecteur résultant. Sa complexité est proportionnelle à $O(n k \log(k))$.

6. Réduction de dimensionnalité

La réduction de dimensionnalité est le procédé permettant de décrire une donnée avec un nombre restreint de variables. Elle permet d'éviter la malédiction de la dimensionnalité [Bellman, 2013], de réduire les temps de traitement et la quantité de données traitées et de faciliter l'interprétation des résultats par des opérateurs humains. Il existe deux grandes familles de méthodes, les méthodes projectives, comme l'analyse en composantes principales [Pearson, 1901], et la sélection de caractéristiques, au cœur du présent chapitre.

6.1. Réduction de dimensionnalité et problème de stabilité

Les **méthodes projectives** sont une suite d'opérations algébriques transformant l'espace de description afin de trouver une ou plusieurs composantes, fusionnant plusieurs caractéristiques, décrivant de manière condensée la donnée. Ces méthodes peuvent ainsi servir pour compresser ou sélectionner des caractéristiques. Néanmoins, l'interprétabilité des composantes, fusions linéaires de plusieurs caractéristiques, est particulièrement complexe à appréhender. De plus, ces méthodes ne permettent pas d'optimiser l'extraction de caractéristique en excluant les espaces algorithmiquement trop complexes à calculer. En effet, si les composantes sélectionnées fusionnent des caractéristiques issues d'extractions coûteuses (comme les caractéristiques d'Haralick Section 5.3.2), le calcul de ces espaces devient obligatoire.

Les **méthodes de sélection de caractéristiques** jugent quant à elles les caractéristiques individuellement ou par sous-ensembles. Ces méthodes permettent une meilleure interprétabilité des espaces de caractéristiques et de l'optimisation de l'étape d'extraction. Pour toutes ces raisons, j'ai donc choisi cette famille de méthodes. Elles jugent la pertinence d'une caractéristique ou d'un groupe de caractéristiques pour un objectif donné et à l'aide d'une métrique spécifique. Une caractéristique qui n'est pas pertinente est appelée **caractéristiques non-informative** parfois, elle est aussi appelée en anglais *garbage feature* (caractéristique détritue) [Kudo et al., 2000].

Le principal problème des méthodes de sélection de caractéristiques est **le manque de stabilité de la sélection pour une méthode d'apprentissage automatique donnée** [Kalousis et al., 2005][Meinshausen et al., 2010][Kuncheva et al., 2018]. Ce **problème implique un manque de généralisation des ensembles de caractéristiques sélectionnées**, ces ensembles étant trop spécifiques par rapport à la base d'apprentissage utilisée. À l'origine, les méthodes étaient en général conçues pour améliorer les performances en termes de qualité de segmentation et de temps de traitement. La stabilité n'était pas la priorité des recherches sur le sujet. En effet, le phénomène de pointe [Sima et Dougherty, 2008] permet d'améliorer parfois les résultats d'évaluation (F-Score par exemple) de la classification en réduisant le nombre de caractéristiques [Kudo et al., 2000][Guyon, 2003]. Pour l'instant, ce phénomène de pointe n'est pas prédictible et l'ensemble de caractéristiques sélectionnées est **un optimum local lié à la**

base d'apprentissage, se focalisant uniquement sur la recherche de la meilleure évaluation possible. De ce fait, il n'est pas forcément généralisable pour des données similaires. Quelques méthodes tentent de pallier ce problème [Meinshausen *et al.*, 2010][Awada *et al.*, 2012][Shah *et al.*, 2012][Kuncheva *et al.*, 2018]. Par exemple, la méthodologie de Meinshausen et Bühlmann (reprise dans un contexte différent par Shah *et al.*) consiste à améliorer les méthodes existantes en introduisant des amorces agrégantes (*bagging*) dans certains procédés de sélection [Breiman, 1994][Meinshausen *et al.*, 2010]. Les amorces agrégantes permettent de moyenner les différents sous-ensembles sélectionnés. Les méthodologies présentées dans l'étude comparative d'Awada et al. sont des méthodes développées pour mesurer la stabilité d'une méthode de sélection, pas pour réaliser une sélection stable [Awada *et al.*, 2012]. À notre connaissance, **il n'existe pas de méthodologie dont le but soit la stabilité de la sélection est le but**. Une méthodologie répondant à ce critère pourrait être utilisée afin d'évaluer et de comparer de manière simple et robuste la stabilité des méthodes de sélection de caractéristiques.

La classification des méthodes de sélection de caractéristiques en trois familles de méthodes a été choisie car elle est le plus souvent présentée dans les études comparatives (Figure 16) [Guyon, 2003][Chandrashekar et Sahin, 2014][Kuncheva *et al.*, 2018]. Les *filter methods* seront appelées « **méthode à filtre** » car elles filtrent les caractéristiques sans utiliser l'apprentissage automatique. Les *embedded methods* seront appelées « **méthodes incluses** » car la sélection est directement incluse dans l'étape d'apprentissage. Enfin, les *wrapper methods* seront appelées « **méthodes intégrées** » car la sélection est réalisée dans le cadre de l'évaluation ou de la validation du modèle d'apprentissage.

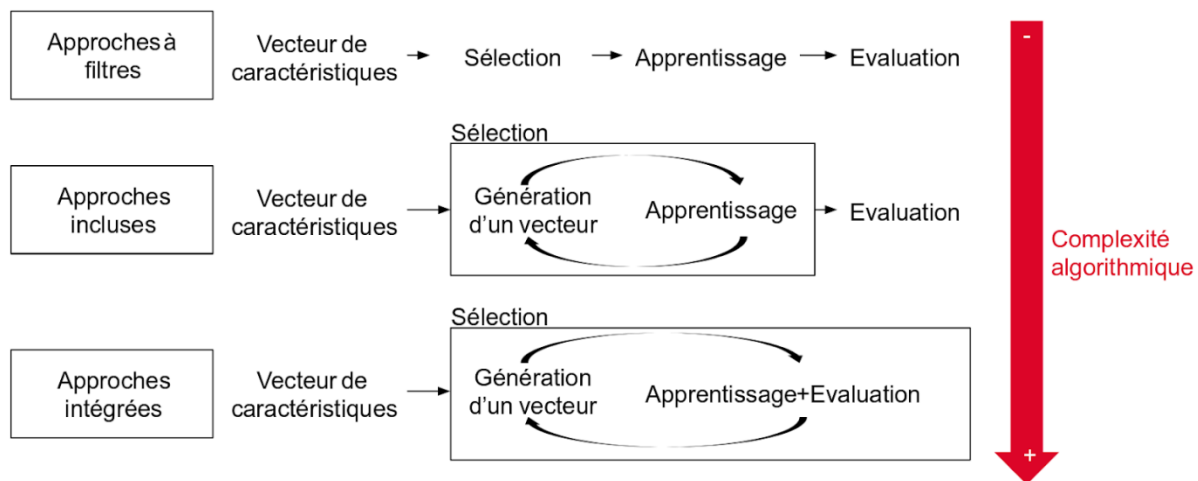


Figure 16 : Familles de méthodes de sélection de caractéristiques. Un vecteur de caractéristiques initial entre dans un procédé de sélection, les approches incluses et intégrées reposant sur la génération d'un sous-ensemble contenu dans le vecteur initial (Figure adaptée, traduite et complétée de la thèse de Julie Hamon [Hamon, 2013]).

6.2. Familles de méthodes de sélection de caractéristiques

6.2.1 Méthode à filtres

Les méthodes à filtres (*filter methods*) utilisent des critères de sélection sur l'ensemble des vecteurs de caractéristiques disponibles pour une base d'apprentissage donnée afin de classer les dites caractéristiques en fonction de leur pertinence pour une classe donnée. Il existe

de nombreuses définitions formelles de cette notion de pertinence [Chandrashekar et Sahin, 2014]. Une des définitions est basée sur la corrélation de Pearson entre une caractéristique et la classe d'intérêt. Plus cette corrélation est importante (c'est à dire proche de 1 ou de -1) plus cette caractéristique sera pertinente pour délimiter la classe d'intérêt [Guyon., 2003]. Un autre critère communément utilisé est l'information mutuelle, représentant la dépendance entre deux variables [Lazar *et al.*, 2012]. Plus l'information mutuelle entre deux variables est proche de 0, plus ces deux variables seront considérées comme indépendantes et inversement. L'objectif est de sélectionner les caractéristiques les plus informatives à une classe d'intérêt donnée.

Bien que ces méthodes soient particulièrement efficaces en termes de temps de calcul et robustes au sur-apprentissage (le classement des caractéristiques n'étant pas lié à un apprentissage spécifique), elles ne prennent pas en compte des informations importantes comme la *pertinence* des caractéristiques pour un algorithme d'apprentissage automatique donné et la *redondance* des caractéristiques entre elles. Une caractéristique est définie comme redondante si elle varie de manière similaire à une autre caractéristique sélectionnée. Ainsi, il est impossible de garantir l'optimalité d'un vecteur de caractéristiques choisi par cette famille de méthode pour un algorithme d'apprentissage automatique spécifique. De plus, ces méthodes ne proposant qu'un classement des caractéristiques, le choix de la cardinalité (le nombre de caractéristiques choisies) est défini arbitrairement.

6.2.2. Méthodes incluses

Les méthodes incluses (*embedded methods*) sont des méthodes de sélection incluses dans l'étape d'apprentissage automatique. Elles permettent, contrairement aux méthodes à filtres, de prendre en compte d'une part l'algorithme d'apprentissage automatique utilisé (voir Section 5.2) et d'autre part la redondance entre les caractéristiques [Chandrashekar et Sahin, 2014]. Ces procédés se basent sur la génération de sous-ensembles de caractéristiques issues d'un vecteur initial, généralement conçu manuellement. Ces sous-ensembles sont jugés pertinents ou non lors de l'étape d'apprentissage, le sous-ensemble le plus pertinent étant alors sélectionné. L'apprentissage du MSV ou des forêts aléatoires sont des bons exemples de méthodes incluses. Le MSV pondère les caractéristiques par le vecteur w tandis que les forêts aléatoires trient les caractéristiques selon l'entropie de Shannon (voir Section 5.2).

Ces méthodes sont des compromis entre les méthodes intégrées et les méthodes à filtres. En effet, elles intègrent un procédé de proposition d'un sous-ensemble de caractéristiques, puis cet ensemble est évalué par l'étape d'apprentissage de la méthode. De plus, leur complexité computationnelle est moins importante par rapport aux méthodes intégrées, facilitant leur implémentation et leur utilisation. Cependant la validation n'est pas prise en compte lors de l'application de ces méthodes, ne permettant donc pas de juger de la stabilité d'un sous-ensemble comme nous l'avons définie dans la Section 6.1.

6.2.3. Méthodes intégrées

Les méthodes intégrées (*wrapper methods*) sont des méthodes utilisant les outils d'évaluation et de validation de classification en tant que critère optimal de sélection du sous-ensemble de caractéristiques. Les méthodes de cette famille sont classées en deux catégories : les recherches séquentielles et les recherches heuristiques.

Une méthode de recherche séquentielle est le minimum de redondance et le maximum de pertinence (mRMP) [Peng *et al.*, 2005]. Pour une classe $c \in C$, un sous-ensemble de taille fixée $S \subset (c_1, c_2, \dots, c_k)$ et $I(a, b)$ information mutuelle de a et b :

$$D_S = \frac{1}{\text{card}(S)} \sum_{c_i \in S} I(c_i, c) \text{ facteur de pertinence pour un sous-ensemble } S \quad (18)$$

$$R_S = \frac{1}{\text{card}(S)^2} \sum_{c_i, c_j \in S} I(c_i, c_j) \text{ facteur de redondance pour un sous-ensemble } S \quad (19)$$

Le sous-ensemble optimal S_{opt} est le sous ensemble maximisant $\varphi_S = D_S - R_S$. La taille de S étant fixée, cette étape est répétée pour toutes les tailles possibles de façon à avoir $S_{opt1} \subset S_{opt2} \subset \dots \subset (c_1, c_2, \dots, c_k)$. Puis, l'ensemble sélectionné parmi les ensembles optimaux est l'ensemble avec l'erreur de classification minimale, une validation croisée étant utilisée. D'autres méthodes de recherche séquentielle particulièrement lourdes en coût calculatoire existent : les stratégies de Sélection Séquentielle par Accumulation (SSA, en anglais *Sequential Forward Selection*) [Whitney, 1971] et les stratégies de Sélection Séquentielle par Décimation (SSD, en anglais *Sequential Backward Selection*) [Marill et Green, 1963]. Les stratégies SSA accumulent les caractéristiques une par une puis évaluent par F-Score ou un autre critère d'évaluation [Taha et Hanbury, 2015]. Si le sous-ensemble formé a une évaluation plus faible que le sous-ensemble précédent, la caractéristique précédemment ajoutée est ôtée et remplacée par une autre caractéristique. Dans le cas contraire, elle est conservée et le procédé est itéré. Les stratégies SSD sont similaires aux SSA sauf qu'elles ôtent progressivement les caractéristiques non-pertinentes. Le mPMR ainsi que les stratégies SSA et SSD sont affectés par l'effet d'imbrication (*nested effect*) [Pudil *et al.*, 1994]. En effet, les sous-ensembles explorés sont imbriqués les uns dans les autres et donc, biaisés par le choix de la première caractéristique ôtée ou ajoutée. L'exploration combinatoire est donc grandement limitée.

Les Algorithmes Génétiques (AG) font partie de la famille des recherches heuristiques [Lucasius et Kateman, 1993]. Ils sont inspirés de la théorie de l'évolution de Darwin [Darwin, 2004] et comparables au procédé du SELEX [Tuerk et Gold, 1990]. Un méta vecteur de caractéristiques est défini comme un vecteur de caractéristiques où les valeurs à chaque position sont binaires : 1 pour la présence d'une caractéristique, 0 pour son absence. Ainsi un méta vecteur indique les positions, donc les caractéristiques, utilisées pour une classification donnée. Un algorithme génétique se déroule en plusieurs étapes. Tout d'abord, N méta-vecteurs sont générés aléatoirement puis la classification résultante est évaluée par un critère choisi [Taha et Hanbury, 2015]. Puis, les méta-vecteurs ayant les meilleures évaluations vont être répliqués (en plus grande quantité) et être utilisés pour remplacer les méta-vecteurs de moins bonnes qualités. Sur les N nouveaux méta-vecteurs sélectionnés, une étape de mutation est appliquée, chaque position ayant une certaine probabilité de changer (c'est à dire passer de 1 à 0 ou de 0 à 1, cette probabilité est généralement de l'ordre de 5%). Enfin, les vecteurs sont de nouveau évalués et un nouveau cycle de sélection commence. Les conditions d'arrêt sont soit un nombre fixe de cycles de mutation, soit un score d'évaluation atteint [Lucasius et Kateman, 1993]. Cette méthode dépend de façon importante de ses hyperparamètres : le nombre de méta vecteurs, la probabilité de mutation et la condition d'arrêt. Elle dépend également de l'initialisation aléatoire, la reproductibilité des résultats n'étant pas garantie.

Les méthodes intégrées sont les méthodes les plus coûteuses en termes de ressources computationnelles requises. En effet, l'intégralité du procédé, de l'apprentissage automatique à l'étape de validation, est prise en compte. De plus, les méthodes de validation, permettant de réduire le risque de sur-apprentissage et de tester la stabilité des sous-ensembles de caractéristiques, sont intégrées dans cette famille de méthodes.

6.3. Objectifs et difficultés de la méthode de sélection souhaitée

Comme évoqué précédemment, de grandes disparités existent dans les dénominations, la classification des méthodes et la formalisation de l'objectif du problème de sélection de caractéristiques. Les méthodes incluses et intégrées sont souvent confondues. On peut citer les disparités terminologiques entre certains articles [Guyon, 2003] (premier article parlant de la triple classification présentée Figure 16) [Gheyas et Smith, 2010] (ne détaille que deux catégories de méthodes) [Peng, Wu et Jiang, 2010] (ne parlant pas de *embedded methods* mais de *hybrid methods*) [Chandrashekar et al., 2014] (classifiant la méthodologie de [Peng et al., 2005] en tant que *embedded method* alors que [Tang et al., 2014] désigne la méthode en tant que *wrapper method*). De plus, certaines études ne développent pas de terminologie [Dash et Liu, 1997][Kudo et al., 2000][Yin et al., 2013]. Je propose ici une traduction française de cette classification. Le domaine de recherche de la sélection de caractéristiques est un problème complexe. Les ambiguïtés terminologiques et méthodologiques rendent la lecture bibliographique dans ce domaine compliquée. De plus, aucune méthode ne permet la comparaison des algorithmes de sélection du point de vue de la stabilité. Néanmoins, nous pouvons définir des propriétés souhaitables pour la méthodologie voulue dans le cadre de mon travail de recherche :

1. La méthode doit fonctionner avec les méthodes d'apprentissage automatique utilisées dans le cadre de la segmentation des images histologiques.
2. La méthode doit avoir pour objectif principal la stabilité des sous-ensembles de caractéristiques sélectionnés.

La méthode doit donc être une méthode intégrée pour satisfaire les deux propriétés précédentes. De plus, l'utilisation de la Force Brute (FB) que je propose apparaît être une piste prometteuse. En effet, la FB a l'avantage de fournir une solution si elle existe [Heule et Kullmann, 2017]. Cette propriété permet de s'assurer d'une part que la formalisation de notre problème est bonne et d'autre part que la stabilité peut être atteinte dans l'étape de sélection.

Finalement, il nous reste à proposer une stratégie globale de sélection de caractéristiques et à définir une métrique afin d'explorer la combinatoire exhaustive fournie par la FB. Dans ce contexte, j'ai proposé plusieurs critères d'évaluation des caractéristiques (Section 9.2.2 et 10.2.2) que j'ai comparés en termes de stabilité et de F-Score. J'ai également développé une stratégie intégrant ces métriques et regroupant les caractéristiques à partir des résultats de la recherche par Force Brute. Cette approche a pu être mise en œuvre sur plusieurs marqueurs différents. Les résultats sont présentés chapitre IV.

III Extraction et sélection de caractéristiques

7. Bases de données biologiques

Les bases de données biologiques permettent de conserver les informations histologiques. Dans le cadre de ce manuscrit, elles vont principalement être constituées des différentes bases d'images histologiques numérisées à haute-résolution par un microscope à fond clair. Ces bases serviront de bases d'apprentissage, de bases de test et de bases de généralisation. Ce chapitre est consacré à la description des bases constituées avec les marqueurs NeuN et anti-Phox2B, marqueurs neuronaux utilisés respectivement sur le cerveau et le tronc cérébral de macaque. Ce chapitre décrit les segmentations manuelles qui vont servir de référence pour les segmentations automatiques.

7.1. Cadre éthique de l'étude

Les animaux de cette étude ont été euthanasiés et leurs cerveaux extraits en respectant les règles d'éthiques européennes. Ces expérimentations sont soumises à la directive européenne 2010/63 et à l'article R 214-87 à 131 du Code Rural français. De plus, le département vétérinaire de MIRCen a été autorisé d'exercer par l'autorisation n° D 92-032-02 et se soumet aux standards de l'expérimentation animale du bureau du bien-être des animaux de laboratoire (OLAW - n°#A5826-01).

7.2. Étude du marqueur NeuN

7.2.1. Production des données histologiques NeuN

Les données de cette étude ont été produites à partir d'un cerveau de macaque sain de neuf ans pour lequel de multiples analyses histologiques ont été menées afin d'étudier l'anatomie de certaines régions d'intérêt dont l'hippocampe. L'étude des neurones a été réalisée en utilisant le marqueur *Neuronal Nuclei* (NeuN), un marqueur spécifique à la majorité des neurones des vertébrés [Kim *et al.*, 2009]. Ce dernier a été utilisé avec le chromogène 3,3'-diaminobenzidine (DAB) afin de révéler le marqueur NeuN dans des nuances de brun. La position et l'organisation des neurones dans un cerveau de primate sont connues *a priori* grâce notamment aux atlas papier, véritables cartographies des régions et sous-régions anatomiques du cerveau [Paxinos *et al.*, 2000].

Le cerveau entier a été coupé en une série de sections coronales de 40 µm d'épaisseur selon l'axe antérieur-postérieur (plus de mille sections produites) et une sous-série de 134 sections (1 section sur 8 avec un espace inter-coupe de 320µm) a été marquée avec le mélange DAB NeuN. La standardisation des procédures de marquage histologique mise en place à MIRCen assure une grande stabilité dans la qualité des sections produites et une bonne reproductibilité du marquage. Trois sections prises au centre du cerveau ont été numérisées avec le microscope Axio Scan.Z1 de Zeiss (Figure 17) avec un facteur de grossissement x20. Ces sections incluent des sous-régions anatomiques d'intérêt comme le cortex, l'hypothalamus, l'hippocampe et le striatum. Les sections numérisées ont une résolution planaire de 0,22x0,22 µm² et ont une taille moyenne d'environ 40 Gigapixels (Gp) chacune (entre 150 et 200 Gigaoctets de mémoire physique). Les images présentent toutes une très grande diversité

de taille et de densité de neurones marqués (Figure 17.**b**). Lors de la coupe du cerveau, des photographies ont été acquises pour permettre *a posteriori* une reconstruction en 3D du cerveau [Annese *et al.*, 2006][Dauguet *et al.*, 2007].

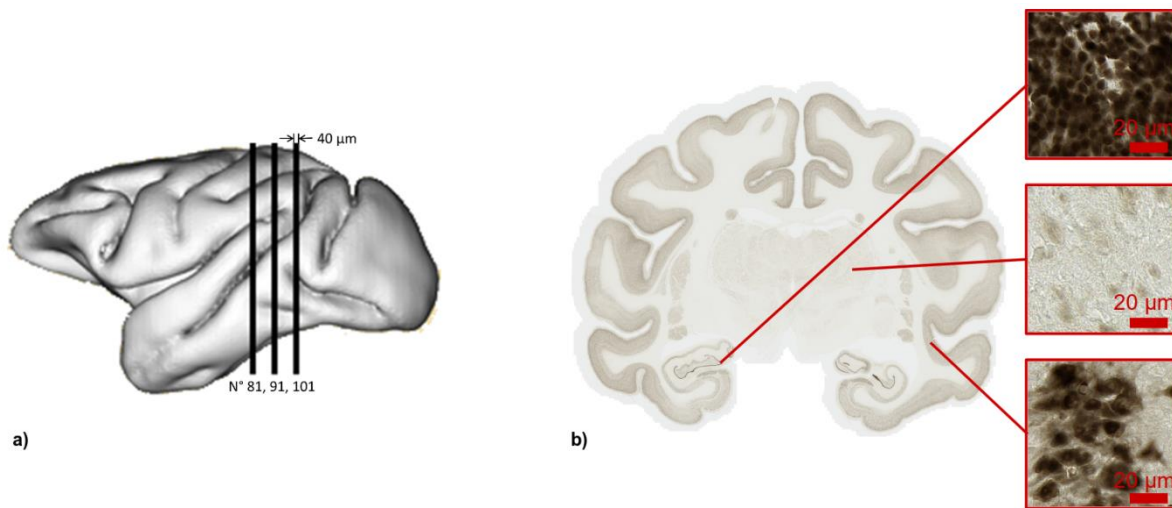


Figure 17 : Présentation des données NeuN. **a)** Surface tridimensionnelle externe du cerveau de macaque ainsi que la position des sections numérisées et leurs numéros. **b)** Section 81 avec des images extraites dans différentes sous-régions du cerveau, présentant la diversité de neurones marqués.

7.2.2. Segmentations manuelles de NeuN

À partir des trois sections numérisées, les deux sections extrêmes ont été utilisées pour générer les bases de notre étude. Ceci permet de limiter les biais d'illumination et les différences antéro-postérieure. De plus, l'ensemble des sous-régions anatomiques d'intérêts sont présentes dans ces deux sections.

Lors de la conception des bases de données, le but était d'avoir une représentation équilibrée des sous-régions, des différents niveaux d'intensité et de densité de marqueur sur l'ensemble des images (Figure 18). Quarante-deux images de dimension 512x512 pixels ont été extraites à l'aide du logiciel Anatomist [Rivière *et al.*, 2011] par un biologiste dans la section numéro 81 et cinquante-huit dans la section numéro 101, amenant le nombre d'images extraites à cent. Ces cent images ont été extraites dans différentes sous-régions d'intérêt du cerveau de macaque choisies pour représenter différents niveaux de densités et d'intensités des neurones marqués avec l'anticorps NeuN (la liste des différentes régions et l'emplacement des images sur les sections numérisées sont donnés dans l'Annexe 1). L'intégralité des pixels des images extraites a été segmentée manuellement par un traiteur d'image en quatre classes : tissu marqué (neurones), tissu non-marqué, fond (pas de tissu) et artéfacts (concentration de marqueur non-spécifique, débris) (Figure 18).

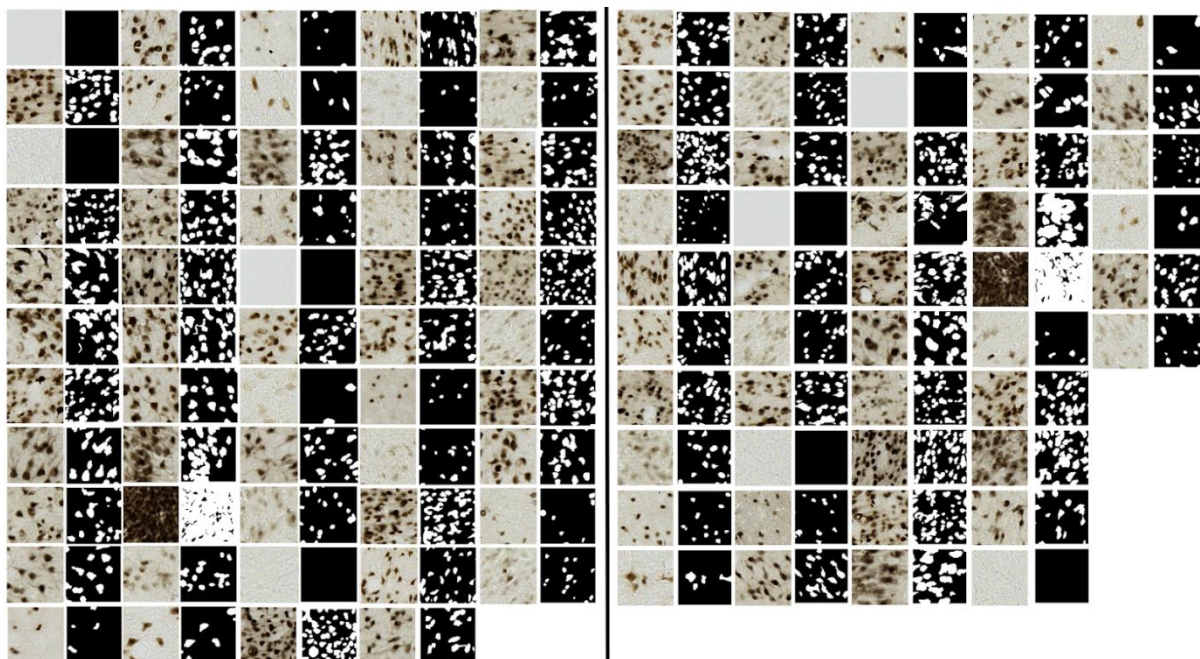


Figure 18 : Marquage NeuN, bases d'apprentissage (à gauche) et de test (à droite) : les données brutes sont associées avec les segmentations manuelles correspondantes mises cote à cote. Les pixels blancs dans les segmentations manuelles sont les régions marquées par l'anticorps NeuN (neurones).

Une base d'apprentissage et une base de test interchangeables ont été constituées à partir de l'ensemble de ces cent images (Figure 18). Ces bases ont été déterminées en suivant trois règles :

- 1) chacune des bases devait avoir autant d'images extraites à partir de l'hémisphère droit que de l'hémisphère gauche,
- 2) chaque base devait avoir au moins un doublon représentant de chaque sous-région anatomique,
- 3) chaque base devait avoir un nombre similaire d'images issues de la section numéro 81 et de la section numéro 101.

Comme le nombre d'images extraites n'est pas identique entre les sections 81 et 101 du fait de la composition anatomique des images (absence de certaines sous-régions de l'hippocampe dans l'une des latéralités de l'une des images), la base d'apprentissage comporte cinquante-quatre images et la base de test quarante-six images.

7.3. Étude du marqueur anti-Phox2B

7.3.1. Production des données histologiques anti-Phox2B

La seconde étude a été réalisée dans le cadre d'une collaboration avec le laboratoire des Handicaps neuromusculaire (UMR-1179 INSERM – Université de Versailles-Saint-Quentin). L'objectif de l'étude menée par cette équipe visait à localiser les neurones appartenant au noyau rétrotrapezoïde (RTN), principal centre respiratoire chémosensible (c'est-à-dire, ayant la capacité d'adapter la respiration à des signaux chimiques, comme le CO_2). Ce noyau a été décrit chez le rongeur (rat et souris) mais pas encore chez le primate (non-humain et humain), ce qui constitue l'intérêt de cette étude. Dans le cadre d'une analyse d'anatomie comparée, la première partie de l'étude visait à cartographier l'ensemble des neurones exprimant le facteur de

transcription Phox2B au sein du tronc cérébral du primate non-humain (*macaca fascicularis*), pour secondairement identifier par des marquages complémentaires, les neurones du RTN. Un tronc cérébral d'un macaque sain de 5 ans a été utilisé pour produire les sections histologiques de cette étude.

Cette protéine (Phox2B) participe au développement de la plupart des centres végétatifs et respiratoires du système nerveux central [Amiel *et al.*, 2003][Gaultier *et al.*, 2005]. Le marqueur utilisé était un anticorps anti-Phox2B. Il s'agit d'un marqueur spécifique des noyaux des neurones exprimant le gène *phox2b*. Toutefois, ce marquage, révélé par la 3,3'-diaminobenzidine (DAB) s'accompagnait d'un marquage non-spécifique important, contrairement au marquage avec l'anticorps NeuN plus spécifique, ce qui est visible Figure 19.**b**. La construction des bases a été réalisée de façon semi-aléatoire. Les neurones d'intérêt doivent être détectés comme étant des zones où anti-Phox2B a des propriétés décelables visuellement par l'expert, parmi le marquage non-spécifique. La différence de marquage est détectée par le contraste entre les différents pixels marqués.

Le tronc cérébral a été coupé en sections de 40 μm d'épaisseur et une série de soixante coupes a été marquée par le marqueur Phox2B (une coupe sur huit soit un espace inter-coupe de 320 μm). L'intégralité des sections a été numérisée avec un microscope Aperio AT2 de Leica biosystems avec un grossissement x20 (Figure 19). La résolution planaire résultante est cependant moins élevée que celle obtenue avec le microscope Axio Scan.Z1 de Zeiss. Les sections numérisées ont une résolution planaire de 0,50x0,50 μm^2 et ont une taille moyenne de 2 Gp. À l'instar des tissus marqués au NeuN, les images des tissus marqués avec l'anti-Phox2B présentent une grande diversité de forme, de densité et d'intensité de marquage (Figure 19). Comme pour le marqueur NeuN, le bloc photographique correspondant au tronc cérébral traité a été acquis et reconstruit en 3D.

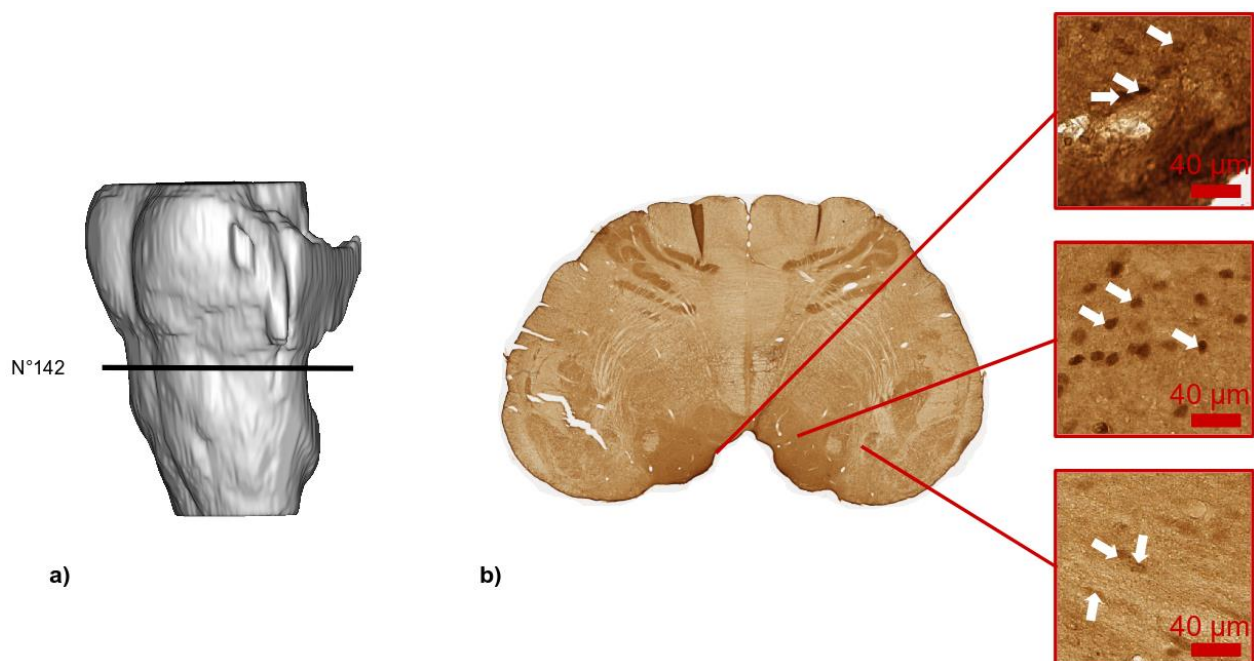


Figure 19 : Présentation des données anti-Phox2B. **a)** Enveloppe externe tridimensionnelle du tronc cérébral étudié et reconstruit en 3D. **b)** Diversité d'intensité de marquage et de densité des neurones étudiés sur la section numéro 142. On peut remarquer que contrairement à la Figure 17, l'intégralité du tissu est teinté en brun. Les flèches blanches indiquent quelques neurones d'intérêt marqués par l'anticorps dirigé spécifiquement contre Phox2b sur les images données en exemple.

7.3.2. Segmentations manuelles d'anti-Phox2B

L'extraction des images de la base de données a été faite de manière semi-aléatoire. Le nombre de neurones d'intérêt étant très faible par rapport au marqueur NeuN dans le cerveau, une première sélection de quinze grandes zones ont été extraites par un histologiste afin qu'elles contiennent des neurones d'intérêts. Pour chacune de ces quinze grandes zones, vingt images 256x256 pixels ont été tirées aléatoirement. Pour augmenter la proportion d'images avec les neurones d'intérêts, 212 images ont été sélectionnées par un histologiste.

Comme pour le marquage NeuN, ces 212 images ont été segmentées manuellement par un expert en quatre classes : neurone (marqueur spécifique anti-Phox2B), tissu (tissu non-marqué et marquage non-spécifique), fond (pas de tissu) et artéfacts (concentration de marquage non-spécifique, débris) (Figure 20). La base d'apprentissage et celle de test ont le même nombre d'images (106 images) et ont à peu près la même proportion de pixels marqués pour les quatre classes définies ci-dessus.

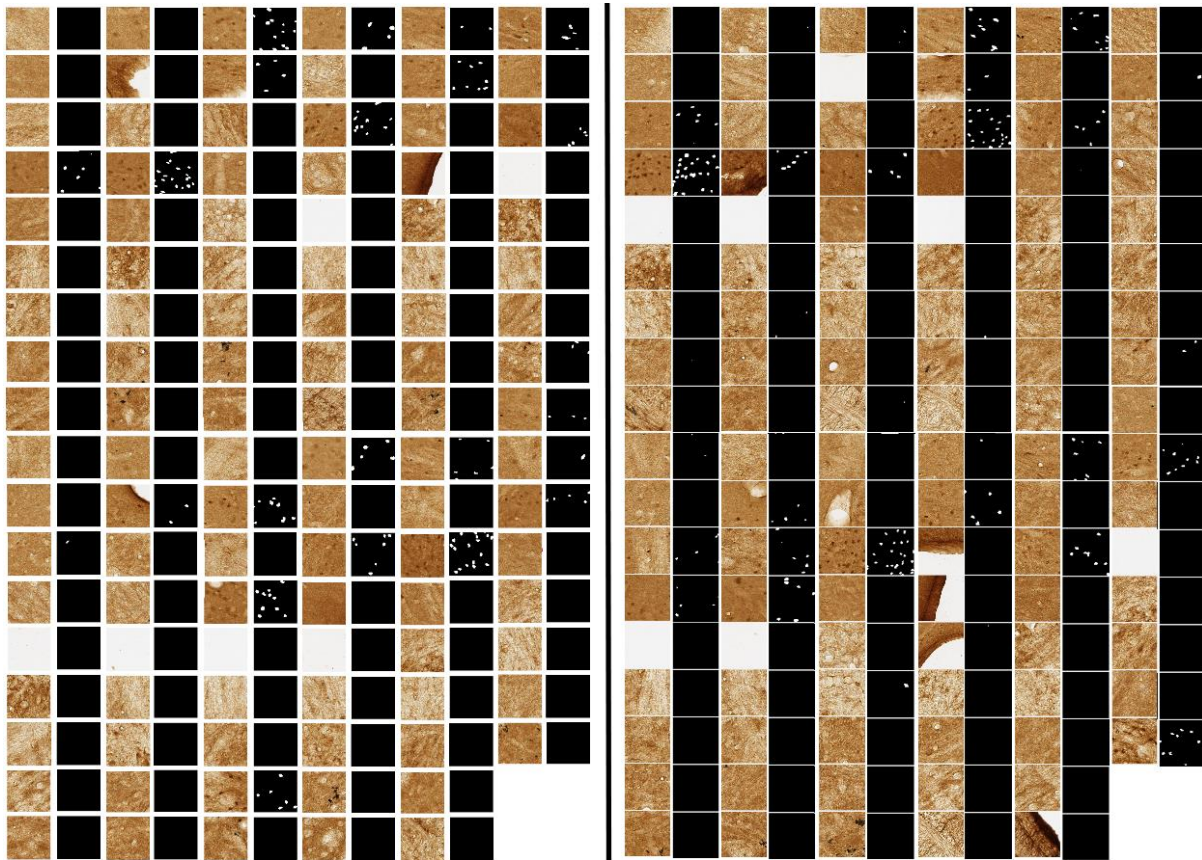


Figure 20 : Marquage anti-Phox2B, bases d'apprentissage (à gauche) et de test (à droite) : les données brutes sont associées avec les segmentations manuelles correspondantes. Les pixels blancs dans les segmentations manuelles sont les pixels considérés comme appartenant à la classe "neurone" par l'opérateur humain. Leur nombre est inférieur par rapport au marquage par NeuN.

7.4. Vecteur initial de caractéristiques utilisé pour NeuN et anti-Phox2B

Une fois les bases de données produites pour ce travail, le vecteur initial de caractéristiques est conçu. Ce vecteur sert de base pour les procédés de sélection décrits Sections 9 et 10. Ce vecteur est aussi utilisé pour les procédés de segmentation.

Le vecteur initial de caractéristiques contient l'ensemble des caractéristiques que nous avons précédemment recensé. Comme présenté dans la Section 5.3, les espaces colorimétriques et de description texturale peuvent décrire les informations de l'image à l'échelle du pixel. Le vecteur initial de caractéristiques contient l'ensemble de ces espaces afin de pouvoir définir leur pertinence et leurs propriétés pour un marquage donné. Les vecteurs initiaux utilisés pour analyser les marquages NeuN et anti-Phox2B sont identiques.

L'ensemble des espaces colorimétriques est inclus dans le vecteur initial, ce qui représente 4 espaces soit 12 caractéristiques. Comme introduit dans la Section 1.1, le rayon d'un neurone varie entre 2,5 et 15 μm [Andersen *et al.*, 2016]. Nous donc avons utilisé ces informations pour définir la taille des éléments structurants utilisés pour l'analyse de la texture (voir Section 8.4).

Comme présenté dans la Section 5.3.2, les filtres de Gabor sont des espaces avec cinq hyperparamètres différents qui doivent être définis. D'abord, le support gaussien est fixé avec un écart-type de 2 pixels (σ , sigma) et un support ellipsoïde de 1,5 pixel (γ , gamma). Ainsi, le lissage est minimal par rapport au rayon des neurones et le support n'est pas contraint par une forme circulaire, un neurone étant rarement parfaitement circulaire mais ellipsoïdal. Ensuite, tous les angles de la connexité 8 sont pris en compte (θ , θ est égal à 0° , 45° , 90° et 135°) avec aucune phase à l'origine (ψ , ψ est égal à 0°) et les longueurs d'ondes λ (λ , λ) sont de 2,5, 5, 10 et 20 pixels. Enfin, les parties réelles et imaginaires sont conservées. Les hyperparamètres des filtres de Gabor sont identiques pour les marquages NeuN et anti-Phox2B. Le support gaussien a un rayon largement inférieur aux deux rayons minimaux en terme de pixels (2 pixels pour des rayons minimaux de 10 et 5 pixels pour respectivement les marquages NeuN et anti-Phox2B). Les longueurs d'ondes sont choisies afin de décrire les plus petits éléments de textures. Nous avons donc un nombre de 32 filtres de Gabor.

Pour le marquage NeuN, la résolution planaire est de $0,22 \times 0,22 \mu\text{m}^2$, ce qui implique qu'un neurone a un rayon compris entre 11 et 68 pixels. L'image moyenne et l'image de variance sont calculées avec un élément structurant sous la forme d'un disque de 10 pixels de rayon afin d'éviter la perte des neurones les plus petits, l'image moyenne étant une image lissée. Pour les espaces descriptifs texturaux, nous choisissons le rayon minimal, moyen et maximal d'un neurone (11, 40 et 68 pixels de rayon correspondant respectivement à 2,42, 8,8 et 14,96 μm). Nous ajoutons aussi le double du rayon maximal (134 pixels correspondant à 29 μm) afin d'avoir des caractéristiques texturales décrivant l'ensemble des neurones de l'image afin de voir si un rayon optimal existe entre le rayon minimum et maximum étudiés. L'ensemble des caractéristiques texturales a une cardinalité de 102 avec 2 caractéristiques issues des images de moyenne et de variance, 4 motifs binaires locaux, 32 filtres de Gabor et 64 caractéristiques d'Haralick. Ceci représente 15 espaces de caractéristiques (soit 88% des espaces totaux). Le vecteur de caractéristiques initial a une taille de 114 (Tableau 3).

Pour le marquage anti-Phox2B, la résolution planaire est de $0,50 \times 0,50 \mu\text{m}^2$. Contrairement aux neurones marqués par NeuN, les neurones marqués avec l'anti-Phox2B ont un rayon estimé

apparent plus faible car NeuN marque également le cytoplasme. Néanmoins, nous avons procédé au même schéma de conception du vecteur initial de caractéristiques. Les rayons choisis en nombre de pixels sont de 5, 10, 20 et 40 correspondant respectivement à 2,5 ; 5 ; 10 et 20 μm . Le vecteur initial de caractéristiques a donc aussi une taille de 114 (Tableau 3).

Espace de caractéristiques	Paramètres pour le marqueur NeuN (en pixels/degrés)	Paramètres pour le marqueur anti-Phox2B (en pixels/degrés)
RVB	/	/
CIE XYZ	/	/
HSV	/	/
CIE L*a*b*	/	/
Moyenne et Variance	10	5
MBL	11, 40, 68, 134	5, 10, 20, 40
Haralick	11, 40, 68, 134	5, 10, 20, 40
Filtres de Gabor	Gamma : 1,5 Sigma : 2 Psi : 0 Theta : 0°, 45°, 90°, 135° Lambda : 2,5, 5, 10, 20	Gamma : 1,5 Sigma : 2 Psi : 0 Theta : 0°, 45°, 90°, 135° Lambda : 2,5, 5, 10, 20

Tableau 3 : Résumé des espaces de caractéristiques utilisés dans le cadre de ma recherche.

8. Développements logiciel réalisés

Ce chapitre est consacré aux choix d'implémentation faits et détaille des considérations techniques. L'objectif est de donner le cadre de réalisation et de développement des résultats de ce travail.

8.1. Matériel informatique utilisé

Les développements et les expérimentations ont été réalisés sur une station de travail Dell avec le système opérateur Ubuntu 14.04 LTS 64-bits, 2 processeurs Intel Xéon E-2630 v3 à 2,40 GigaHertz avec 16 cœurs de calcul (soit 32 threads de calcul avec l'*hyper threading*) et 128 Gigaoctets de mémoire vive.

Nous avons eu un accès privilégié au Très Grand Centre de Calcul (TGCC) du CEA (<http://www-hpc.cea.fr/fr/complexe/tgcc.htm>). Notamment, nous avons pu utiliser le supercalculateur Irène possédant 1656 nœuds de calculs Intel *Skylake* à 2,7 GigaHertz (AVX512) avec 48 cœurs de calculs et 192 Gigaoctets par nœud. Au total, 79 488 cœurs de calculs sont disponibles mais nous avons limité nos calculs à 4032 cœurs.

8.2. Environnement logiciel BrainVISA

BrainVISA est la plateforme logicielle collaborative permettant la gestion, l'analyse et la visualisation des données d'imagerie cérébrale [Rivière *et al.*, 2009]. La totalité des développements présentés dans ce manuscrit ont été réalisés sur cette plateforme. BrainVISA est organisée en plusieurs parties fonctionnant de manière indépendante. Plusieurs formats de fichiers sont acceptés par BrainVISA dont les principaux formats utilisés en imagerie biomédicale. Nous pouvons citer notamment le format CZI (format pyramidale de Zeiss), NIfTI-1, SVS et GIS. Ce dernier est le format maison utilisé dans la majorité des traitements de BrainVISA.

Aims (*Analysis of Images and Signals*) est la librairie de traitements de base de BrainVISA. Elle est consacrée à l'implémentation des traitements d'images les plus coûteux en termes calculatoires et est codée en C/C++. Parmi l'ensemble des méthodes implémentées que comporte Aims (302 programmes au mois d'octobre 2018), nous nous servons principalement de la conversion des fichiers, du recalage linéaire et non-linéaire, du seuillage et de l'algorithme BioVision dans le cadre de mon travail (voir Section 5.2.2).

SomaWorkflow est un logiciel qui gère les mécanismes de parallélisation et d'ordonnancement des procédés [Laguitton *et al.*, 2011]. Il permet de créer des chaînes de traitements séquentielles et parallèles intégrant les programmes Aims. SomaWorkflow, a été utilisé pour les développements parallélisables produits dans le cadre de mon travail.

Axon est l'interface entre les différentes parties de BrainVISA. Il est ainsi possible de combiner les tâches Aims avec la parallélisation et l'ordonnancement de SomaWorkflow. La subdivision des données massives est rendue possible grâce à la lecture et l'écriture partielles [Balbastre *et al.*, 2014]. Grâce à ces mécanismes d'accès partiel, seule une partie de l'image de taille compatible avec la taille de la mémoire vive disponible peut être chargée en mémoire. Une fois traitée, cette donnée est écrite dans le volume final qui aura été préalablement alloué sur le disque. Ce procédé est facilement parallélisable si l'on dispose d'architecture multi-cœur (station de travail, grappe de calcul). Il est important d'optimiser les dimensions des images traitées en fonction du nombre de cœurs de calculs et de la mémoire vive disponible en particulier lorsque des lames de calcul dans des infrastructures de calcul haute performance sont utilisées (Figure 21). En effet, chaque lame possède une quantité limitée de mémoire vive. Pour optimiser l'acheminement des données à traiter vers les cœurs de calcul, la quantité de mémoire consommée par les procédés ne doit pas excéder la limite de mémoire vive des lames. De plus, nous découpons la donnée selon la deuxième dimension de l'image afin que l'accès aux bandes de données soit contigu (c'est-à-dire que chaque pixel dans les bandes est enregistré consécutivement).

Enfin, Anatomist est le logiciel de visualisation des images pour tous les formats de BrainVISA. Ce logiciel permet aussi de réaliser des segmentations manuelles, de gérer différents référentiels spatiaux et de superposer des images en mode fusion.

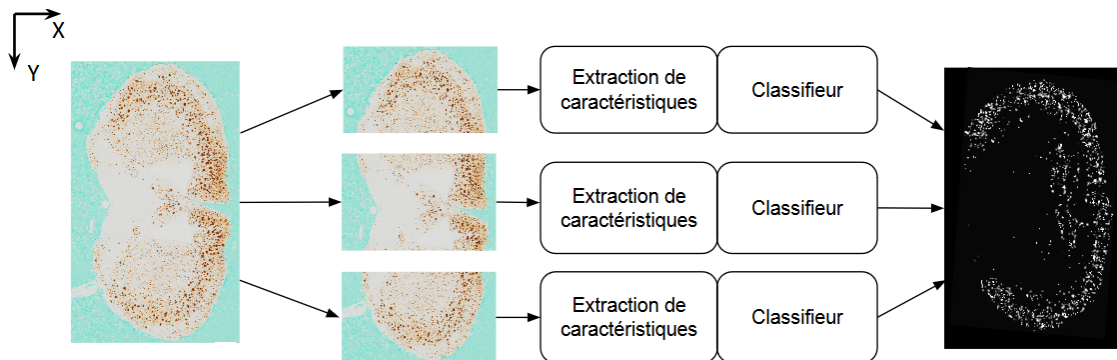


Figure 21 : Exemple d'utilisation des bibliothèques de lecture et d'écriture par accès partiels, pour la parallélisation de l'extraction de caractéristiques et de la segmentation. À partir de l'image brute, plusieurs zones contiguës sont extraites puis traitées séparément. L'image finale est produite lorsque les différents résultats de segmentation parallèles sont retournés et intégrés sur le disque.

8.3. Algorithmes de segmentation

BrainVISA intègre des méthodes de segmentation et de sélection de caractéristiques que j'ai utilisées et modifiées dans le cadre de ma thèse. Cette Section présente les implémentations des algorithmes d'apprentissage automatique.

8.3.1. BioVision

BioVision est une méthode de segmentation d'images histologiques couleur qui a été proposée par Chubb et associés en 2006 [Chubb *et al.*, 2006]. Elle repose sur une modélisation de la distribution des valeurs de caractéristiques des pixels par mixtures de gaussiennes. Un apprentissage supervisé est réalisé à partir de la segmentation d'experts et produit un modèle basé sur les caractéristiques de colorimétrie (R, V, B et image moyenne). Une implémentation de ce code en C/C++ a été réalisée à l'occasion du stage de Master 2 de Rémi Mariani et fait partie de la boîte à outils Aims (voir Section 5.2.2). Originellement, l'extraction de caractéristiques était intégrée dans l'étape d'apprentissage et de classification. J'ai redécoupé le code afin de la rendre plus modulaire.

Un algorithme de recuit simulé possédant plusieurs paramètres est utilisé par BioVision lors de l'étape d'apprentissage (voir Section 5.2.2). Le résultat de l'apprentissage est enregistré sous un format texte. L'avantage de ce format est qu'il condense l'information et qu'il est facile à lire et à écrire, contrairement au format PICKLE contenant les modèles des forêts aléatoires (voir Section 8.3.3).

Avec le modèle généré lors de l'étape d'apprentissage, la partie de généralisation de BioVision est capable de donner la probabilité de chaque pixel d'appartenir à chacune des classes apprises. La classe avec la plus forte probabilité devient la classe affectée au pixel. Cette étape est parallélisée en utilisant le même schéma que celui présenté dans la Figure 21.

8.3.2. Multi-seuillage

Dans de nombreuses études biologiques, la segmentation est réalisée par simple seuillage car les marquages histologiques sont souvent très spécifiques. L'opération définit un seuil qui sera utilisé sur l'ensemble des données. Nous avons décidé d'utiliser cette méthode

pour la comparer à notre stratégie de segmentation. Afin d'éviter de fixer nous-même le seuil, ce qui pourrait être source d'erreur (variabilités intra- et inter-opérateur), nous proposons de tester différents niveaux de seuils en couvrant toute la dynamique d'intensité de l'image convertie en niveaux de gris. Afin de réduire leur nombre, un échantillonnage a été réalisé en testant une valeur de seuil tous les 25 niveaux de gris (la dynamique étant l'écart entre le plus petit et le plus grand niveau de gris, [0 ; 255] correspondant à une dizaine de valeurs). Cette méthode est utilisée pour comparer les méthodes d'apprentissage automatique que nous proposons avec un seuillage optimisé pour la base d'apprentissage, un optimum unique pouvant être trouvé par cette méthode pour la détection de marqueurs histologiques.

8.3.3. Forêt aléatoire

Scikit-learn est une librairie Python spécialisée dans l'apprentissage automatique et la classification [Pedregosa et al., 2012]. Elle regroupe plusieurs méthodes d'apprentissage automatique et différentes métriques d'évaluation de la qualité de segmentation. L'étape d'apprentissage et de généralisation de la forêt aléatoire a été codée en Python dans BrainVISA à l'aide de *scikit-learn*, et parallélisée en utilisant le logiciel SomaWorkflow et les librairies d'Entrée/Sortie avec des accès partiels.

L'étape d'apprentissage comporte trois paramètres pour cette méthode (voir Section 5.2.3). La profondeur et le nombre d'arbres permettent d'ajuster la qualité et la stabilité de la segmentation. Pour les forêts aléatoires pondérées, le poids de la classe d'intérêt revêt une importance particulière quand la proportion des classes est fortement déséquilibrée. Nous en verrons un exemple avec la base de segmentation anti-Phox2B et l'optimisation de la forêt aléatoire associée (Section 11.2). Le modèle d'apprentissage est enregistré au format PICKLE, un format permettant d'enregistrer les objets Python.

Dans l'étape de généralisation, la parallélisation présentée Figure 21 a été implémentée. La première étape est de lire le modèle au format PICKLE. Puis la segmentation est réalisée en prenant en compte les différents paramètres (nombre d'arbres, profondeur d'arbre et pondération des différentes classes) définis pendant l'apprentissage.

8.4. Extraction de caractéristiques

Lors du démarrage de mon projet doctoral, la méthode de segmentation supervisée nommée BioVision (Section 8.3.1) était disponible dans le laboratoire avec un nombre de caractéristiques limité (R, V, B et image moyenne). Le programme intégrait le calcul des caractéristiques choisies ainsi que le procédé d'apprentissage au sein d'un même programme. Pour plus de généricité, ma première tâche a été de désolidariser l'étape d'extraction de caractéristiques des algorithmes d'apprentissage automatique. L'objectif était d'avoir un unique programme commun à toutes les méthodes. Ce programme est décrit Figure 22, il consiste en plusieurs extracteurs instanciés par le programme principal.

Les différentes caractéristiques sont enregistrées dans la quatrième dimension du format d'image utilisé (les trois premières dimensions servent à décrire les données spatiales). Cette dimension généralement exploitée pour coder le temps est inutilisée dans le cas des sections histologiques numérisées. L'extraction de caractéristiques est donc générique et utilisable sur des images bi- et tridimensionnelles. Nous pouvons également noter que le nombre de pixels est proportionnel aux nombre de caractéristiques. Dans le cas d'une section de cerveau de macaque numérisée de 40Gp, chaque caractéristique ajoute 40Gp au volume de données total.

Les caractéristiques sont toutes enregistrées sous la forme d'un nombre à virgule flottante simple précision. En effet, ce format permet de représenter aussi bien les nombres entiers, spécifiques aux espaces colorimétriques que des paramètres texturaux non entiers. Le choix d'une représentation et d'une dynamique commune pour toutes les caractéristiques extraites facilite les étapes d'apprentissage automatique.

L'extraction des caractéristiques, qui est une étape coûteuse en temps de calcul, peut être réalisée en parallèle avec les mécanismes d'accès partiels (Figure 21). Des bordures sont générées dans les images pour gérer les effets de bords se produisant lorsque le calcul de la caractéristique implique l'utilisation d'un élément structurant. Pour optimiser les temps d'accès mémoire, une sous-partie extraite à partir de l'image entière est une séquence de lignes consécutives ce qui produit une série de bandes à traiter. Pour chaque sous-partie, les valeurs des bordures sont remplies par les valeurs des sous-parties contiguës lorsque qu'elles sont incluses dans l'image d'origine. Pour les bordures externes de l'image initiale, la valeur des bordures est fixée par symétrisation de l'information interne à l'image. La taille de la bordure est égale à la taille du rayon du plus grand élément structurant.

Le premier extracteur concerne la couleur. Dans le cas d'une image codée dans l'espace RVB, l'extracteur est capable, d'une part, d'extraire les caractéristiques individuelles R, V et B et d'autre part de convertir ces caractéristiques RVB en caractéristiques HSV, XYZ et $L^*a^*b^*$ (Section 5.3.1). Il est aussi capable de convertir l'image en une image en niveaux de gris en faisant la moyenne de toutes les caractéristiques RVB. Cette image en niveaux de gris est utilisée par les autres extracteurs pour calculer les caractéristiques texturales, ce qui est fait traditionnellement dans les articles de référence [Haralick *et al.*, 1973][Ojala, 2002].

Le deuxième extracteur calcule la moyenne et la variance des niveaux de gris sur un voisinage autour d'un pixel traité. Le voisinage est défini comme étant l'ensemble des pixels inclus dans un élément structurant, dont les caractéristiques (taille et forme) sont définies par l'utilisateur. Les formes 2D-3D disponibles dans Aims sont diverses : croix, disque, cercle, carré, cube et sphère.

Le troisième extracteur est entièrement consacré à la génération et à l'application des filtres de Gabor. Les paramètres principaux sont les hyperparamètres de Gabor présentés Section 5.3.2. Ils se présentent sous forme de listes de paramètres, permettant l'extraction de plusieurs filtres de Gabor de manière consécutive.

Enfin, le quatrième extracteur est consacré aux dernières caractéristiques texturales présentées dans la Section 5.3.2. La forme des éléments structurants pour les caractéristiques d'Haralick et le MBL sont définis par les articles de référence [Haralick *et al.*, 1973][Ojala, 2002]. Néanmoins, il est possible d'analyser la texture de l'image en niveaux de gris à plusieurs échelles, en définissant plusieurs rayons.

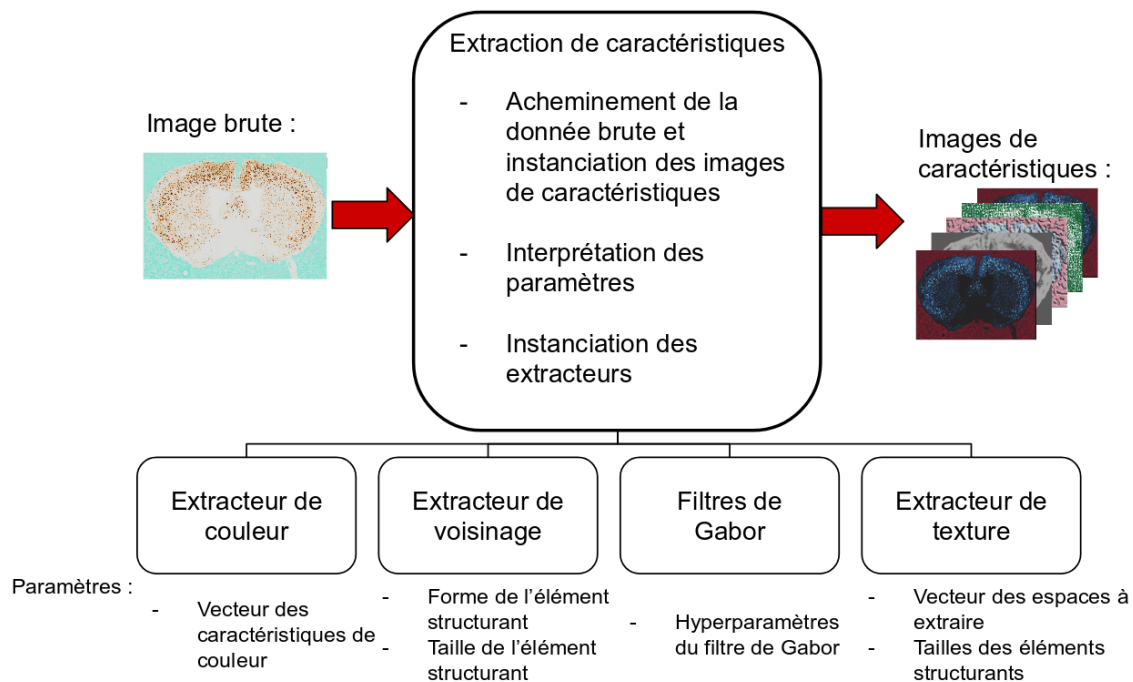


Figure 22 : Schéma de l'implémentation proposée pour l'extraction des caractéristiques.

9. Exploration des espaces de caractéristiques par Force Brute

Dans ce chapitre sont abordées la génération et l'analyse des résultats de la recherche par Force Brute. La méthode que j'ai proposée permet de comparer des milliers voire des millions de combinaisons de caractéristiques et d'évaluer les capacités de certains espaces pour segmenter les données étudiées. En plus de la méthode de génération des combinaisons par Force Brute, j'ai proposé de nouvelles métriques pour réaliser une analyse quantitative robuste de ces caractéristiques afin de les classer. Enfin, j'ai également proposé une nouvelle subdivision des espaces de caractéristiques texturales sous la forme de familles de caractéristiques.

9.1. Principe et objectifs de l'exploration par Force Brute

La Force Brute consiste à tester l'ensemble des combinaisons possibles pour résoudre un problème donné et est souvent considérée comme un moyen non-élégant de trouver la solution optimale [Heule et Kullmann, 2017]. Cette pensée est souvent étayée par des considérations de coût calculatoire, la Force Brute nécessitant des infrastructures de calcul permettant de réaliser plusieurs milliers voire millions d'opérations simultanément. De plus, la Force Brute est souvent remplacée avantageusement par des heuristiques de recherche permettant d'explorer efficacement un grand nombre d'éléments [Heule et Kullmann, 2017]. Dans notre cas et comme détaillé dans la Section 6, la Force Brute paraît être le meilleur moyen de proposer une sélection de caractéristiques stables de par son caractère exhaustif, ce qui n'est pas garanti par les autres méthodes [Meinshausen *et al.*, 2010][Awada *et al.*, 2012][Shah *et al.*, 2012][Kuncheva *et al.*, 2018].

Pour les deux bases de données, nous avons préliminairement extraits les 114 caractéristiques. La combinatoire totale de combinaisons de caractéristiques à explorer est donc de 2^{114} ce qui correspond à plusieurs centaines de milliards de milliards de combinaisons, soit

2.10^{34} . Une combinaison tirée à partir des 114 caractéristiques est traitée en minimum 40 minutes (de l'apprentissage au calcul du F-Score final permettant d'évaluer la segmentation automatique, le temps d'extraction de caractéristique n'est pas compté), le traitement avec 1 000 cœurs de calcul en parallèle (accessible avec des infrastructures de calcul haute performance) représenterait 10^{27} années de traitement, chaque combinaison (ou sous-ensemble) étant appris et classifié. Les segmentations résultantes sont évaluées par le calcul du F-Score, puis triées par ordre de F-Score décroissant, l'ordre des combinaisons résultant est appelé vecteur de positions.

Plus généralement, l'exploration exhaustive par Force Brute correspond à un nombre total de combinaisons de 2^N où N est la taille du vecteur initial de caractéristiques. En limitant la taille du vecteur de caractéristiques explorée, le nombre de combinaisons à calculer devient $\sum_{k=1}^M \binom{N}{k}$ avec M entier tel que $0 < M \leq N$.

J'ai donc décidé de réduire la combinatoire explorée. En limitant la taille des vecteurs à 3, il résulte 247 019 combinaisons à analyser, ce qui représente à peu près 10 jours de traitement avec 1 000 cœurs de calculs travaillant en parallèles. Néanmoins, avec une station de travail classique (dans le cadre de ce manuscrit, 30 cœurs en parallèles, 2 cœurs sont gardés en réserve), la combinatoire est encore trop grande (Figure 23). Ainsi, en limitant la taille des vecteurs à 2, le nombre de combinaisons à analyser est de 6 555 ce qui représente 5 jours avec 30 cœurs de calcul. Nous considérons comme raisonnable un calcul dont le temps de traitement est inférieur à une semaine.

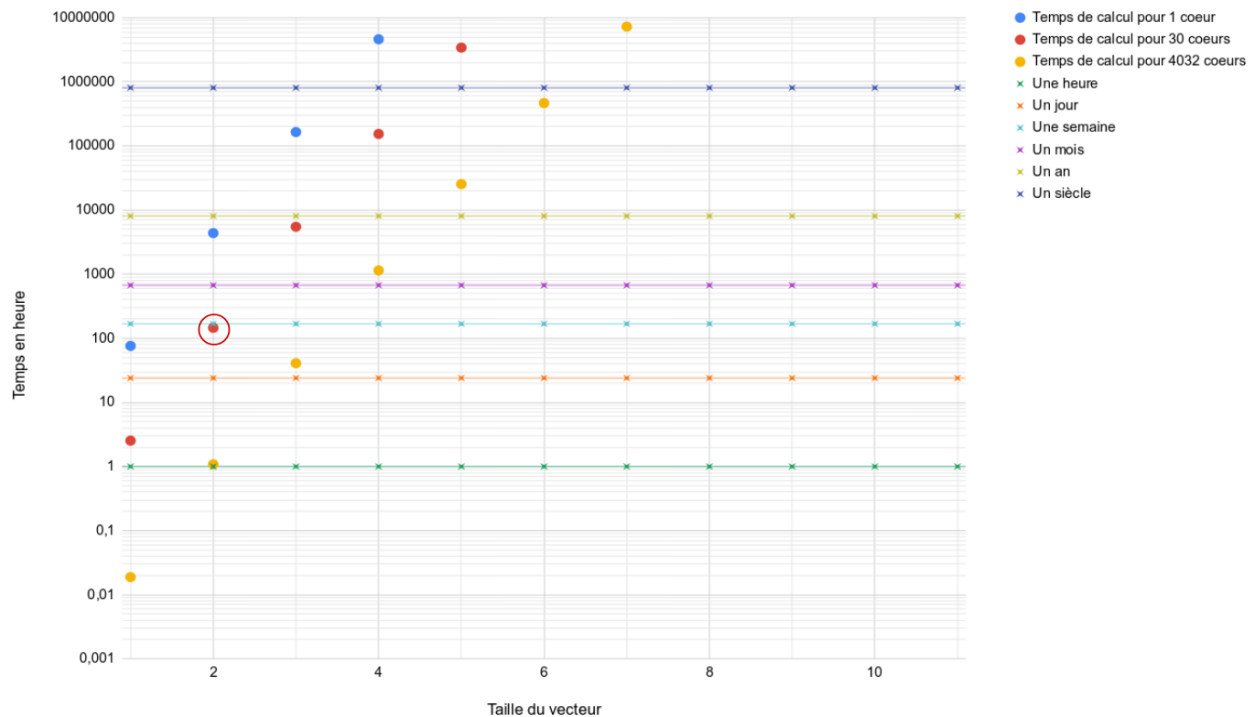


Figure 23 : Temps de traitement (échelle logarithmique) pour différentes conditions pour le vecteur de 114 caractéristiques. Les points ronds colorés représentent les différents temps de traitement pour les conditions exploitées. Les lignes représentent les différentes échelles de temps (de l'heure au siècle). Le cercle rouge représente la condition de traitement avec la limite de taille de vecteur choisie pour 30 cœurs et moins d'une semaine de calcul.

Ainsi, nous observons une infime quantité de la combinatoire en limitant la taille des vecteurs analysés. Pour exploiter au mieux l'information obtenue par l'exploration par Force Brute, je propose de subdiviser les espaces de caractéristiques ayant une grande cardinalité

(dans notre cas les espaces de description de la texture) en familles de caractéristiques. Il y a deux principaux intérêts à cette subdivision. Tout d'abord, une granularité intermédiaire est créée entre la caractéristique et l'espace de caractéristique. En effet, les espaces de caractéristiques regroupent des caractéristiques ayant des propriétés parfois très différentes les unes des autres. Par exemple, l'espace d'Haralick regroupe quatre statistiques (le second moment angulaire, le contraste, la variance et la corrélation) n'exprimant pas la même information. Le deuxième intérêt est de permettre la comparaison des différentes familles entre elles. À l'échelle des espaces de caractéristiques, il est difficile de comparer l'effet des 16 caractéristiques d'Haralick avec la caractéristique unique du Motif Binaire Local ou même aux trois caractéristiques des espaces colorimétriques. La Section 9.2.1 détaille la conception des familles de caractéristiques que je propose.

Une fois les caractéristiques regroupées en familles, j'ai également proposé plusieurs métriques afin de les évaluer. L'état de l'art de la sélection de caractéristique indique que les critères d'évaluation traditionnels comme le F-Score confèrent peu de stabilité aux combinaisons sélectionnées. J'ai choisi d'évaluer les caractéristiques selon leur présence dans les meilleurs résultats de segmentation selon le F-Score. L'idée est de bénéficier de la grande quantité de données fournie par l'exploration par Force Brute afin d'évaluer une caractéristique selon sa distribution dans les différentes combinaisons triées de la meilleure à la moins bonne en terme de F-Score. Nous proposons ainsi une alternative originale aux critères d'évaluation classiques permettant de juger la stabilité d'une caractéristique. Une caractéristique présente dans les meilleures combinaisons étant jugée plus stable qu'une caractéristique présente dans les pires combinaisons en termes de F-Score. La Section 9.2.2 et la Section 10.2.2 détaillent les différents critères mis en place dans ce travail de thèse.

La comparaison des familles de caractéristiques avec les métriques proposées va nous permettre de réaliser une première étape de sélection à l'échelle des familles de caractéristiques. Cette étape réduit le nombre de caractéristiques à considérer et de ce fait la combinatoire. Ainsi, une exploration par Force Brute avec une limite sur la taille des vecteurs plus grande est rendue possible. La seconde étape de sélection est réalisée cette fois à l'échelle des caractéristiques. Nous détaillons cette seconde étape de sélection dans la Section 10.

Nous avons mis en place la méthode d'exploration par Force Brute en tirant parti des outils et logiciels de BrainVISA comme SomaWorkflow ainsi que les outils de *scikit-learn*. Nous générons toutes les combinaisons possibles de caractéristiques à partir d'un vecteur initial à étudier. Les résultats sont compilés dans un tableau unique puis analysés par les méthodes détaillées dans la Section suivante (voir Figure 24). Cette exploration est aussi réalisée en validation croisée (Section 5.1) afin d'évaluer la stabilité des familles de caractéristiques.

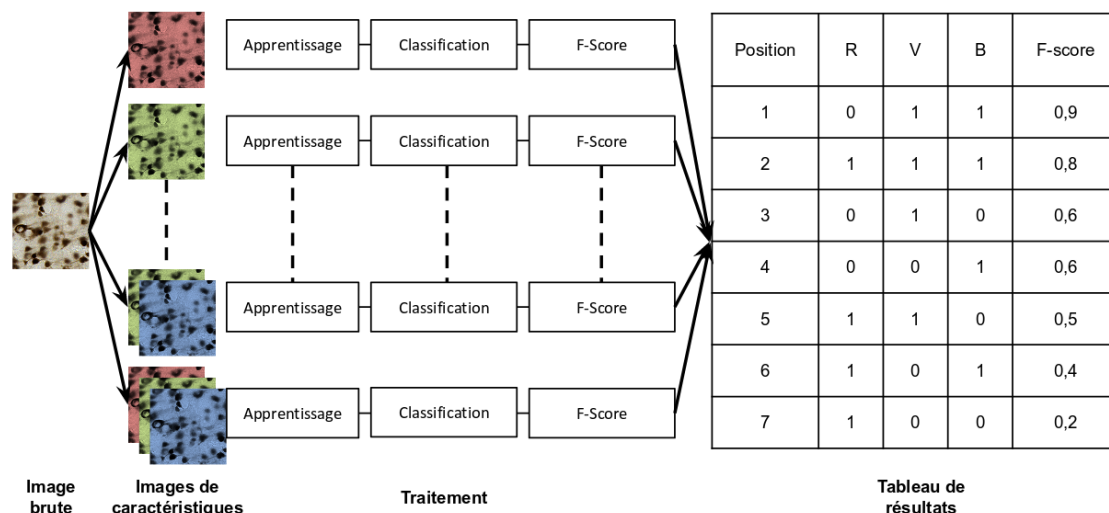


Figure 24 : Exemple de génération des résultats de l'exploration de Force Brute pour l'espace RVB uniquement. Sur le tableau final, le '0' représente l'absence et le '1' la présence d'une caractéristique. La meilleure combinaison est [V ; B] avec un F-Score de 0,9. Le tableau résultant classe dans l'ordre les combinaisons testées du grand F-Score au plus petit.

9.2. Méthode de sélection des familles de caractéristiques

9.2.1. Définition d'une famille de caractéristiques

Pour regrouper les caractéristiques sous le terme de famille, j'ai choisi de suivre plusieurs règles afin de construire ces combinaisons de manière logique. Tout d'abord, chaque famille regroupe des caractéristiques avec les mêmes paramètres d'échelle. Ce que nous appelons paramètres d'échelles sont les paramètres définissant l'échelle d'analyse sur laquelle elles ont été calculées. Pour Haralick et le MBL, il n'y a qu'un seul paramètre d'échelle qui est le rayon de l'élément structurant. Ce premier regroupement est donc directement équivalent aux espaces de caractéristiques pour le MBL et les caractéristiques d'Haralick. Pour les filtres de Gabor, ce sont tous les paramètres définissant le support gaussien et la longueur d'onde d'analyse.

Cette première étape forme des groupes de 16 caractéristiques d'Haralick (l'espace d'Haralick à une échelle spécifique) et 8 caractéristiques de Gabor. Nous allons de nouveau regrouper les caractéristiques mais cette fois selon leurs propriétés intrinsèques (propres à un deuxième regroupement de caractéristiques).

Les 16 caractéristiques d'Haralick restantes sont les quatre statistiques calculées sur quatre directions différentes. Nous appelons famille d'Haralick la subdivision d'un espace d'Haralick selon ses quatre statistiques (le second moment angulaire, le contraste, la variance et la corrélation). Les familles ainsi formées ont une cardinalité de 4.

Les 8 caractéristiques de Gabor restantes sont la partie réelle et imaginaire de quatre filtres de Gabor. Un seul hyperparamètre diffère entre ces quatre filtres, c'est l'angle d'analyse Θ . Néanmoins, je trouve plus judicieux de regrouper les caractéristiques selon le couple partie réelle et imaginaire car elles ont des propriétés différentes. La partie réelle est un détecteur de région et la partie imaginaire est un détecteur de contours [Mehrotra *et al.*, 1992]. Nous appelons familles de Gabor les groupes ainsi formés, elles ont une cardinalité de 4.

Le Tableau 4 résume la conception des familles de caractéristiques. Elles ont toutes des cardinalités similaires. Cette propriété des familles est importante afin de permettre leur comparaison entre elles. Les métriques utilisées pour cette comparaison sont détaillées dans la Section 9.2.2 et la Section 10.2.2.

Espace de caractéristiques	Famille de caractéristiques à une échelle spécifique	Cardinalité
RVB	Idem	3
HSV	Idem	3
CIE XYZ	Idem	3
CIE L*a*b*	Idem	3
Espace descriptif d'Haralick	Deuxième moment angulaire (A), Contraste (Con), Corrélation (Cor) ou Variance (V) des quatre MCNG	4
Filtres de Gabor	Partie imaginaire (I) ou partie réelle (R) dans les quatre directions du filtre de Gabor (les autres hyperparamètres sont fixés)	4
MBL	Idem	1

Tableau 4 : Tableau résumant les éléments et la taille des familles de caractéristiques par espace de caractéristiques.

9.2.2. Critères de sélection des familles de caractéristiques

Pour comparer les familles de caractéristiques entre elles, nous proposons d'utiliser les critères de la valeur médiane ou moyenne de positions (VMeP et VmoP respectivement) d'une famille de caractéristiques qui est la moyenne de l'ensemble des valeurs médianes ou moyennes de positions de ses éléments (Figure 25). Cette valeur est proche de zéro si la caractéristique est présente dans les meilleures combinaisons de caractéristiques au sens du F-Score.

Les derniers critères pris en compte pour les familles de caractéristiques sont leur cardinalité (Section 9.2.1). Ces critères jugent de la quantité de temps et de mémoire nécessaires. Plus ces critères ont des valeurs élevées, moins la famille sera intéressante à calculer pour l'objectif de minimisation des coûts qui a été fixé pour objectif de ce travail.

Ces critères fournissent des éléments pour sélectionner des familles de caractéristiques pertinentes pour notre problème de segmentation. Ce procédé de sélection est détaillé Section 9.2.3. Cette optimisation est nécessaire avant de réaliser une sélection optimale de caractéristique, l'exploration par Force Brute étant coûteuse en ressources de calcul (voir Section 9.2.3).

Position	R	V	B	F-score		Caractéristiques	R	V	B	
1	0	1	1	0,9	→	Vecteur de positions	2	1	1	
2	1	1	1	0,8			5	2	2	
3	0	1	0	0,6			6	3	4	
4	0	0	1	0,6			7	5	6	
5	1	1	0	0,5		Valeur médiane de position	5,5	2,5	3	Moyenne : 3,66
6	1	0	1	0,4		Valeur moyenne de position	5	2,75	3,25	Moyenne : 3,66
7	1	0	0	0,2						

Tableau de résultats

Analyse du tableau

Figure 25 : Exemple du calcul des valeurs médianes et moyennes de positions réalisé à partir des résultats de l'exploration par Force Brute pour les caractéristiques de l'espace RVB. Dans le tableau de résultats, le '0' indique l'absence et le '1' indique la présence d'une caractéristique. L'ensemble RVB est une famille de caractéristiques et ses valeurs médianes et moyennes de positions sont toutes les deux de 3,66.

9.2.3. Sélection des meilleures familles de caractéristiques

Nous calculons les moyennes globales des VMeP et des VMoP à l'échelle des caractéristiques et des espaces, nous les appelons valeurs de positions globales. Les caractéristiques seront considérées comme avantageuses si leur valeur de positions est inférieure à la valeur de position globale des caractéristiques. De la même manière, un espace sera considéré comme avantageux si la valeur de position est inférieure à la valeur de position globale des espaces.

Ainsi, nous pouvons procéder à une sélection des familles de caractéristiques donnant les meilleurs résultats afin de réduire la complexité algorithmique (Section 10). Une famille est sélectionnée lorsqu'elle est issue d'un espace avantageux et que l'intégralité de ses caractéristiques sont avantageuses. Nous prenons en compte l'union des familles respectant toutes ces propriétés dans les deux étapes de validation (l'ensemble des familles sélectionnées par validation directe et par validation croisée). Les familles sélectionnées fournissent le nouveau vecteur de caractéristiques réduit par rapport à l'initial. Il est utilisé dans le procédé de sélection de caractéristiques suivant.

10. Sélection de caractéristiques par Force Brute

La sélection de caractéristiques est opérée en ne conservant que les meilleures familles de caractéristiques issues de l'analyse de la première étape de sélection. Le but est de réduire la complexité algorithmique de la sélection afin d'augmenter la limite de taille des vecteurs à explorer. Je propose dans ce paragraphe un nouvel algorithme de sélection utilisant les critères proposés dans la Section 9.2.2 et 10.2.2 afin d'agréger les caractéristiques optimales dans une combinaison de caractéristiques de taille minimale.

10.1. Principe et objectifs de la sélection par Force Brute

À partir des familles sélectionnées dans la première étape (voir Section 9), entre 9 et 10 familles sont sélectionnées ce qui représente une vingtaine de caractéristiques pour les deux marquages (voir Section 12 et 13). Cette réduction permet d'augmenter la limite de taille des vecteurs explorés par FB. Je propose une méthode de sélection utilisant les résultats d'exploration par Force Brute sur les combinaisons des caractéristiques issues des familles sélectionnées.

Ma méthode de sélection utilise les VmeP et VMoP utilisées précédemment dans l'étape de sélection des familles. Je propose une méthode de sélection intégrée. Cette méthode agrège séquentiellement les meilleurs caractéristiques les unes après les autres selon leurs valeurs de positions (voir Section 10.2.1 pour l'algorithme). Cette manière de procéder permet à chaque caractéristique sélectionnée de ne conserver que les combinaisons possédant les caractéristiques déjà sélectionnées pour la prochaine itération. Cette stratégie réduit le nombre de combinaisons traitées à chaque itération et réduit également la complexité algorithmique.

À l'instar de la sélection des familles de caractéristiques par Force Brute, la complexité algorithmique de l'exploration par Force Brute est la même, à savoir $\sum_{k=1}^M \binom{N}{k}$ avec M entier tel que $0 < M \leq N$ où N est la taille du vecteur de caractéristiques initial. Néanmoins, l'algorithme de sélection a une complexité décroissante avec le nombre d'itérations. Sa complexité algorithmique est donc $\sum_{i=0}^{M-1} \sum_{k=1}^{M-i} \binom{N}{k}$.

Afin d'évaluer notre approche, je propose de comparer les combinaisons sélectionnées par notre méthodologie avec les combinaisons sélectionnées par deux autres méthodes de sélection de caractéristiques (Tableau 4). La première est le résultat de la sélection par Force Brute simple, la combinaison avec le meilleur F-Score est la combinaison sélectionnée. La deuxième est le résultat de l'importance des caractéristiques calculée par la méthode incluse dans l'apprentissage de la forêt aléatoire. Les X premières caractéristiques sont sélectionnées, X étant la taille du vecteur maximal exploré (notre limite d'exploration par Force Brute). Je propose d'utiliser un critère dérivé du facteur d'importance dans l'algorithme détaillé dans la Section 10.2.1. Ce critère est la moyenne des facteurs d'importance de chaque caractéristique.

Toutes les méthodes testées sont aussi utilisées en validation croisée afin de tester la stabilité des sous-ensembles sélectionnés.

Méthode de sélection	Critère de sélection	Combinaison sélectionné
Force Brute simple	F-Score	La combinaison avec le meilleur F-Score
Facteur d'importance (FA)	-	Les premières caractéristiques en termes d'importance
Algorithme de sélection proposé Section 10.2.1	Valeur médiane de position	Combinaison résultant
	Valeur moyenne de position	Combinaison résultant
	Moyenne des facteurs d'importances	Combinaison résultant

Tableau 4 : Les différentes méthodes de sélection de caractéristiques à comparer. La recherche par Force Brute simple est la plus élémentaire des méthodes intégrée et le facteur d'importance est la méthode incluse dans l'apprentissage de la forêt aléatoire. Notre algorithme de sélection ainsi que les critères utilisés sont détaillés Section 10.2.

10.2. Méthode de sélection des caractéristiques

10.2.1. Algorithme de la méthode proposée

Nous proposons un algorithme de sélection inspirée de la stratégie de Sélection Séquentielle Agrégante (SSA) présentée Section 6.4. Le principe est d'agréger progressivement les caractéristiques pour former une combinaison optimale. Contrairement aux stratégies SSA, notre condition d'arrêt est déterminée par l'algorithme et pas par un hyperparamètre fixé arbitrairement.

À partir de la liste des combinaisons issues de l'exploration par Force Brute, les caractéristiques sont évaluées par un des critères détaillé dans la Section 10.2.2. La meilleure caractéristique au sens du critère est sélectionnée et ajoutée à la combinaison. Si le F-Score calculé avec la nouvelle combinaison augmente, l'algorithme continue le processus de sélection. Seules les combinaisons de plus grandes tailles incluant la combinaison courante sont considérées pour la suite du procédé. L'algorithme s'arrête lorsque le F-Score décroît pour la première fois. La dernière caractéristique ajoutée n'est alors pas prise en compte et la combinaison de sélection résultant devient la combinaison optimale (Figure 26).

Si la condition d'arrêt n'est pas rencontrée, le procédé est itéré jusqu'à son terme. Il ne reste alors plus qu'une combinaison à tester à la dernière itération. Le critère de sélection n'est donc plus la VMeP, la VMoP ou la moyenne des facteurs d'importance (MFI voir Section 10.2.2) mais le F-Score (une seule combinaison par caractéristique). La combinaison considérée comme sélectionnée est donc celle de l'avant dernière itération (où la VMeP, la VMoP et la MFI sont utilisées).

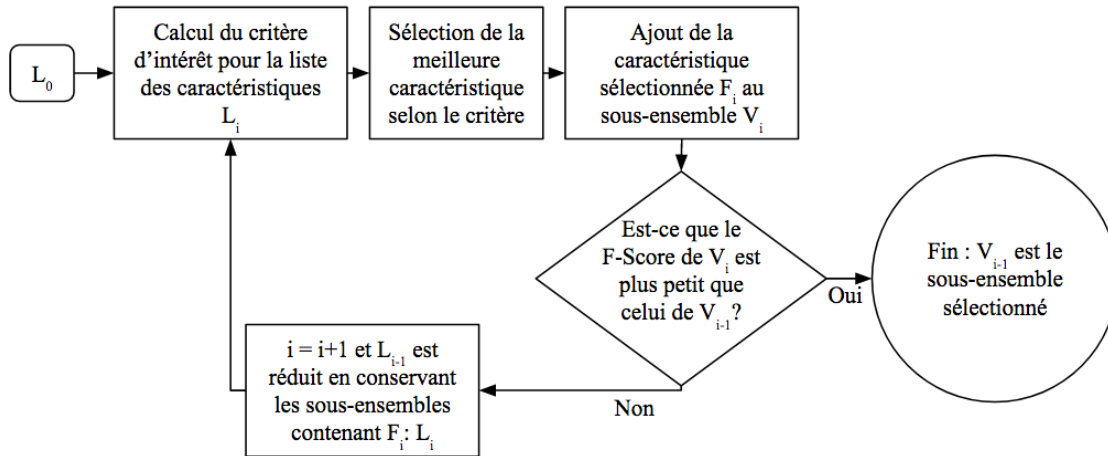


Figure 26 : Algorithme de sélection proposé. L_0 est l'ensemble des résultats de la recherche par Force Brute sur les familles sélectionnées (voir Section 9) avec une taille maximale de 4. L_i est la liste des combinaisons restantes à l'itération i . F_i est la meilleure caractéristique selon notre critère à l'itération i . V_i est la combinaison de sélection à l'itération i (à la première itération, V_0 est vide).

Le but de mon algorithme est de grouper séquentiellement des caractéristiques travaillant en synergie selon un critère donné et de s'arrêter à la première caractéristique non-informative ajoutée. Notre hypothèse est que les caractéristiques travaillant en synergie tendent à mieux décrire la donnée.

10.2.2. Critères de sélection de caractéristiques

Nous utilisons les critères de VMeP et VMoP que j'ai proposé et décrit dans la Section 9.2.2. De plus, un autre critère résultant de la sélection de caractéristiques incluse dans l'apprentissage de la forêt aléatoire est la moyenne des facteurs d'importance pour chaque caractéristique. Plus cette moyenne est grande, plus la caractéristique a eu une importance lors de l'apprentissage de la forêt aléatoire. Bien que les facteurs d'importance seuls aient été utilisés comme critères de sélection [Yaquib *et al.*, 2014], nous ne connaissons pas de travaux ayant utilisé la moyenne des facteurs d'importance sur plusieurs combinaisons de caractéristiques pour analyser sa pertinence en terme de stabilité. C'est ce que je propose de réaliser dans le cadre de ce travail (Figure 27).

Des trois critères décrits précédemment, deux ensembles sont discernables. Le premier est l'ensemble des valeurs médianes et moyennes de positions à minimiser. Ces critères ont été proposés afin de quantifier la présence ou non d'une caractéristique dans les meilleurs résultats. Le deuxième ensemble regroupe uniquement la moyenne des facteurs d'importances. Ce critère nous permettra d'utiliser les facteurs d'importances de toutes les combinaisons afin de conclure sur la stabilité de la méthode de sélection incluse dans la forêt aléatoire.

Position	R	V	B	F-score
1	0	0,8	0,2	0,9
2	0,1	0,6	0,3	0,8
3	0	1	0	0,6
4	0	0	1	0,6
5	0,3	0,7	0	0,5
6	0,5	0	0,5	0,4
7	1	0	0	0,2

Tableau de résultats

Caractéristiques	R	V	B
Vecteur de positions	2	1	1
	5	2	2
	6	3	4
	7	5	6
Valeur médiane de position	5,5	2,5	3
Valeur moyenne de position	5	2,75	3,25
Moyenne des facteurs d'importances	0,475	0,775	0,5

Analyse du tableau

Figure 27 : Exemple des critères calculés pour l'exemple présenté Figure 24 avec la moyenne des facteurs d'importance. Contrairement à la Figure 25, chaque facteur d'importance dont la valeur est différentes de '0' est conservé dans le Tableau de résultats. Dans ce tableau, le '0' indique l'absence et un facteur d'importance non nul indique la présence d'une caractéristique. Par exemple dans la première combinaison, R est absent, V a un facteur d'importance de 0,8 et B de 0,2 respectivement.

IV Mise en œuvre et validation de la méthode

11. Comparaison des méthodes de segmentation

Cette étude compare les résultats des forêts aléatoires avec ceux de BioVision et du multi-seuillage. Le but est de comparer les trois algorithmes présentés pour la segmentation de tissu marqué.

11.1. Résultats obtenus avec un marqueur spécifique : NeuN

Le vecteur initial de 114 caractéristiques détaillé Section 7.4 a été utilisé pour entraîner une forêt aléatoire et l'algorithme BioVision. Le multi-seuillage (Section 8) a été utilisé sur les images des deux bases NeuN converties en niveaux de gris. Les résultats produits avec BioVision avec le vecteur de 114 caractéristiques n'étaient pas concluant, BioVision n'étant pas adapté aux caractéristiques texturales. Le vecteur de caractéristiques RVBM qui contenait trois caractéristiques colorimétriques et l'image moyenne proposé par Chubb et associés fonctionne mieux que le vecteur initial que nous avons pu tester (Tableau 5).

F-Score	BioVision avec RVBM [Chubb <i>et al.</i> , 2006]	Seuillage au niveau 200	Forêt aléatoire avec le vecteur initial de caractéristiques
Validation directe	0,82	0,86	0,88
Validation croisée	0,84	0,85	0,89

Tableau 5 : F-Scores de la classe "tissu marqué" (neurones) pour BioVision, les forêts aléatoires et le seuillage optimal pour le marqueur NeuN. Le vecteur de 114 caractéristiques a été utilisé pour les forêts aléatoires et le vecteur RVBM pour BioVision. Les segmentations en validation directe et croisée sont semblables pour le marquage NeuN.

L'algorithme des forêts aléatoires a été testé avec des valeurs par défaut de ses hyperparamètres : une profondeur d'arbres de 4, 100 arbres de décisions et le même poids pour toutes les classes. Nous avons comparé les résultats de segmentation de cette forêt aléatoire avec les résultats du multi-seuillage sur les images de la base en niveaux de gris comme expliqué Section 8.3.2 (Figure 28). Les multi-seuillages sur la base d'apprentissage (Seuil VC) et sur la base de test (Seuil VD) ont donné des résultats inférieurs en termes de F-Score par rapport aux segmentations réalisées avec la forêt aléatoire en validation directe (FA VD) et en validation croisée (FA VC). Le meilleur résultat de seuil pour le multi-seuillage de la base d'apprentissage et de test était le seuil au niveau de gris 200, les scores étant de 0,86 et 0,85 respectivement. Les F-Scores de la forêt aléatoire étaient de 0,88 en validation directe et 0,89 en validation croisée. La forêt aléatoire représentait une augmentation de F-Score de 6,2% en moyenne par rapport à BioVision et une augmentation de 3,3% en moyenne par rapport au multi-seuillage. Les différents F-Score sont présentés dans le Tableau 5. Comme nous pouvons le voir Figure 29, la segmentation par BioVision avec le vecteur initial RVBM n'a pas permis de segmenter les neurones les plus clairs.

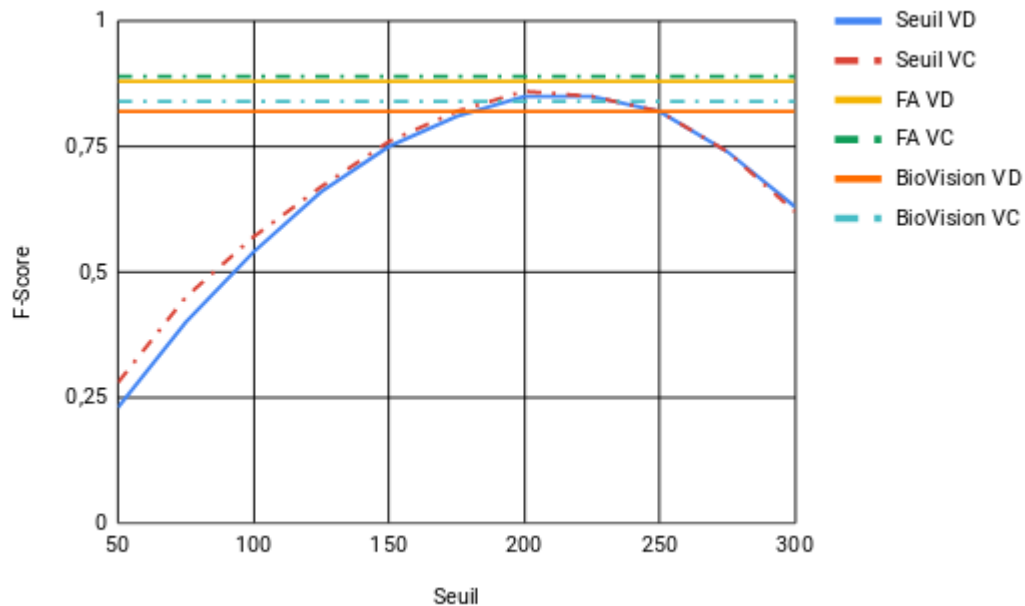


Figure 28 : Comparaison entre le multi-seuillage (Seuil) de la base d'apprentissage et la base de test, de BioVision en validation directe (VD) et en validation croisée (VC) et la forêt aléatoire (FA) en validation directe (VD) et la validation croisée (VC) pour le marqueur NeuN.

Comme nous pouvons le voir Figure 29, la segmentation avec BioVision est moins performant que la forêt aléatoire pour détecter les neurones plus clairs présents sur la deuxième image.

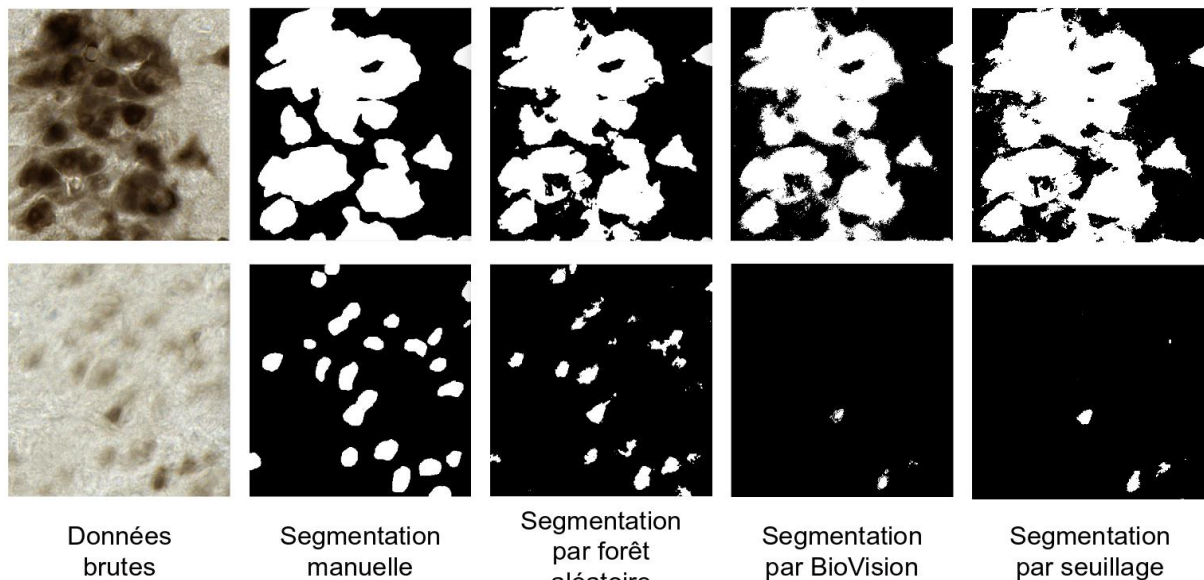


Figure 29 : Exemple d'images marquées avec l'anticorps NeuN extraites dans deux sous-régions différentes (l'hippocampe et le corps grenouillé latéral) avec les segmentations manuelles et automatiques.

11.2. Résultats obtenus avec un marqueur non-spécifique : anti-Phox2B

Le vecteur initial de caractéristiques détaillé Section 7.4 a été utilisé pour entraîner une forêt aléatoire et l'algorithme BioVision. Le multi-seuillage a été utilisé sur les images des deux bases converties en niveaux de gris (Section 8). Les bases utilisées ont été marquées à l'anti-Phox2B (Section 7.3).

Les résultats obtenus avec le marqueur anti-Phox2B sont similaires à ceux obtenus avec le marqueur NeuN. BioVision n'a pas fonctionné avec le vecteur de 114 caractéristiques et la forêt aléatoire a obtenu des meilleurs F-Score finaux que BioVision avec le vecteur RVBM et le multi-seuillage (voir Tableau 6). Entre les F-Scores résultant de la validation directe (VD et la validation croisée (VC), un écart de 10% est observé.

F-Score	BioVision avec RVBM	Seuillage au niveau 150	Forêt aléatoire de base avec le vecteur initial de caractéristiques	Forêt aléatoire optimisée avec le vecteur initial de caractéristiques
Validation directe	0,12	0,15	0,49	0,65
Validation croisée	0,21	0,14	0,42	0,58

Tableau 6 : F-Scores de la classe "neurone" pour BioVision, le seuillage optimal et les forêts aléatoires sans et avec optimisation pour le marqueur anti-Phox2B. Le vecteur de 114 caractéristiques a été utilisé pour les forêts aléatoires et le vecteur RVBM pour BioVision.

Néanmoins, les F-Scores étant inférieurs à 0,5 pour les hyperparamètres de base. Il a donc été décidé d'optimiser les hyperparamètres de la forêt aléatoire afin d'augmenter les valeurs de F-Score dans les deux conditions de validation. Tout d'abord, la profondeur d'arbre a été augmentée jusqu'à atteindre un plateau à partir d'une profondeur de 10 (Figure 30). Puis le poids de la classe "neurone" a été augmenté car cette classe est très minoritaire dans les images. Un optimum a été trouvé lorsque son poids était égal à 2. Il a donc été choisi de fixer la profondeur d'arbre à 10, le nombre d'arbres à 100 (le nombre d'arbre changeant peu les F-Score résultant) et le poids de la classe "neurone" à 2. Le nouveau F-Score calculé en validation directe était de 0,65 et celui en validation croisée de 0,58. Cela représentait une augmentation de 26% du F-Score grâce à l'optimisation.

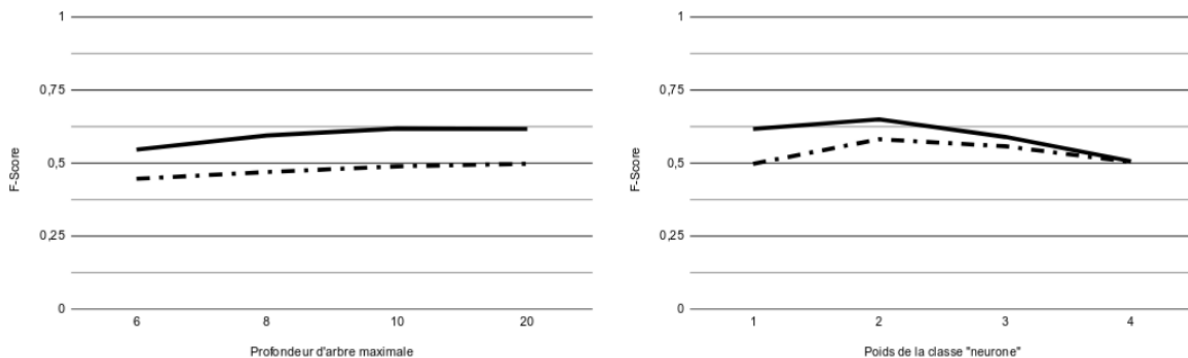


Figure 30 : Optimisation des paramètres de la forêt aléatoire pour la segmentation du marqueur anti-Phox2B.

Une fois optimisée, la forêt aléatoire présente une augmentation de F-Score d'un facteur 4 en moyenne par rapport à BioVision et au multi-seuillage, dont le seuil optimal était 150 (Figure 31). Contrairement aux images marquées avec l'anticorps NeuN, les images anti-Phox2B sont très difficiles à segmenter (Figure 32). On peut voir des effets de bruit de segmentation sur certaines images et une partie des neurones n'est pas détecté, ce que l'on peut voir dans la deuxième image présentée Figure 32. En effet, le marqueur n'étant pas spécifique et se fixant dans quasiment tout le tissu (Section 7.4), les algorithmes de segmentation ont beaucoup de difficultés à détecter correctement les éléments de la classe "neurone". La forêt aléatoire optimisée ici sera utilisée pour l'ensemble des étapes de sélection pour le marqueur anti-Phox2B.

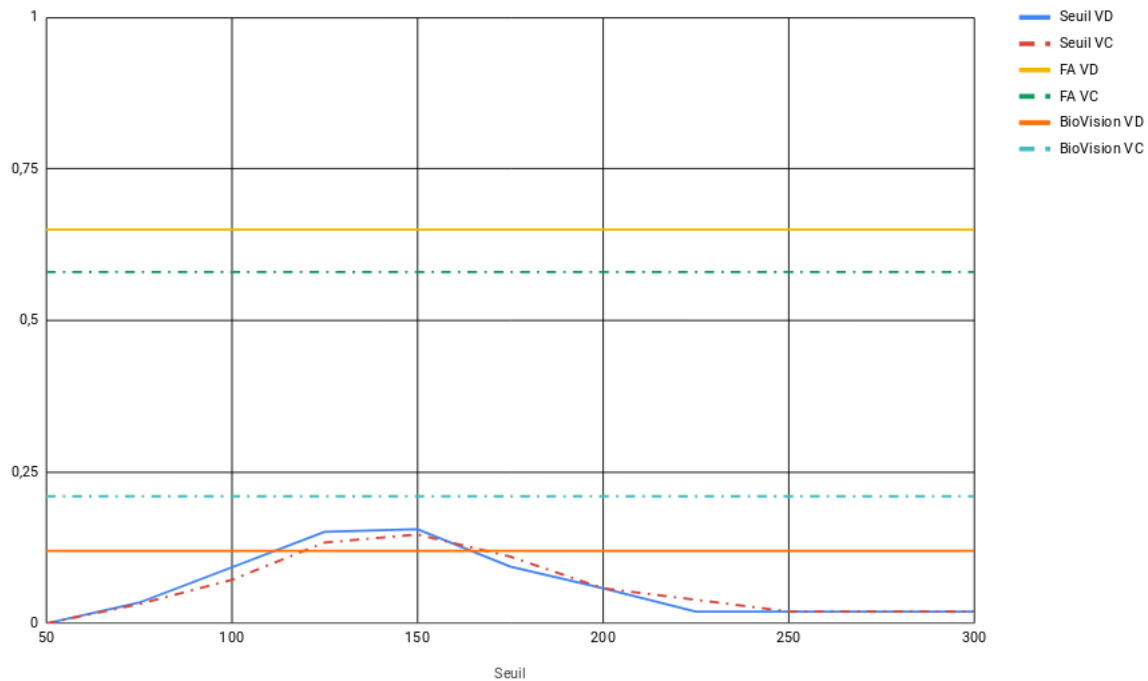


Figure 31 : Comparaison des résultats obtenus entre le multi-seuillage de la base d'apprentissage et la base de test, BioVision dans les deux conditions de validation et la forêt aléatoire en validation directe et en validation croisée pour le marqueur anti-Phox2B.

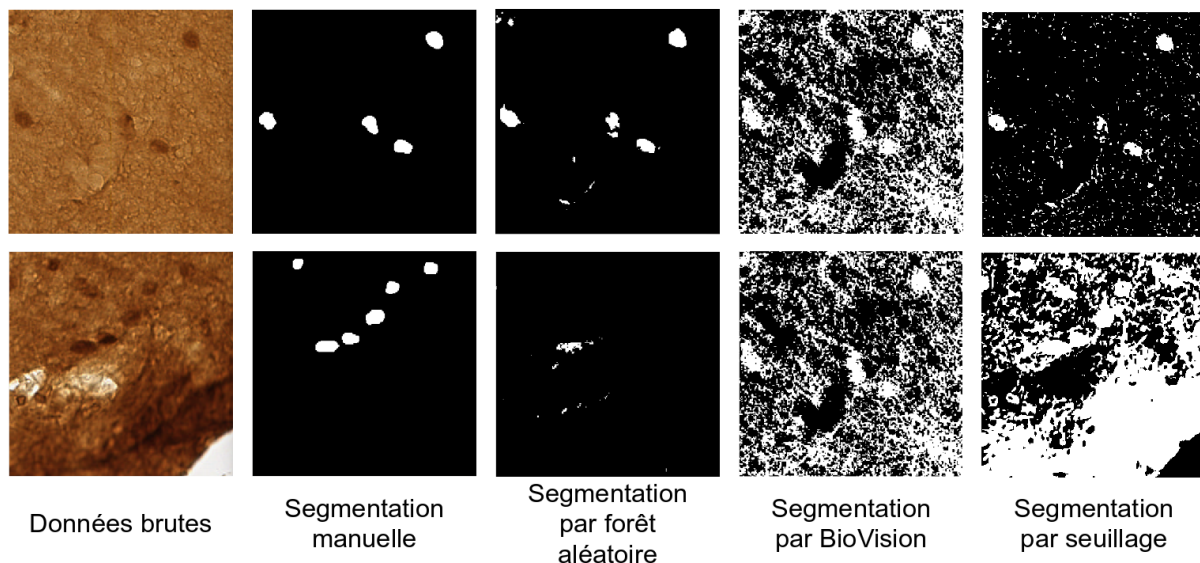


Figure 32 : Exemple d'images marquées à l'anti-Phox2B avec les segmentations manuelles et automatiques.

La forêt aléatoire semble être la méthode indiquée pour segmenter au mieux les neurones marqués à l'anti-Phox2B. Pour la suite de ce manuscrit, la forêt aléatoire sera utilisée pour segmenter les deux marquages présentés.

12. Sélection de caractéristiques pour le marqueur spécifique NeuN

12.1. Rappel de l'objectif de la segmentation

Le but de l'étude est de segmenter la classe "tissu marqué" correspondant aux neurones marqués par le marqueur NeuN. L'objectif est d'avoir un minimum de caractéristiques pour une qualité de segmentation proche de la qualité de la segmentation résultante du vecteur initial de caractéristiques. Cependant, l'utilisation des 114 caractéristiques n'est pas envisageable pour traiter des données massives. Cela nous amène à chercher un vecteur réduit de caractéristiques pour rendre possible la segmentation. Sur les 114 caractéristiques utilisées pour segmenter le marqueur NeuN, lesquelles sont les plus pertinentes pour la segmentation ? Si ces caractéristiques existent, forment-elles un sous-ensemble stable ?

L'étude que nous proposons est multi-échelle, nous nous intéresserons tout d'abord à l'ensemble des familles de caractéristiques que nous avons introduites dans la Section 9.2.1. Les critères basés sur la position des caractéristiques sont considérés pour classer les familles. Les meilleures familles sélectionnées forment un nouveau vecteur sur lequel une nouvelle étape de sélection est appliquée (Section 10).

12.2. Sélection basée sur la Force Brute

12.2.1. Familles sélectionnées pour NeuN

La première étape de sélection est appliquée aux familles de caractéristiques sur le vecteur initial de 114 caractéristiques avec une profondeur d'exploration de 2 (la limite de taille de vecteur est de 2). Nous avons proposé deux critères afin de trier les caractéristiques selon leur densité de présence dans les meilleures combinaisons : la valeur médiane de positions (VMeP) et la valeur moyenne de positions (VMoP) (Section 9.2.2). La moyenne des facteurs d'importance est utilisée lors de la deuxième étape de sélection (à l'échelle des caractéristiques).

Une fois les 114 caractéristiques et les 34 familles triées, la distribution entre la valeur médiane et la valeur moyenne correspondante pour chaque élément a été évaluée (Figure 33). Pour définir un sous-ensemble des caractéristiques présentant les plus petites valeurs médianes de positions, nous avons séparé l'ensemble des valeurs par la valeur moyenne des 114 VMeP. Pour les valeurs inférieures à la moyenne des VMeP, la dispersion est quasiment nulle avec une corrélation globale de 0,93 en validation directe et de 0,92 en validation croisée à l'échelle des caractéristiques. Pour les familles de caractéristiques, nous avons observé un comportement très similaire avec une corrélation globale de 0,96 dans les deux conditions de validation. Pour étudier la densité de représentation des caractéristiques dans les meilleures combinaisons, la VMeP et la VMoP sont équivalentes pour le marqueur NeuN. Cependant la médiane sera préférée car sa dynamique est plus grande et permet d'être plus discriminant par rapport à celle de la moyenne (la dynamique étant la différence moyenne de la différence entre la plus grande et la plus petite valeur dans les deux conditions de validation). La dynamique de la VMeP est de 5 138,5 et celle de la VMoP de 4 064,44.

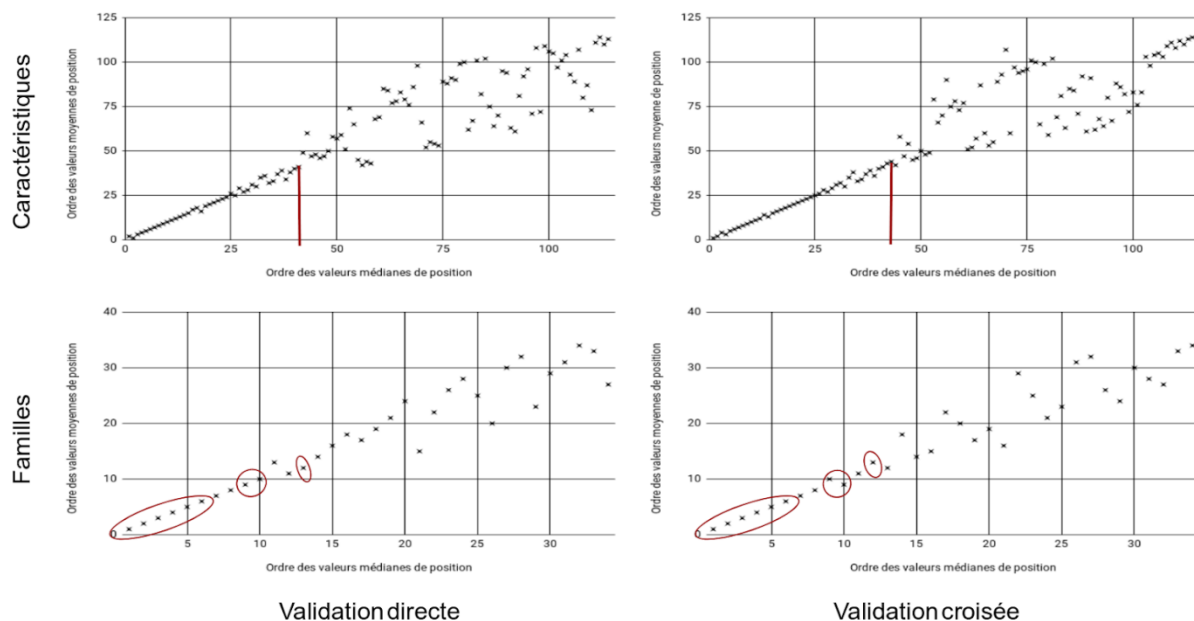


Figure 33 : Corrélation de la distribution des valeurs médianes et moyennes de positions pour le marqueur NeuN. À l'échelle des caractéristiques, la barre rouge représente la moyenne de toutes les valeurs médianes de positions. À l'échelle des familles, les familles entourées en rouge sont les familles sélectionnées (voir Figures 34 et 35).

Pour sélectionner les meilleures familles de caractéristiques, celles-ci ont été représentées sous la forme d'un tableau de distribution en considérant les espaces et les caractéristiques. Les espaces et les caractéristiques ont été triés de façon décroissante de la valeur médiane de positions, puis les familles ont été représentées par rapport à ces axes discrets. La dispersion, la répartition des caractéristiques et la prévalence dans les meilleures combinaisons de chaque famille sont visibles graphiquement. Comme expliqué Section 9.2.3, les moyennes des valeurs médianes de positions sont calculées à l'échelle des espaces et à l'échelle des caractéristiques (deux axes) pour définir les meilleures candidates. Les familles sélectionnées sont celles dont les caractéristiques ont toutes une VMeP inférieure à la moyenne à l'échelle des caractéristiques et issue des espaces ayant une VMeP inférieure à la moyenne à l'échelle des espaces. L'annexe 2 contient les tableaux numériques des classements des familles de caractéristiques par ordre de moyenne des VMeP de leurs caractéristiques. Plus la famille est en haut à gauche de notre graphique plus la famille est considérée comme pertinente pour la segmentation du marqueur NeuN. Ce graphique a été produit pour chaque condition de validation (Figures 34 et 35).

Ainsi, les familles incluses en validation directe étaient dans l'ordre de leur prévalence dans les meilleures combinaisons (la nomenclature détaillée Annexe 2 est entre parenthèse) :

- la partie réelle du filtre de Gabor avec un lambda de 20 pixels (G_20_R)
- la partie réelle du filtre de Gabor avec un lambda de 10 pixels (G_10_R)
- CIE XYZ
- RVB
- l'image moyenne (Moyenne)
- HSV
- MBL avec un rayon de 11 pixels (MBL 11)

En validation croisée, les familles de caractéristiques sélectionnées sont :

- la partie réelle du filtre de Gabor avec un lambda de 20 pixels (G_20_R)
- la partie réelle du filtre de Gabor avec un lambda de 10 pixels (G_10_R)
- CIE XYZ
- RVB
- l'image moyenne (Moyenne)
- HSV
- MBL avec des rayons de 40, 11 et 68 pixels (MBL 11, MBL 40 et MBL 68)

Aucune famille d'Haralick n'a été sélectionnée dans les deux conditions de validation. Néanmoins, les composantes d'une famille d'Haralick (la corrélation d'Haralick pour un rayon de 11 pixels) avaient toutes une VMP inférieure à la moyenne à l'échelle des caractéristiques (les caractéristiques de cette famille sont considérée comme avantageuses).

L'intégralité des familles sélectionnées en validation directe ont été retrouvées en validation croisée avec en plus deux familles de MBL. L'union des familles sélectionnées a intégré 9 familles dont 7 familles étaient communes aux deux conditions de validation (77% des familles). Sur ces 9 familles, nous avons recensé 3 familles colorimétriques et 6 familles texturales, ce qui correspond à 9 caractéristiques colorimétriques et 12 caractéristiques texturales. Ces 21 caractéristiques ont donc été sélectionnées pour l'étape suivante de sélection ce qui représentait 18,42% de l'ensemble des caractéristiques de départ. Il a été ainsi possible

de réduire la complexité algorithmique totale à 2^{21} soit 2 097 152 combinaisons (division du nombre de combinaisons total à explorer d'un facteur 1025). Malgré ce gain significatif, il était nécessaire de réduire encore ce nombre de caractéristiques pour obtenir des temps de traitements acceptables (entre 1 et 2 semaines). À partir des 21 caractéristiques restantes, toutes les combinaisons jusqu'à une taille maximale de 4 des vecteurs ont été explorées (soit 7 542 combinaisons de caractéristiques, 20 349 combinaisons devraient être ajoutées si la taille maximale étaient de 5). Cette étude correspond à la deuxième étape de la méthodologie de sélection que nous proposons.

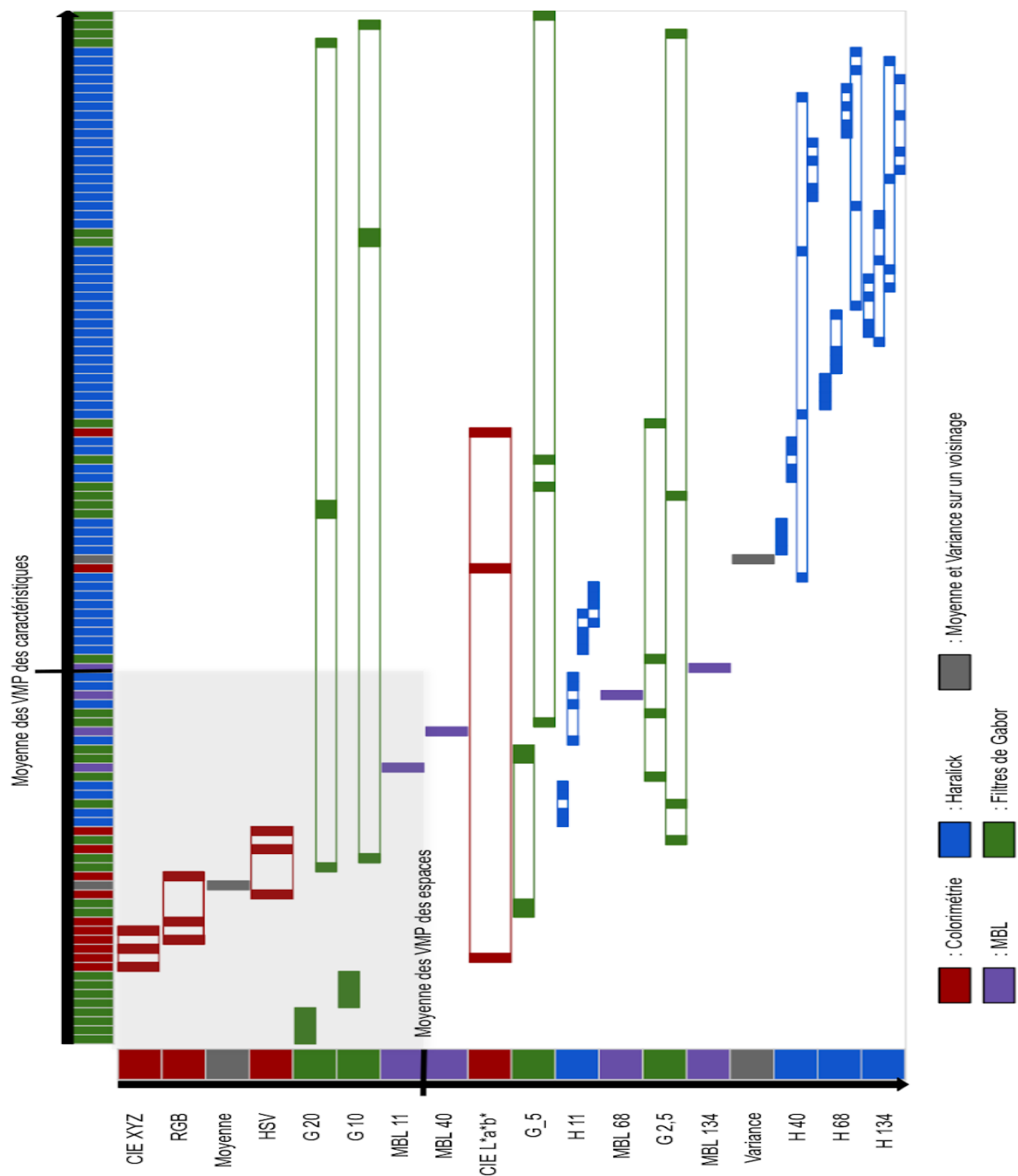


Figure 34 : Répartition des familles de caractéristiques en fonction des meilleures caractéristiques et des meilleurs espaces selon la valeur médiane de positions en validation directe pour le marqueur NeuN. Les caractéristiques sont triées de la plus faible VMeP (à gauche) à la plus forte VMeP (à droite). Les espaces sont triés de la plus faible moyenne des VMeP sur l'ensemble de leurs caractéristiques (en haut) à la plus forte (en bas). Les familles définies Section 9.2.1 sont mises en relation avec leurs espaces d'origine et la répartition des caractéristiques des familles est indiquée par les aires remplies. L'étendue des familles représente l'étendue des familles sur l'ensemble des caractéristiques et leur couleur, l'espace d'origine. La zone grisée inclut les familles ayant à la fois des caractéristiques avec de faibles VMeP et issues d'espaces avec une faible VMeP.

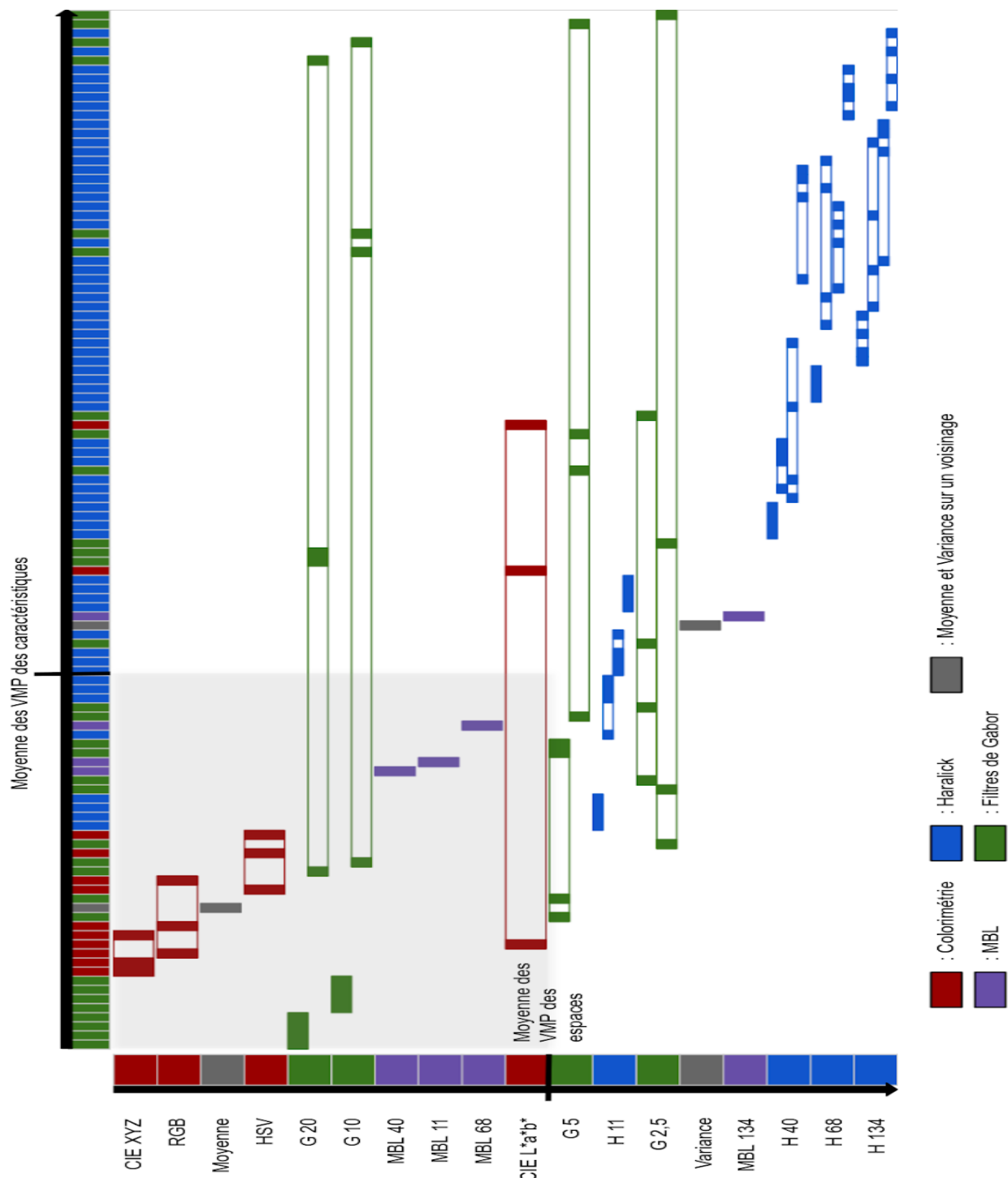


Figure 35 : Répartition des familles de caractéristiques en fonction des meilleures caractéristiques et des meilleurs espaces selon la valeur médiane de positions en validation croisée pour le marqueur NeuN. Les caractéristiques sont triées de la plus faible VMeP (à gauche) à la plus forte VMP (à droite). Les espaces sont triés de la plus faible moyenne des VMeP sur l'ensemble de leurs caractéristiques (en haut) à la plus forte (en bas). Les familles définies Section 9.2.1 sont mises en relation avec leurs espaces d'origine et la répartition des caractéristiques des familles est indiquée par les aires remplies. L'étendue des familles représente l'étendue des familles sur l'ensemble des caractéristiques et leur couleur, l'espace d'origine. La zone grisée inclut les familles ayant à la fois des caractéristiques avec de faibles VMeP et issues d'espaces avec une faible VMeP.

12.2.2. Vecteur de caractéristiques stables

La deuxième étape consiste à tester les combinaisons des 21 caractéristiques sélectionnées avec une limite de taille de vecteur allant jusqu'à 4. À l'instar du vecteur de 114 caractéristiques, nous avons tracé les temps de traitement en fonction de la limite de taille afin de motiver notre choix (Figure 36).

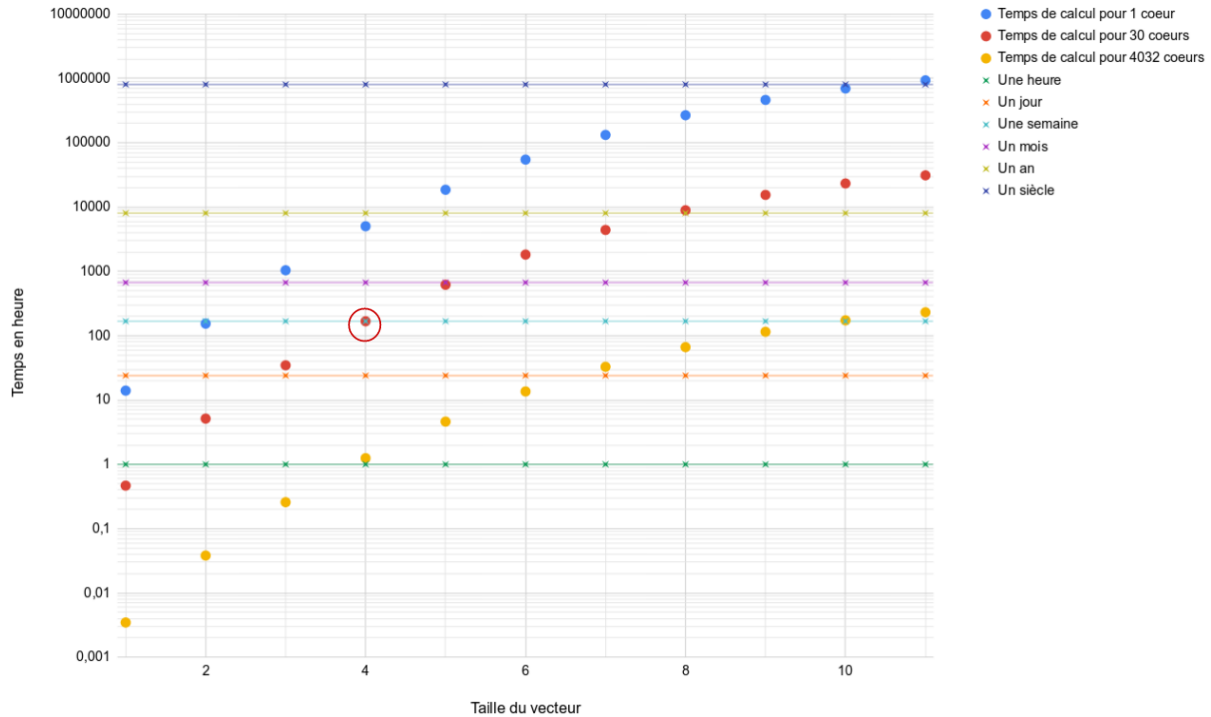


Figure 36 : Temps de traitement (échelle logarithmique) pour différentes conditions après la première étape de sélection (familles) pour le marquage NeuN. Les points ronds colorés représentent les différents temps de traitement pour les conditions exploitées. Les lignes représentent les différentes échelles de temps (de l'heure au siècle). Le cercle rouge représente la condition de traitement avec la limite de taille de vecteur de 4 choisie (30 cœurs et une semaine de calcul).

À partir des résultats obtenus, la répartition des F-Scores en fonction des tailles de vecteurs a été représentée dans la Figure 37. La valeur obtenue pour le vecteur de 114 caractéristiques est donnée à titre de référence. La moyenne augmente et la variance diminue avec la taille croissante des vecteurs de caractéristiques étudiés. L'écart entre le F-Score minimal et maximal diminue lui aussi avec le nombre croissant de caractéristiques. À partir d'une taille de deux caractéristiques, aucun F-Score n'est en dessous de 0,6 et la majorité est au-dessus de 0,8.

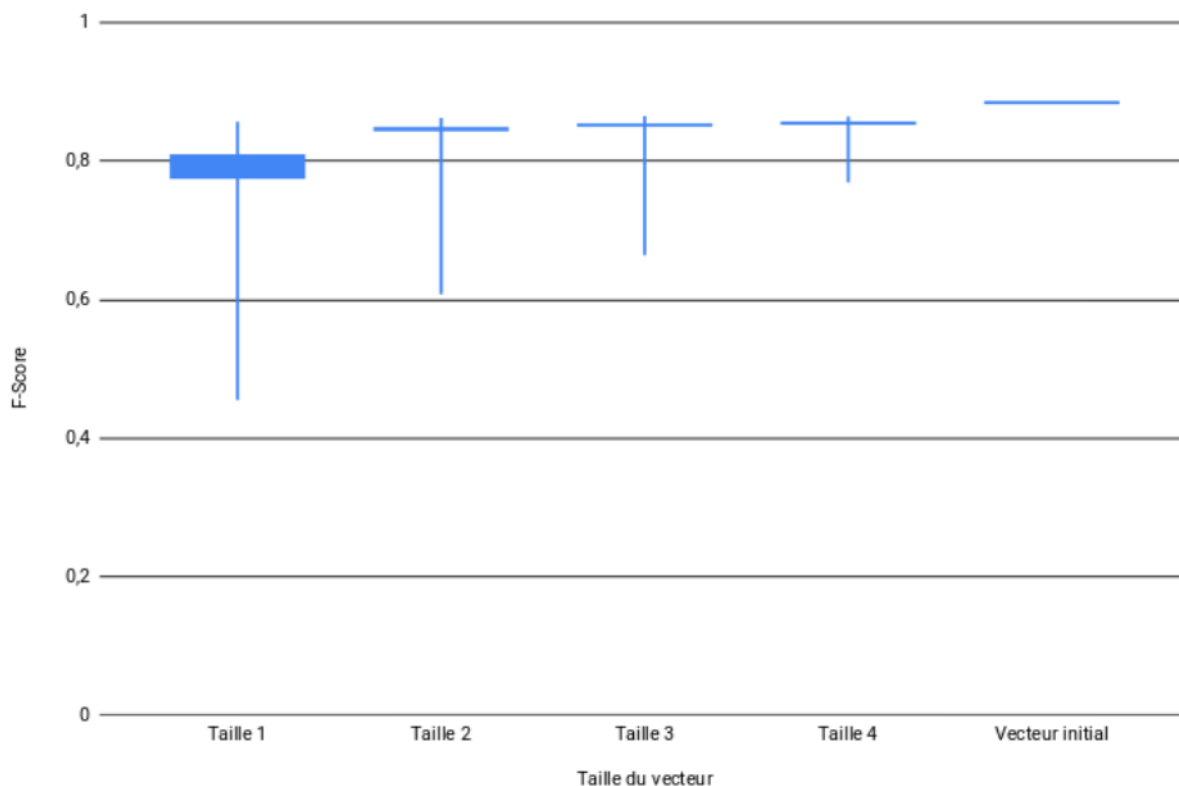


Figure 37 : Répartition des F-Scores des différentes combinaisons en fonction de leur taille pour le marquage NeuN en validation directe. Les extrémités des tiges verticales représentent le minimum (en bas) et le maximum (en haut) du F-Score. La limite basse de la boîte en bleu plein représente la moyenne des F-Scores moins sa variance et la limite haute la moyenne plus sa variance.

Les tableaux de 7 à 9 présentent les résultats de sélection obtenus avec les trois critères détaillés Section 10.2.2 (VMeP, VMoP et moyenne des facteurs d'importance). Pour la valeur médiane de positions (Tableau 7), 2 caractéristiques ont été sélectionnées dans les deux conditions de validation :

- MBL avec un rayon de 68 pixels
- une caractéristique issue de la partie réelle du filtre de Gabor avec une longueur d'onde de 20 pixels

Les caractéristiques sélectionnées après la condition d'arrêt sont différentes entre les deux conditions de validation. Les vecteurs sélectionnés ont une moyenne de F-Score de 0,862 ce qui représente environ 97% du F-Score obtenu avec le vecteur initial de 114 caractéristiques et constitue un écart faible entre les deux F-Scores.

Numéro de l'itération	Condition de validation	Vecteur de caractéristiques sélectionné par VMeP	F-Score
1	Directe	[MBL 68]	0,4552
	Croisée	[MBL 68]	0,5000
2	Directe	[MBL 68 ; G_20_R_135]	0,8584
	Croisée	[MBL 68 ; G_20_R_0]	0,8657
3	Directe	[MBL 68 ; G_20_R_135 ; S]	0,8476
	Croisée	[MBL 68 ; G_20_R_0 ; MBL 40]	0,8643
4	Directe	[MBL 68 ; G_20_R_135 ; S ; G_10_R_135]	0,8592
	Croisée	[MBL 68 ; G_20_R_0 ; MBL 40 ; MBL 11]	0,8672

Tableau 7 : Tableau de sélection des caractéristiques avec comme critère la VMeP pour le marquage NeuN. À chaque itération, le vecteur de caractéristiques produit par agrégation est indiqué avec le F-Score correspondant pour chaque condition de validation (VD et VC). Les vecteurs en gras sont les vecteurs sélectionnés.

Pour la valeur moyenne de positions (Tableau 8), le vecteur sélectionné à l'itération numéro 2 est quasiment le même que pour la valeur médiane de position. Dans ce cas, la condition d'arrêt est atteinte à la troisième itération. Le vecteur de caractéristiques sélectionné est :

- MBL avec un rayon de 68 pixels
- une caractéristique de la partie réelle du filtre de Gabor avec une longueur d'onde de 20 pixels
- MBL avec un rayon de 40 pixels

Les vecteurs sélectionnés en VD et VC ont une moyenne de F-Score de 0,866 ce qui représente environ 97% du F-Score obtenu avec le vecteur initial de 114 caractéristiques. Cela est équivalent avec les résultats obtenus avec le critère VMeP.

Numéro de l'itération	Condition de validation	Vecteur de caractéristiques sélectionné par VMoP	F-Score
1	Directe	[MBL 68]	0,4552
	Croisée	[MBL 68]	0,5000
2	Directe	[MBL 68 ; G_20_R_45]	0,8594
	Croisée	[MBL 68 ; G_20_R_90]	0,8659
3	Directe	[MBL 68 ; G_20_R_45 ; MBL 40]	0,8645
	Croisée	[MBL 68 ; G_20_R_90 ; MBL 40]	0,8689
4	Directe	[MBL 68 ; G_20_R_45 ; MBL 40 ; MBL 11]	0,8643
	Croisée	[MBL 68 ; G_20_R_90 ; MBL 40 ; MBL 11]	0,8676

Tableau 8 : Tableau de sélection des caractéristiques avec comme critère la VMoP pour le marquage NeuN. À chaque itération, le vecteur de caractéristiques produit par agrégation est indiqué avec le F-Score correspondant pour chaque condition de validation (VD et VC). Les vecteurs en gras sont les vecteurs sélectionnés.

Enfin la moyenne des facteurs d'importance produit un vecteur avec une seule caractéristique. Cette dernière est issue de la partie réelle du filtre de Gabor avec une longueur d'onde de 20 pixels. Dans les étapes de sélection postérieure à la condition d'arrêt, seules des caractéristiques de Gabor et l'image moyenne ont été sélectionnées. Avec un F-Score moyen de 0,8599 la moyenne des facteurs d'importance fournit le vecteur de caractéristiques le plus petit mais également celui fournissant le plus faible F-Score des trois critères testés.

Numéro de l'itération	Condition de validation	Vecteur de caractéristiques sélectionné par la moyenne des facteurs d'importances	F-Score
1	Directe	[G_20_R_90]	0,8569
	Croisée	[G_20_R_90]	0,8629
2	Directe	[G_20_R_90 ; G_20_R_45]	0,8567
	Croisée	[G_20_R_90 ; G_20_R_45]	0,8627
3	Directe	[G_20_R_90 ; G_20_R_45 ; G_20_R_0]	0,8565
	Croisée	[G_20_R_90 ; G_20_R_45 ; Moyenne]	0,8615
4	Directe	[G_20_R_90 ; G_20_R_45 ; G_20_R_0 ; G_10_R_135]	0,8567
	Croisée	[G_20_R_90 ; G_20_R_45 ; Moyenne ; G_20_R_0]	0,8627

Tableau 9 : Tableau de sélection des caractéristiques avec comme critère la moyenne des facteurs d'importance pour le marquage NeuN. À chaque itération, le vecteur de caractéristiques produit par agrégation est indiqué avec le F-Score correspondant pour chaque condition de validation. Les vecteurs en gras sont les vecteurs sélectionnés.

Le tableau 10 synthétise et compare les résultats des différentes méthodes de sélection en fonction des différentes conditions de validation (VD et VC) ainsi que les valeurs moyennes de F-Score des vecteurs résultants. Les vecteurs sélectionnés sont similaires entre les conditions de validation sauf pour le vecteur sélectionné par Force Brute simple. On peut observer que la sélection de caractéristiques incluse dans l'apprentissage de la forêt aléatoire tend à sélectionner des caractéristiques issues des familles de Gabor. Les F-Scores résultats entre la méthode incluse dans l'apprentissage des forêts aléatoires et la sélection avec la moyenne des facteurs d'importance sont identiques et inférieurs aux autres F-Scores. Le F-Score du vecteur sélectionné par VMeP est très proche du F-Score de celui sélectionné par VMoP. Cependant le vecteur sélectionné par VMeP a une cardinalité plus petite.

Dans ce contexte, le vecteur optimal que nous avons retenu est celui de la VMoP :

- MBL avec un rayon de 68 pixels
- une caractéristique issue de la partie réelle du filtre de Gabor avec un rayon de 20 pixels

Méthode de sélection	Validation directe	Validation croisée	Moyenne des F-Score
Force Brute simple	[G_20_R_135 ; MBL 40 ; MBL 68]	[G_20_R_45 ; MBL 11 ; MBL 40 ; MBL 68]	0,8679 ± 0,2%
Facteur d'importance	[G_20_R_90 ; G_20_R_45 ; G_20_R_135 ; Moyenne]	[G_20_R_90 ; G_20_R_45 ; G_20_R_135 ; Moyenne]	0,8598 ± 0,3%
Valeur médiane de positions	[MBL 68 ; G_20_R_135]	[MBL 68 ; G_20_R_0]	0,8620 ± 0,3%
Valeur moyenne de positions	[MBL 68 ; G_20_R_45 ; MBL 40]	[MBL 68 ; G_20_R_90 ; MBL 40]	0,8667 ± 2,2%
Moyenne des facteurs d'importance	[G_20_R_90]	[G_20_R_90]	0,8599 ± 0,3%

Tableau 10 : Tableau de comparaison pour les différentes méthodes de sélection pour le marquage NeuN (Section 10.1).

Des exemples d'images de différentes caractéristiques pour le marquage NeuN sont présentés dans la Figure 38. La partie imaginaire du filtre de Gabor avec une longueur d'onde de 5 pixels à 135° est la pire caractéristique selon la valeur médiane de positions. À partir des images générées, nous pouvons très difficilement discerner les neurones, contrairement aux autres images. La pire caractéristique après la sélection des familles est le Rouge de l'espace RVB. Néanmoins, l'image générée par cette caractéristique reste visuellement intéressante, les neurones y sont clairement identifiables. Ceci s'explique par le fait que toutes les caractéristiques sélectionnées après la première étape de notre protocole donnent de très bons résultats avec les métriques que nous avons proposées.

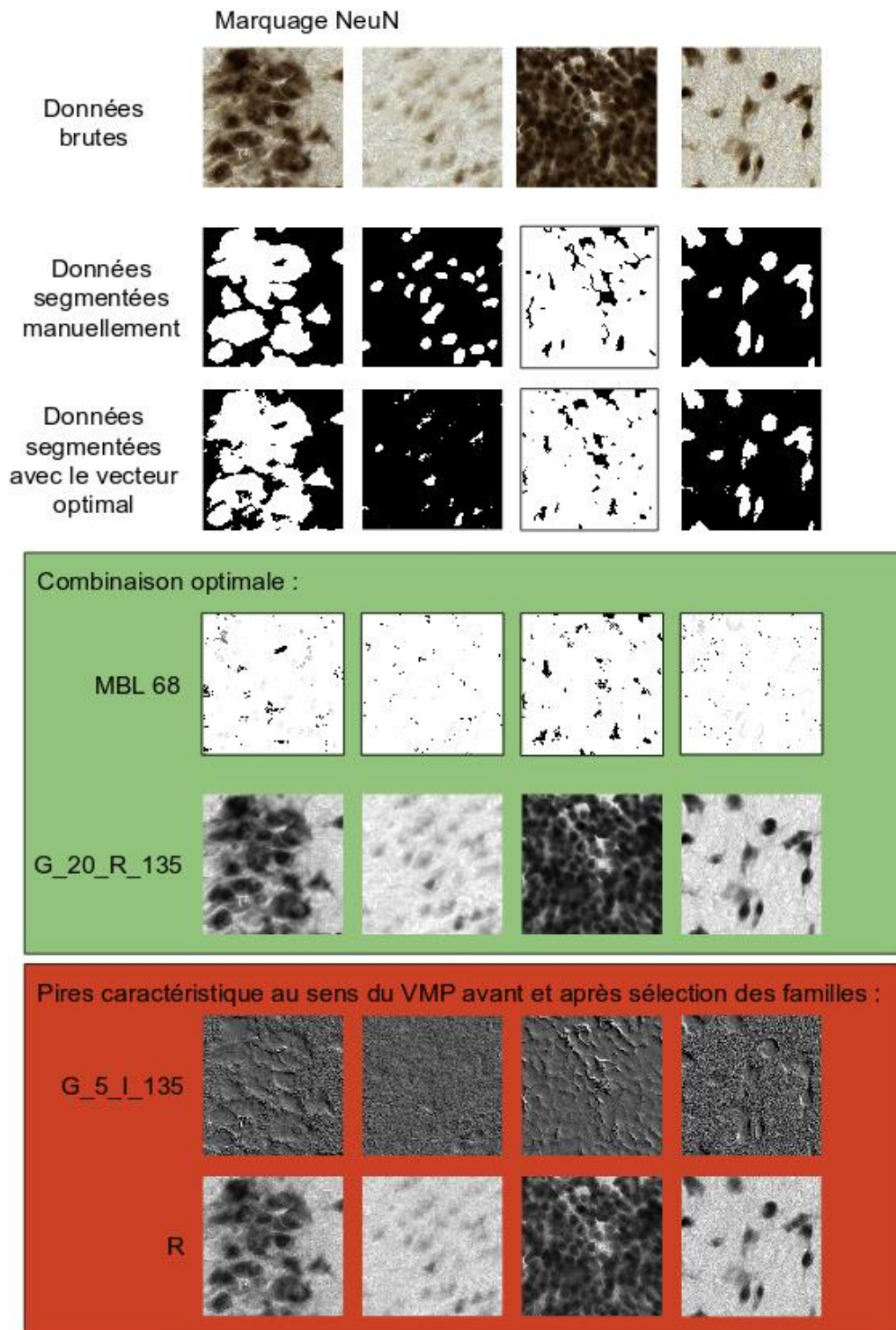


Figure 38 : Les images brutes, les segmentations manuelles et la segmentation automatique en validation directe sont sur les trois premières lignes pour plusieurs exemples d'images NeuN. Images des caractéristiques optimales et des pires caractéristiques avant et après la première étape de sélection selon la valeur médiane de position pour le marquage NeuN. Dans le cadre vert, les images des caractéristiques optimales du marquage NeuN sont présentées. La pire des 114 caractéristiques initiales est sur la première ligne dans le cadre rouge. La pire des caractéristiques après la première étape de sélection (familles) est sur la deuxième ligne dans le cadre rouge.

12.3. Gain computationnel

La réduction du nombre de caractéristiques permet une diminution : 1) de la quantité de mémoire vive utilisée lors de l'étape de segmentation et 2) du temps de segmentation. Il a été ainsi possible de limiter la quantité des données intermédiaires. L'ordinateur utilisé pour réaliser ces calculs est décrit dans Section 8.1.

La segmentation d'une coupe de cerveau de macaque (section 91, voir Section 8) a été réalisée en exploitant le vecteur optimal de caractéristiques. Le Tableau 11 présente les temps de calcul qui ont été nécessaires avec 1 et 30 cœurs de calcul et en utilisant le TGCC avec 4032 cœurs de calcul pour les différentes étapes de traitement :

- phase de sélection du vecteur optimal
- extraction des caractéristiques pour l'image de la coupe entière du cerveau de macaque
- segmentation par forêt aléatoire avec le vecteur initial et optimal de la coupe de cerveau de macaque

Par rapport au vecteur initial de 114 caractéristiques, le vecteur optimal de 2 caractéristiques introduit une réduction d'un facteur 57 et permet de diviser le temps de traitement total par 2 044 en traitement séquentiel (rapport des sommes des temps d'extraction et de segmentation entre le traitement avec et sans sélection par Force Brute). Le procédé de sélection constitue plus des deux tiers du temps de traitement total pour le traitement d'une coupe entière de cerveau de macaque. Le coût de notre optimisation en termes de temps de calcul peut être rentabilisé dès le traitement de 2 coupes entières de cerveau de macaque. De plus, la quantité de données produites et stockées sur disque lors de la mise en production de la segmentation est considérablement réduite même si ces données n'ont pas vocation à être conservées durablement.

	Nombre de cœurs de calcul utilisés	Temps de la sélection par notre méthode Force Brute	Nombre de caractéristiques considérées	Temps d'extraction des caractéristiques pour la coupe entière NeuN	Temps de segmentation par forêt aléatoire de la coupe entière NeuN
Station de travail	1	NA	114	~2 ans 6 mois*	~7 mois*
	1	~300 jours*	2	~125 jours*	~30h*
	30	NA	114	~1 mois 26 jours*	~14 jours*
	30	~10 jours	2	~4 jours 4h	~1h
Irène Skylake (TGCC)	4032	~2h	2	~45min*	~1min*

Tableau 11 : Tableau des temps de traitement pour une image entière de cerveau de macaque marquée avec NeuN et numérisée à haute résolution (~150 Go). Les deux dernières colonnes correspondent aux étapes de segmentation de la coupe (extraction des caractéristiques optimales et classification par forêt aléatoire). Les temps avec une étoile sont des estimations.

13. Sélection de caractéristiques pour un marqueur non-spécifique : anti-Phox2B

13.1. Rappel de l'objectif de la segmentation

L'objectif est le même que pour le marquage NeuN à savoir de segmenter la classe "neurone" correspondant aux neurones marqués par anti-Phox2B avec un minimum de caractéristiques et pour une qualité de segmentation comparable à la segmentation initiale. Sur les 114 caractéristiques utilisées pour segmenter les "neurones" marqués avec l'anticorps anti-Phox2B, lesquelles sont les plus pertinentes pour la segmentation ? Si ces caractéristiques existent, forment-elles un sous-ensemble stable ?

13.2. Sélection basée sur la Force Brute

13.2.1. Familles sélectionnées pour anti-Phox2B

La première étape de sélection est appliquée aux familles de caractéristiques. Nous appliquons la même méthodologie que celle présentée dans la Section 12.2.1 (Figure 39). Pour des valeurs inférieures à la moyenne des VMeP, la dispersion est quasiment nulle avec une corrélation globale de 0,95 en validation directe et de 0,90 en validation croisée à l'échelle des caractéristiques, et 0,96 en validation directe et 0,94 en validation croisée à l'échelle des familles. Comme pour NeuN, la médiane sera préférée car sa dynamique est plus large que celle de la moyenne. La dynamique de la valeur médiane de position est de 5 141,5 et celle de la valeur moyenne de position est de 4 215,23.

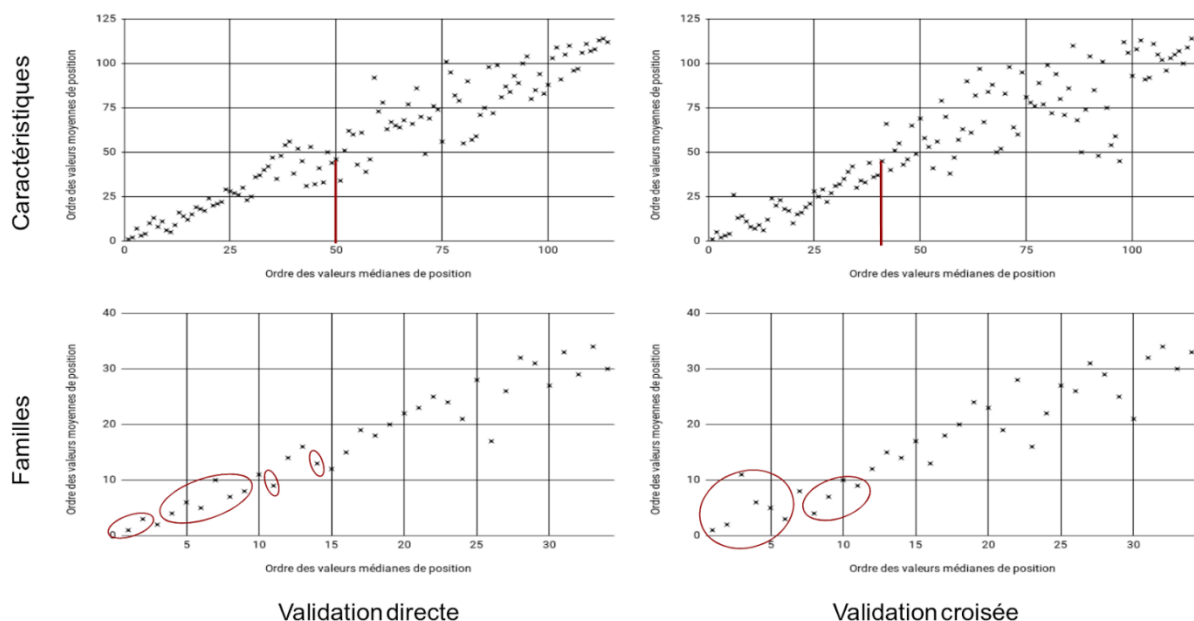


Figure 39 : Corrélation de la distribution des valeurs médianes et moyennes de positions pour le marqueur anti-Phox2B. À l'échelle des caractéristiques, la barre rouge représente la moyenne de toutes les valeurs médianes de positions. À l'échelle des familles, les familles entourées en rouge sont les familles sélectionnées lors de cette première étape (voir Figures 40 et 41).

À l'instar du marqueur NeuN (Section 12.2.1), les mêmes graphiques de mise en correspondance des familles avec les espaces et les caractéristiques sont produits pour chaque condition de validation pour le marqueur anti-Phox2B (Figures 40 et 41). L'annexe 2 contient les tableaux numériques des classements des familles de caractéristiques par ordre de moyenne des VMP de ses caractéristiques. On peut observer une plus forte dispersion et une corrélation plus faible pour le marquage anti-Phox2B par rapport au marquage NeuN.

En prenant en compte les 114 caractéristiques et en étudiant les combinaisons jusqu'à une taille de 2, les familles sélectionnées en validation directe étaient dans l'ordre de leur VMeP moyenne (avec entre parenthèse la nomenclature présentée Annexe 2) :

- le MBL avec un rayon de 40 pixels (MBL 40)
- le MBL avec un rayon de 20 pixels (MBL 20)
- la partie réelle du filtre de Gabor avec un lambda de 20 pixels (G_20_I)
- la partie réelle du filtre de Gabor avec un lambda de 10 pixels (G_10_I)
- HSV
- le MBL avec un rayon de 10 pixels (MBL 10)
- RVB
- CIE XYZ
- l'image moyenne (Moyenne)

En validation croisée, les familles sélectionnées étaient :

- le MBL avec un rayon de 40 pixels (MBL 40)
- le MBL avec un rayon de 20 pixels (MBL 20)
- le MBL avec un rayon de 10 pixels (MBL 10)
- la partie réelle du filtre de Gabor avec un lambda de 20 pixels (G_20_I)
- RVB
- HSV
- CIE XYZ
- CIE L*a*b*

Aucune famille d'Haralick n'est sélectionnée dans les différentes conditions de validation. Néanmoins, les composantes de deux familles d'Haralick (la corrélation et la variance d'Haralick pour un rayon de 5 pixels) ont des caractéristiques considérées comme avantageuses (dont la VMeP était inférieure à la moyenne de toutes les VMeP).

À part CIE L*a*b*, le reste des familles sélectionnées en validation croisée est retrouvé en validation directe avec en plus un filtre de Gabor et l'image moyenne. Donc sur l'union des familles sélectionnées, soit 10 familles, 7 familles étaient communes aux deux conditions de validation (70% des familles). Sur les 10 familles, nous avons recensé 4 familles colorimétriques et 6 familles texturales, ce qui correspond à 12 caractéristiques colorimétriques et 12 caractéristiques texturales. Les 24 caractéristiques sélectionnées pour l'étape suivante représentaient 21,05% de l'ensemble des caractéristiques de départ ce qui est un ordre de grandeur proche de celui trouvé pour NeuN. Cette étape réduit la complexité algorithmique totale à 2^{24} soit 16 777 216 combinaisons (soit une division du nombre total de combinaisons à traiter d'un facteur de 1024). Afin d'avoir des temps de traitement acceptables (de l'ordre de la semaine), la taille maximale des vecteurs explorés choisie est de 4, soit 12 950 combinaisons (42 504 combinaisons seraient ajoutées avec une taille maximale de 5).

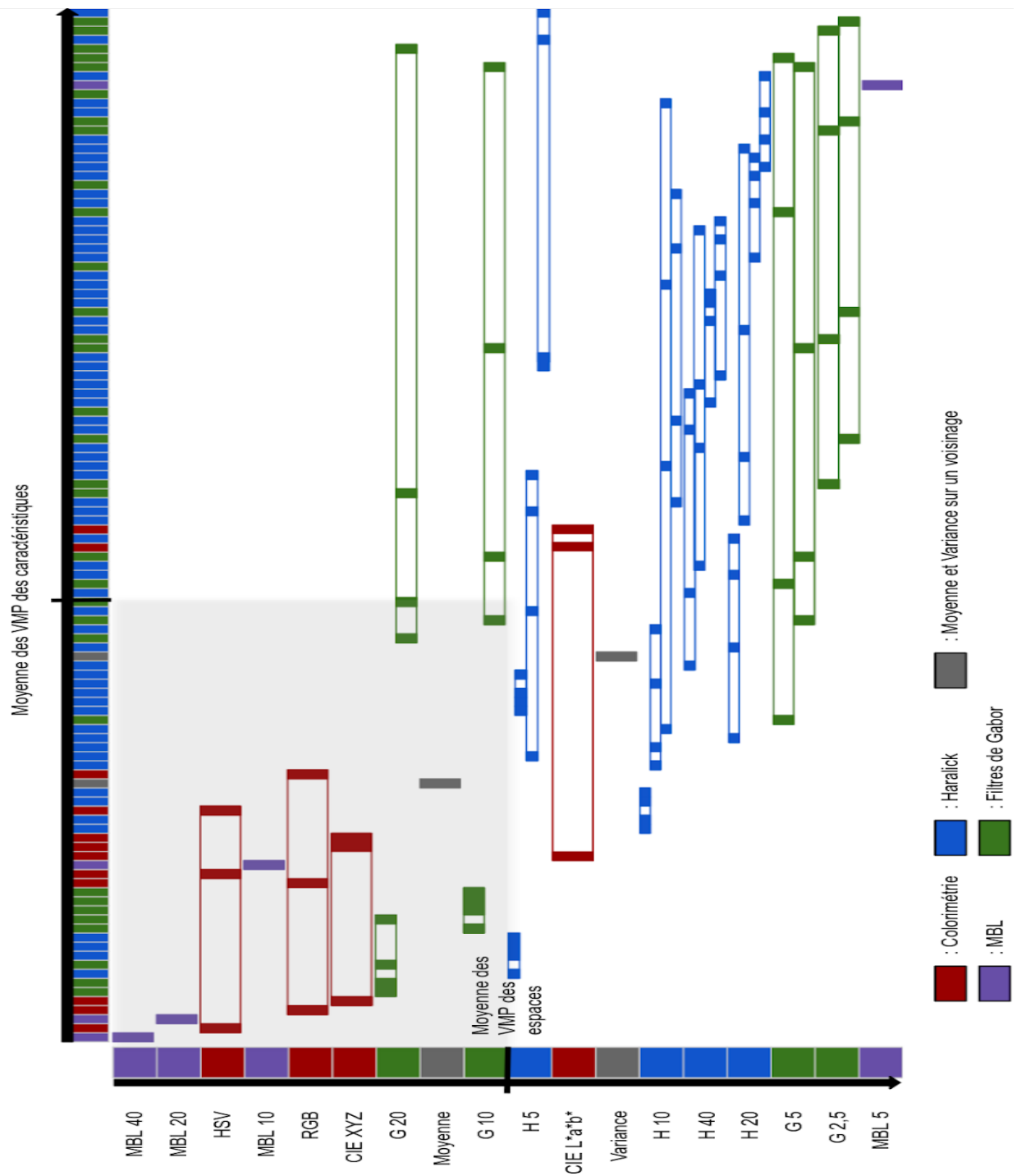


Figure 40 : Répartition des familles de caractéristiques en fonction des meilleures caractéristiques et des meilleurs espaces selon la valeur médiane de positions en validation directe pour le marqueur anti-Phox2B. Les caractéristiques sont triées de la plus faible VMP (à gauche) à la plus forte VMP (à droite). Les espaces sont triés de la plus faible moyenne des VMP sur l'ensemble de leurs caractéristiques (en haut) à la plus forte (en bas). Les familles définies Section 9.2.1 sont mises en relation avec leurs espaces d'origine et la répartition des caractéristiques des familles est indiquée par les aires remplies. L'étendue des familles représente l'étendue des familles sur l'ensemble des caractéristiques et leur couleur, l'espace d'origine. La zone grisée inclut les familles ayant à la fois des caractéristiques avec de faibles VMP et issues d'espaces avec une faible VMP soient les meilleures familles candidates pour notre seconde étape de sélection.

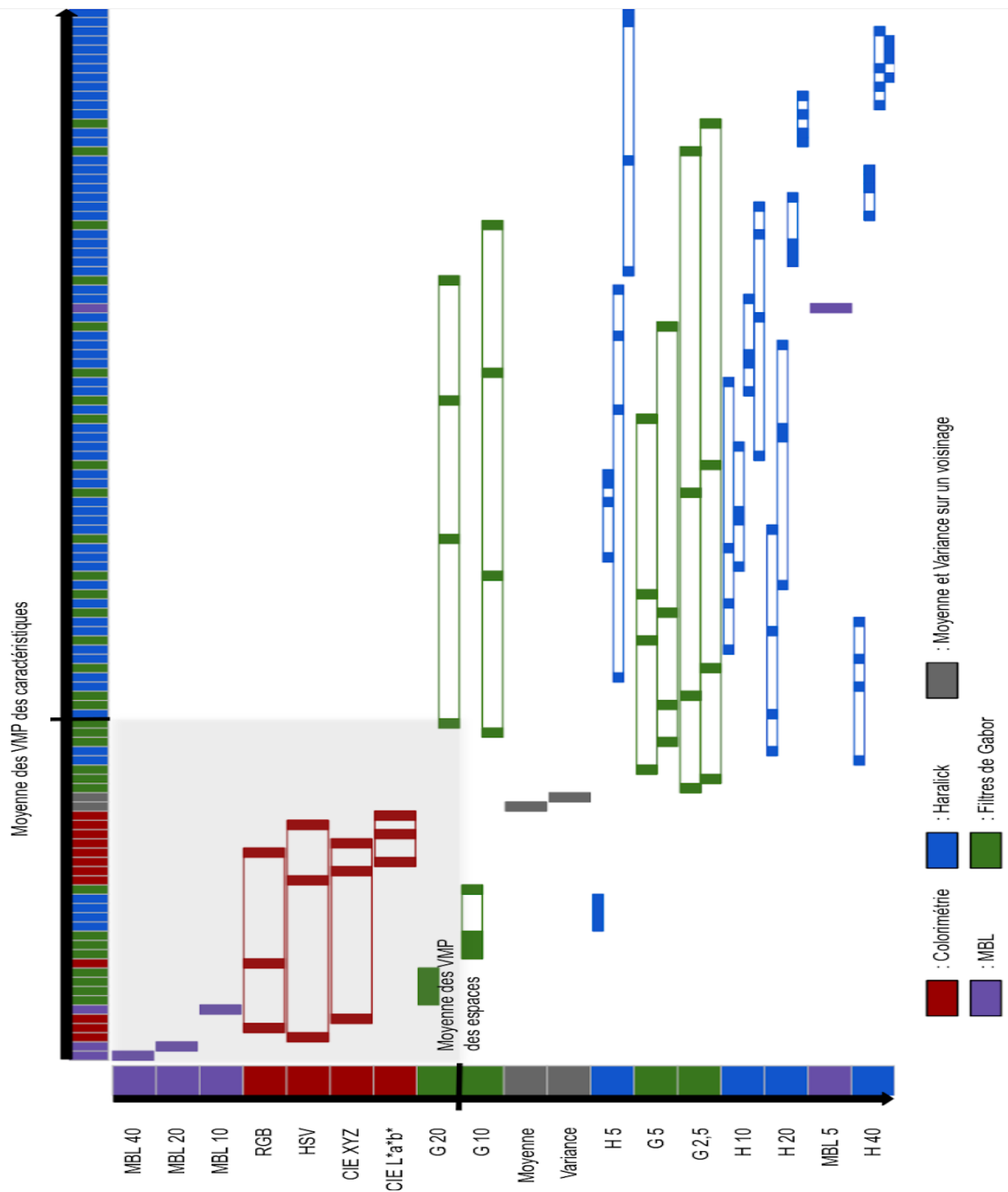


Figure 41 : Répartition des familles de caractéristiques en fonction des meilleures caractéristiques et des meilleurs espaces selon la valeur médiane de positions en validation croisée pour le marqueur anti-Phox2B. Les caractéristiques sont triées de la plus faible VMP (à gauche) à la plus forte VMP (à droite). Les espaces sont triés de la plus faible moyenne des VMP sur l'ensemble de leurs caractéristiques (en haut) à la plus forte (en bas). Les familles définies Section 9.2.1 sont mises en relation avec leurs espaces d'origine et la répartition des caractéristiques des familles est indiquée par les aires remplies. L'étendue des familles représente l'étendue des familles sur l'ensemble des caractéristiques et leur couleur, l'espace d'origine. La zone grisée inclut les familles ayant à la fois des caractéristiques avec de faibles VMP et issues d'espaces avec une faible VMP soient les meilleures familles candidates pour notre seconde étape de sélection.

13.2.2. Vecteur de caractéristiques stables

La deuxième étape de sélection consiste à tester les combinaisons des 24 caractéristiques sélectionnées avec des tailles de vecteurs allant jusqu'à 4. La même figure que pour la Section 12.2.2 est produite pour justifier notre choix (Figure 42). Nous pouvons observer que la limite de taille de 4 a un temps de traitement supérieur à la semaine avec 30 cœurs de calcul en parallèle. Les résultats présentés sont produits avec l'utilisation du TGCC.

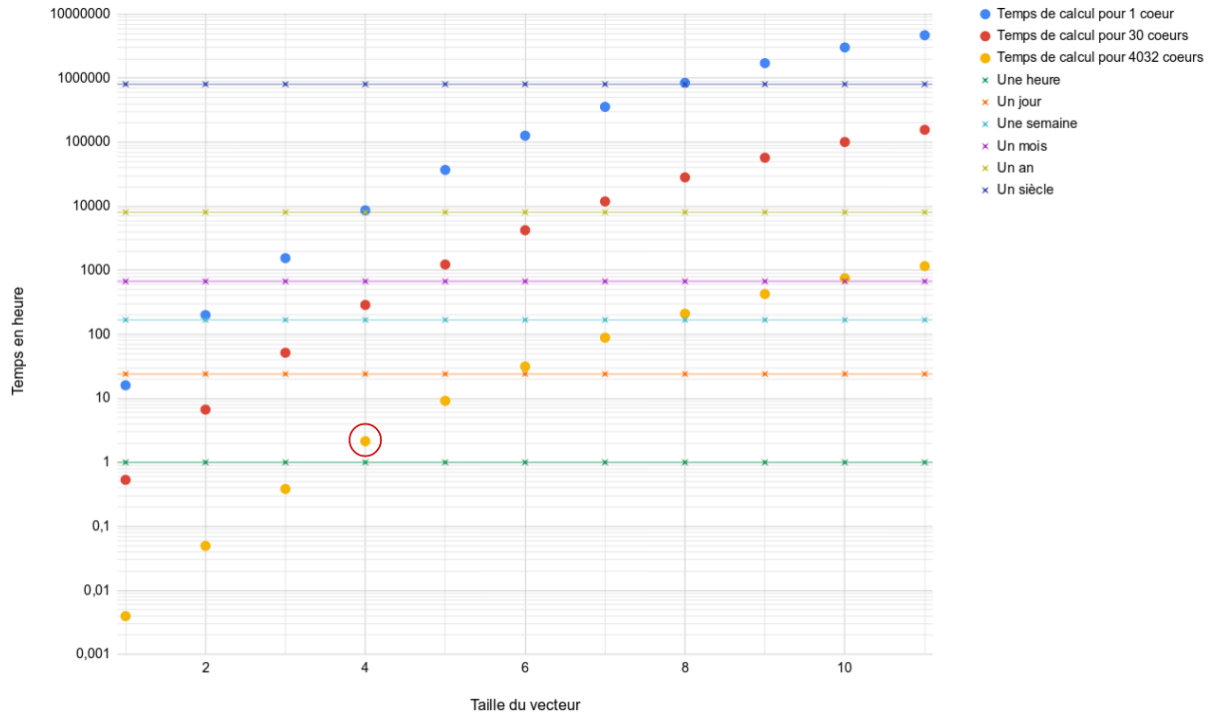


Figure 42 : Temps de traitement (échelle logarithmique) pour différentes conditions pour le marquage anti-Phox2B. Les points ronds colorés représentent les différents temps de traitement pour les conditions exploitées. Les lignes représentent les différentes échelles de temps (de l'heure au siècle). Le cercle rouge représente la condition de traitement avec la limite de taille de vecteur.

À partir des résultats obtenus, la répartition des F-Scores en fonction des tailles de vecteurs a été représentée dans la Figure 43. Le F-Score du vecteur initial (114 caractéristiques) est donné à titre indicatif. La moyenne et la variance augmentent avec la taille croissante des vecteurs de caractéristiques étudiés. L'écart entre le F-Score minimal et maximal augmente lui aussi avec le nombre de caractéristiques. Nous pouvons observer que l'évolution des variances des F-Scores avec la taille des combinaisons analysées est différent de celui observé avec le marquage NeuN (qui décroît quant à lui comme présenté Section 12.2.2). Même avec une Taille de vecteur de 4, le F-Score moyen est de 0,2306, ce qui représente 35% du F-Score obtenu avec le vecteur des 114 caractéristiques initiales.

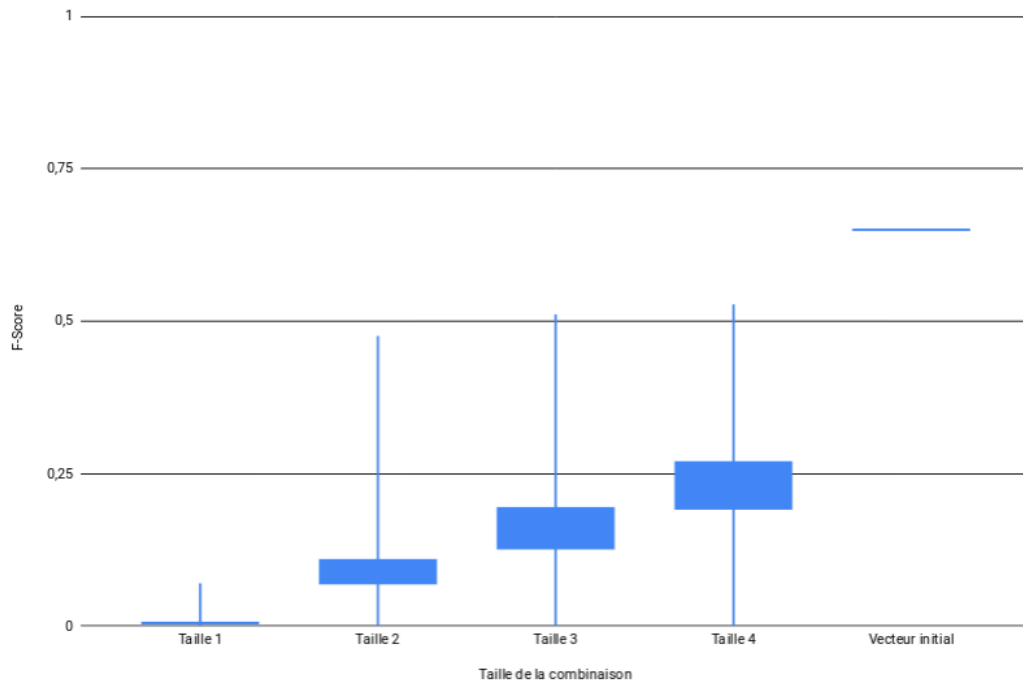


Figure 43 : Répartition des F-Scores des différentes combinaisons testées pendant la seconde étape de sélection en fonction de leur taille pour le marquage anti-Phox2B en validation directe. Les extrémités des tiges verticales représentent le minimum (en bas) et le maximum (en haut) du F-Score. La limite basse de la boîte en bleu plein représente la moyenne des F-Scores moins sa variance et la limite haute la moyenne plus sa variance.

Les tableaux de 12 à 14 montrent les résultats obtenus avec les trois critères détaillés Section 10.2.2. Aucune des sélections produites avec les différents critères ne rencontrent la condition d'arrêt. En effet, les F-Scores des combinaisons sélectionnées avec les valeurs de positions augmentent, ce qui est visible dans les tableaux 12 et 13. Pour ces sélections, seule la dernière caractéristique diffère entre la validation directe et la validation croisée. Néanmoins, cette dernière caractéristique n'est pas sélectionnée par les valeurs de position mais par le F-Score car l'antépénultième combinaison est sélectionnée, ce qui est expliqué Section 10.1.

Pour la valeur médiane de positions, la combinaison sélectionnée en validation directe est (la nomenclature présentée Annexe 2 est entre parenthèse) :

- le MBL avec un rayon de 20 pixels (MBL 20)
- b^* de l'espace colorimétrique CIE $L^*a^*b^*$ (b^*)
- la partie réelle du filtre de Gabor avec une longueur d'onde de 20 pixels et un angle de 90° (G_20_R_90)

Le F-Score de cette combinaison est de 0,5083 (0,65 pour 114 caractéristiques), ce qui représente 78% du F-Score du vecteur initial de 114 caractéristiques (Tableau 12).

En validation croisée, le vecteur est :

- le MBL avec un rayon de 20 pixels (MBL 20)
- b^* de l'espace colorimétrique CIE $L^*a^*b^*$ (b^*)
- la partie réelle du filtre de Gabor avec une longueur d'onde de 20 pixels et un angle de 0° (G_20_R_0)

Le F-Score de cette combinaison est de 0,3984 (0,58 pour 114 caractéristiques), ce qui représente 68% du F-Score du vecteur initial de 114 caractéristiques (Tableau 12).

Numéro de l'itération	Condition de validation	Vecteur de caractéristiques sélectionné par VMeP	F-Score
1	Directe	[MBL 20]	0
	Croisée	[MBL 20]	0
2	Directe	[MBL 20 ; b*]	0,0518
	Croisée	[MBL 20 ; b*]	0,0001
3	Directe	[MBL 20 ; b* ; G_20_R_90]	0,5083
	Croisée	[MBL 20 ; b* ; G_20_R_0]	0,3984
4	Directe	[MBL 20 ; b* ; G_20_R_90 ; MBL 10]	0,5277
	Croisée	[MBL 20 ; b* ; G_10_R_0 ; MBL 40]	0,4229

Tableau 12 : Tableau de sélection avec comme critère la Valeur médiane de positions pour le marquage anti-Phox2B. À chaque itération, le vecteur de caractéristiques agrégé est indiqué avec le F-Score correspondant pour chaque condition de validation. Les vecteurs en gras sont les vecteurs sélectionnés.

Pour la valeur moyenne de positions, la combinaison sélectionnée en validation directe est (la nomenclature présentée Annexe 2 est entre parenthèse) :

- le MBL avec un rayon de 20 pixels (MBL 20)
- la valeur de saturation (S) de l'espace colorimétrique HSV
- le MBL avec un rayon de 10 pixels (MBL 10)

Le F-Score de cette combinaison est de 0,4969 (0,65 pour 114 caractéristiques), ce qui représente 76% du F-Score du vecteur initial de 114 caractéristiques (Tableau 13).

En validation croisée, le vecteur est :

- le MBL avec un rayon de 20 pixels (MBL 20)
- la valeur de saturation (S) de l'espace colorimétrique HSV
- le MBL avec un rayon de 10 pixels (MBL 10)

Le F-Score de cette combinaison est de 0,3868 (0,58 pour 114 caractéristiques), ce qui représente 66% du F-Score du vecteur initial de 114 caractéristiques (Tableau 13).

Numéro de l'itération	Condition de validation	Vecteur de caractéristiques sélectionné par VMoP	F-Score
1	Directe	[MBL 20]	0
	Croisée	[MBL 20]	0
2	Directe	[MBL 20 ; S]	0,4759
	Croisée	[MBL 20 ; S]	0,3602
3	Directe	[MBL 20 ; S ; MBL 10]	0,4969
	Croisée	[MBL 20 ; S ; MBL 10]	0,3868
4	Directe	[MBL 20 ; S ; MBL 10 ; G_20_R_90]	0,5233
	Croisée	[MBL 20 ; S ; MBL 10 ; Moyenne]	0,4195

Tableau 13 : Tableau de sélection avec comme critère la Valeur moyenne de positions pour le marquage anti-Phox2B. À chaque itération, le vecteur de caractéristiques agrégé est indiqué avec le F-Score correspondant pour chaque condition de validation. Les vecteurs en gras sont les vecteurs sélectionnés.

Pour la moyenne des facteurs d'importance, les caractéristiques sélectionnées sont les caractéristiques de la partie réelle du filtre de Gabor avec une longueur d'onde de 20 pixels et, pour le dernier tour de validation croisée, l'image moyenne. À la fin des quatre tours de sélection, le F-Score est de 0,15 en moyenne, ce qui représente 24% de la moyenne des F-Score avec 114 caractéristiques.

Numéro de l'itération	Condition de validation	Vecteur de caractéristiques sélectionné par moyenne des facteurs d'importance	F-Score
1	Directe	[G_20_R_0]	0
	Croisée	[G_20_R_0]	0
2	Directe	[G_20_R_0 ; G_20_R_135]	0
	Croisée	[G_20_R_0 ; G_20_R_90]	0
3	Directe	[G_20_R_0 ; G_20_R_135 ; G_20_R_45]	0
	Croisée	[G_20_R_0 ; G_20_R_90 ; G_20_R_90]	0
4	Directe	[G_20_R_0 ; G_20_R_135 ; G_20_R_45 ; G_20_R_90]	0,0013
	Croisée	[G_20_R_0 ; G_20_R_90 ; G_20_R_45 ; Moyenne]	0,2310

Tableau 14 : Tableau de sélection avec comme critère la moyenne des facteurs d'importance du FA pour le marquage anti-Phox2B. À chaque itération, le vecteur de caractéristiques agrégé est indiqué avec le F-Score correspondant pour chaque condition de validation. Les vecteurs en gras sont les vecteurs sélectionnés.

Le tableau 15 compare les différentes méthodes de sélection en fonction des conditions de validation directe et croisée et le F-Score moyen résultant associé. Les combinaisons sélectionnées sont similaires pour les valeurs de positions mais pas pour la sélection par Force Brute simple et les sélections utilisant le facteur d'importance. A l'instar du marquage NeuN, la méthode de sélection incluse dans l'apprentissage de la forêt aléatoire tend à sélectionner des caractéristiques issues des familles de Gabor. Les combinaisons sélectionnées par valeurs de positions ont une cardinalité plus faible que les autres combinaisons. De plus, la combinaison sélectionnée par la valeur médiane de positions possède le F-Score le plus élevé. Cette combinaison est la combinaison considérée comme optimale pour le marquage anti-Phox2B.

Méthode de sélection	Validation directe	Validation croisée	Moyenne des F-Scores
Force Brute simple	[b* ; G_20_R_90 ; MBL 10 ; MBL 20]	[a* ; Moyenne; MBL 20 ; MBL 40]	0,4875 ± 8,2%
Facteur d'importance	[G_20_R_0 ; G_20_R_135 ; G_20_R_90 ; G_20_R_45]	[G_20_R_0 ; G_20_R_45 ; G_20_R_90 ; G_20_R_135]	0,0006 ± 66%
Valeur médiane de positions	[MBL 20 ; b* ; G_20_R_90]	[MBL 20 ; b* ; G_20_R_0]	0,4533 ± 12%
Valeur moyenne de positions	[MBL 20 ; S ; MBL 10]	[MBL 20 ; S ; MBL 10]	0,4418 ± 12%
Moyenne des facteurs d'importance	[G_20_R_0 ; G_20_R_135 ; G_20_R_45 ; G_20_R_90]	[G_20_R_0 ; G_20_R_90 ; G_20_R_45 ; Moyenne]	0,1161 ± 66%

Tableau 15 : Tableau de comparaison pour les différentes méthodes de sélection pour le marquage anti-Phox2B (Section 10.1).

Un aperçu des images de caractéristiques optimales sélectionnées est présenté dans la Figure 44. Contrairement au marquage NeuN, la pire caractéristique sur les 114 initiales selon la valeur médiane de positions (le second moment angulaire d'Haralick avec un rayon de 5 pixels à 90°) génère une image difficile voire impossible à interpréter visuellement. Suite à la sélection à l'échelle des familles, la pire caractéristique parmi les caractéristiques sélectionnées (la partie réelle du filtre de Gabor d'une longueur d'onde de 10 pixels avec un angle θ de 0°) génère une image peu contrastée mais où il est possible de discerner les neurones d'intérêt.

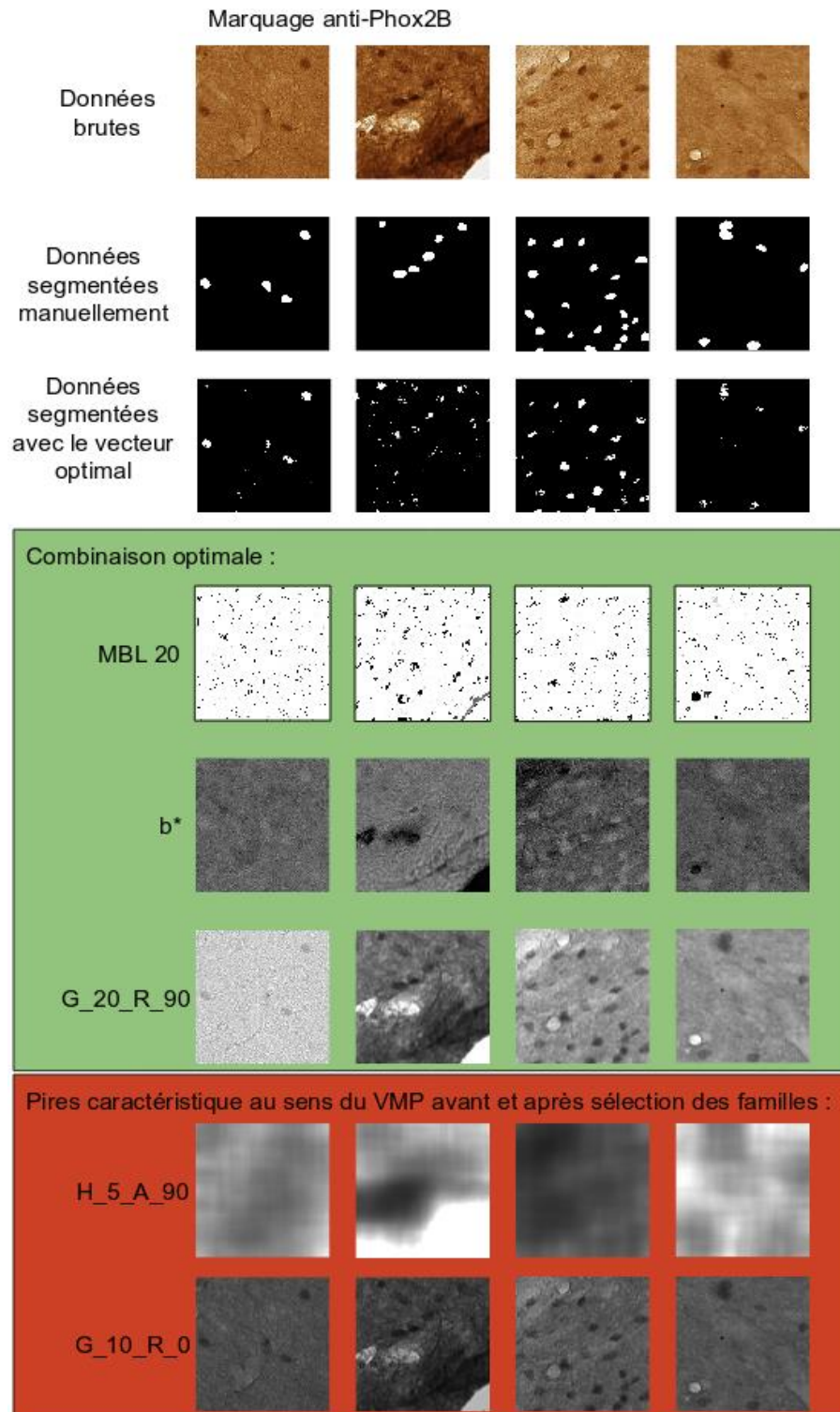


Figure 44 : Les images brutes, les segmentations manuelles et les segmentations automatiques en validation directe sont sur les trois premières lignes. Images des caractéristiques optimales et des pires caractéristiques avant et après la première étape de sélection (familles) selon la VMeP pour le marquage anti-Phox2B. Dans le cadre vert, les images des caractéristiques optimales du marquage anti-Phox2B sont présentées. La pire des 114 caractéristiques initiales est sur la première ligne dans le cadre rouge. La pire des caractéristiques après la première étape de sélection est sur la deuxième ligne dans le cadre rouge.

13.3. Gain computationnel

La réduction du nombre de caractéristiques permet de diminuer : 1) la quantité de mémoire physique utilisée lors de l'étape de segmentation et 2) le temps de segmentation. Il a été ainsi possible de limiter la quantité de données intermédiaires. L'ordinateur utilisé est décrit dans la Section 8.1. Pour ce marquage, nous n'avons pas encore pu utiliser les ressources du TGCC au moment de la rédaction de ce mémoire.

Pour estimer les temps d'extraction de caractéristiques et de segmentation, nous nous sommes basés sur les temps d'extraction et de segmentation de nos bases d'apprentissage et de test. Nos bases combinées ont un poids de 55 Mo. L'extraction des 3 caractéristiques s'est effectuée sur ces deux bases en 2 minutes avec 30 cœurs de calcul et la segmentation en quelques secondes (pour les estimations, nous avons pris 2 secondes). Pour estimer le temps d'extraction et de segmentation sur une coupe entière, nous avons calculé le rapport de proportionnalité.

Contrairement au marquage NeuN, notre optimisation n'est pas rentabilisée dès la première section traitée. En effet, pour une section entière de tronc cérébral, le temps de traitement avec 30 cœurs de calcul est de 2 jours et 16 heures sans l'optimisation contre 10 jours et 4 heures avec l'optimisation. Néanmoins, dès la quatrième section traitée, notre optimisation devient rentable car l'optimisation ne s'effectue qu'une seule fois et que le temps de l'extraction des caractéristiques et de la segmentation avec le vecteur issu de la sélection est très inférieur à celui sans optimisation.

Nombre de cœurs de calcul utilisés	Temps de sélection par notre méthode Force Brute	Nombre de caractéristiques considérées	Temps d'extraction des caractéristiques pour la coupe entière anti-Phox2B	Temps de la segmentation par forêt aléatoire de la coupe entière anti-Phox2B
1	NA	114	~1 mois 15 jours*	~10 jours*
1	~300 jours*	3	~5 jours*	~1h 30min*
30	NA	114	~2 jours*	~16h*
30	~10 jours	3	~4h*	~5min*

Tableau 16 : Tableau des temps de traitement pour une seule image entière de tronc cérébral marquée avec l'anti-Phox2B et numérisée à haute résolution (~7Go). Les deux dernières colonnes correspondent aux étapes de segmentation de la coupe (extraction des caractéristiques optimales et classification par forêt aléatoire). Les temps avec une étoile sont des estimations.

Discussion et Perspectives

La méthode de sélection que j'ai proposée permet **d'améliorer notre compréhension des effets des caractéristiques sur la qualité de la segmentation**. De plus, elle nous permet **de comparer des espaces de caractéristiques très différents, de retrouver des propriétés conformes avec la littérature et d'en découvrir de nouvelles**. Elle permet aussi de **sélectionner de manière stable les caractéristiques** afin de **réduire les besoins en ressources de calcul** pour la segmentation des images histologiques numérisées à haute-résolution. Néanmoins, elle n'est pas très efficace avec le marqueur non-spécifique étudié. La sélection par Force Brute est pensée comme un standard qui nous permettra d'identifier ou de proposer une meilleure méthode de sélection. Cette méthodologie a été appliquée à deux marqueurs différents de neurones de macaque. Le marqueur NeuN, qui a une fixation sur les tissus très spécifique, marque les noyaux de la grande majorité des neurones de la plupart des espèces animales. Le marqueur anti-Phox2B, quant à lui, marque de manière peu spécifique des neurones responsables de certaines fonctions végétatives. Une synthèse de la stratégie de réduction de caractéristiques que je propose est présentée dans la Figure 45.

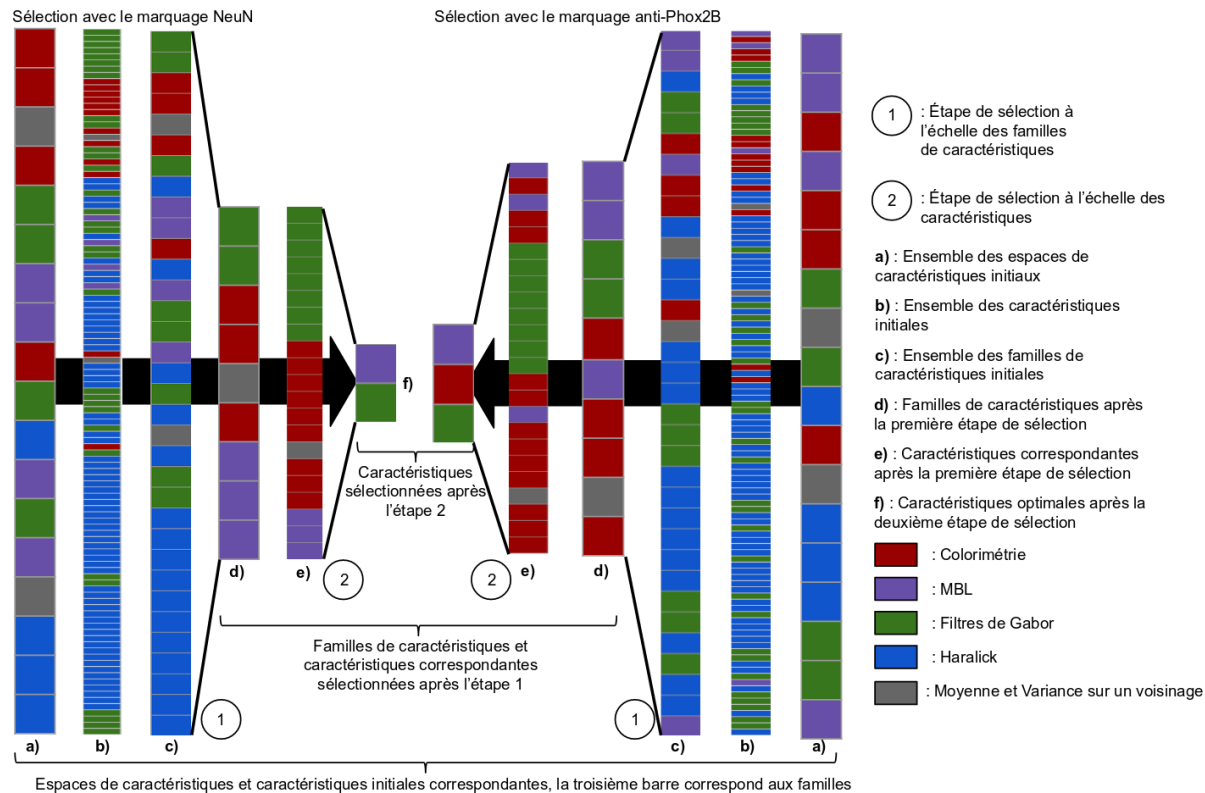


Figure 45 : Synthèse de la stratégie de réduction de caractéristiques proposée. Pour les deux marquages étudiés, deux étapes de sélection sont effectuées afin de réduire progressivement le nombre de caractéristiques à deux pour le marquage NeuN et à trois pour le marquage anti-Phox2B à partir d'un vecteur initial de 114 caractéristiques.

Algorithmes d'apprentissage automatique

Les algorithmes de segmentation ne sont pas tous adaptés aux cas étudiés au cours de ma recherche. Dans le cas du marquage NeuN, les trois algorithmes que nous avons testés

(multi-seuillage, BioVision et Forêt Aléatoire) sont parvenus à segmenter correctement les bases d'apprentissage et de test. BioVision ne prend pas en compte les caractéristiques texturales à cause des profils non-gaussiens de ces caractéristiques. Par exemple, le MBL ayant une valeur pour chaque motif uniforme, ses valeurs ne présentent pas forcément de répartition gaussienne dans l'image, contrairement à l'espace linéaire RVB. Pour le marquage NeuN, les caractéristiques colorimétriques semblent être suffisantes pour réaliser une segmentation de qualité par BioVision. L'algorithme des forêts aléatoires avec ou sans pondération sur la proportion des classes étudiées fonctionne moins bien pour le marquage anti-Phox2B mais mieux que les autres méthodes testées. Cet algorithme est largement utilisé dans le domaine de la segmentation d'images biomédicales [DiFranco *et al.*, 2011][Kallenbach-Thieltges *et al.*, 2013][Yaqub *et al.*, 2014][Vandenbergh *et al.*, 2015] et est reconnu tant pour sa stabilité que pour la reproductibilité de la classification qu'il produit. Les résultats obtenus avec cette méthode pour anti-Phox2B ne sont pas entièrement satisfaisants (les F-Score ne dépassent pas la valeur de 0,5) et il a été nécessaire d'optimiser certains hyperparamètres pour améliorer les F-Scores (Section 11.2). La variation des hyperparamètres a conduit à une variation monotone des F-Scores résultants. Dans la figure représentant la variation du F-Score en fonction de la profondeur d'arbre et celle en fonction du poids de la classe d'intérêt (Figure 30), la courbe du F-Score décrit un plateau quand la profondeur d'arbre devient grande. Dans le cas du poids de la classe d'intérêt, il existe une valeur pour laquelle le F-Score est maximum (après optimisation le F-Score augmente de 26%). Aucun comportement chaotique des courbes de F-Score en fonction des hyperparamètres n'est observable dans ce cas-là, nous avons donc un comportement monotone et le procédé d'optimisation est robuste aux hyperparamètres de l'apprentissage des forêts aléatoires. Ceci n'est pas le cas avec d'autres algorithmes d'apprentissage automatique, notamment les méthodes d'apprentissage profond [Pascanu *et al.*, 2012].

Pour la méthode de multi-seuillage et BioVision appliquées au marquage avec anti-Phox2B, les F-Scores calculées n'ont pas dépassés la valeur de 0,21 ce qui est très insuffisant. Les espaces descriptifs de texture ont été déterminants pour la segmentation de la classe "neurone" marquée par anti-Phox2B avec la forêt aléatoire car il a été possible d'atteindre des F-Scores supérieurs à 0,55 en utilisant cette méthode. De plus, la détection de ces neurones a été difficile car le problème de classification est très complexe et des cas où la segmentation de l'expert était incertaine ont pu être constatés. Cela a pu avoir un impact sur les bases utilisées et peut expliquer la faiblesse des F-Scores résultants par rapport à la segmentation du marquage NeuN. On peut donc imaginer que la segmentation manuelle comporte une part plus importante d'erreur par rapport à celle de NeuN expliquant avec la non-spécificité d'anti-Phox2B la différence de qualité de segmentation obtenue entre les deux marqueurs. Dans ce contexte-là, les F-Scores obtenus avec la forêt aléatoire optimisée en validation directe et croisée (respectivement 0,65 et 0,58 avec 114 caractéristiques) ont été considérés comme satisfaisants, ce qui nous a encouragé à appliquer notre méthode de sélection de caractéristiques. En effet, malgré ces difficultés, il nous a paru pertinent de mener le test de notre méthodologie jusqu'à son terme pour en évaluer les limites. En parallèle de ce travail, le tronc cérébral a été reconstruit en trois dimensions à l'aide de méthodes de recalage et les neurones exprimant Phox2B ont été pointés manuellement par un expert sur les 60 sections histologiques constituant ce jeu de données (~20 000 neurones). Il a été ainsi possible de produire une cartographie 3D de la distribution de ces neurones (Figure 46). Grâce à ce travail, il sera possible de tester le potentiel de notre méthode de segmentation automatique pour produire une cartographie similaire malgré les faibles valeurs de F-Score obtenues. La comparaison de la cartographie manuelle et automatique pourrait être un critère d'évaluation à grande échelle de notre méthodologie de sélection pour le marquage anti-Phox2B. Un article de neuro-anatomie est d'ores et déjà en préparation concernant la détection et la localisation du marqueur anti-Phox2B dans le tronc cérébral de macaque.

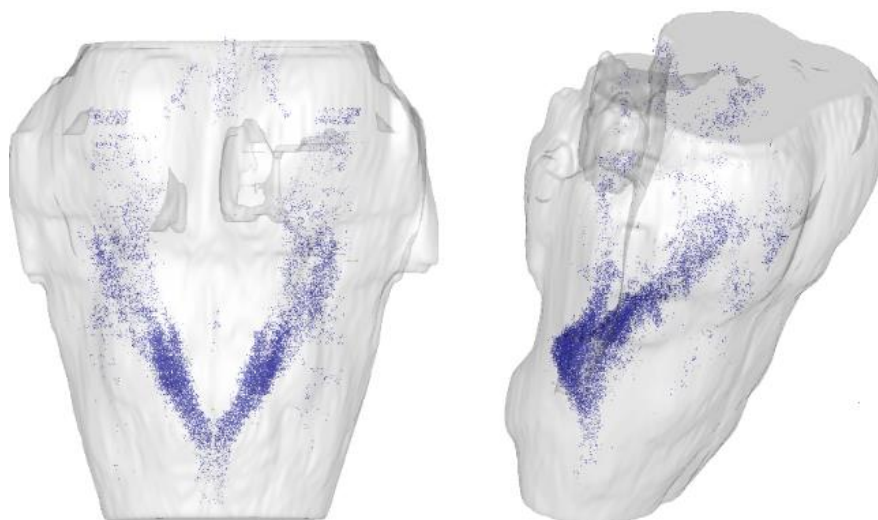


Figure 46 : Distribution des neurones marqués avec anti-Phox2B (points bleus) dans le tronc cérébral d'un macaque reconstruit en 3D (vue de face à gauche et vue oblique à droite).

Pour améliorer la qualité de la segmentation de la classe “neurone” pour le marquage anti-Phox2B, l'utilisation d'algorithmes d'augmentation de contraste ou de normalisation colorimétrique constituent des pistes prometteuses à explorer sur nos données [Li et Plataniotis, 2015]. Une étape de pré- ou post-traitement pourrait être aussi une piste intéressante. En effet, une étape de régularisation par filtre médian de la segmentation automatique a déjà permis d'augmenter la qualité de segmentation pour le marquage NeuN [You, 2017]. Un filtre médian ou morphologique pourrait donc être utilisé dans le cas du marquage anti-Phox2B.

Cohérence des sélections proposées avec les spécificités des marquages

J'ai proposé la création de familles de caractéristiques, une contribution originale de ce travail, afin de proposer une granulométrie intermédiaire entre les caractéristiques seules et leurs espaces associés. Il devient ainsi possible de comparer des groupes de caractéristiques avec des cardinalités semblables. Il n'y aurait aucun sens à comparer l'information fournie par les 16 caractéristiques de l'espace d'Haralick avec celle d'un Motif Binaire Local. Dans le pire des cas, les familles comparées ont une cardinalité de 4 et 1. Les familles sont des sous-espaces de caractéristiques étant non-orientés et partageants les mêmes propriétés d'échelles (tailles des éléments structurants). La prévalence de certaines familles de caractéristiques dans les meilleures combinaisons montre que les espaces linéaires sont les espaces colorimétriques privilégiés dans le cas du marqueur NeuN (CIE XYZ et RVB) et que les espaces CIE $L^*a^*b^*$ et HSV sont privilégiés quant à eux dans le cas du marqueur anti-Phox2B. En effet, les deux caractéristiques sélectionnées pour l'anti-Phox2B sont b^* pour la valeur médiane de positions et S pour la valeur moyenne de positions. La caractéristique b^* représente l'axe de jaune-bleu intégrant la saturation exprimée par S dans l'espace HSV. La saturation est une information importante pour le marquage anti-Phox2B. Les neurones ciblés ne sont pas détectés spécifiquement par ce marqueur, une différence d'intensité de marquage est alors recherchée. Le niveau de saturation permet donc à l'algorithme d'apprentissage automatique de mieux classifier les neurones d'intérêt. De plus, la sélection de la caractéristique b^* est cohérente avec les conclusions émises dans notre publication à ICPRAI 2018 [Bouvier *et al.*, 2018]. Nous avons alors étudié le marquage NeuN sans les filtres de Gabor en nous concentrant principalement sur les espaces colorimétriques. La caractéristique Z représentant la teinte de bleu dans CIE XYZ avait alors été sélectionnée. Nous avons dès lors confirmé la cohérence de cette sélection avec

la théorie du mélange de couleur. En effet, les teintes brunes sont obtenues avec des niveaux (quantité) de rouge constants, les niveaux des teintes vertes et bleues servant à ajuster la teinte de brun [Witowski et Brown, 1977]. Le bleu étant la couleur primaire la plus discriminante pour les teintes brunes, il est intéressant de constater sa présence dans les caractéristiques sélectionnées par notre méthode pour le marquage anti-Phox2B, un marquage qui est globalement de couleur brune.

Les parties réelles de Gabor ont également été sélectionnées automatiquement pour les deux marqueurs histologiques que nous avons étudiés. Ce choix est d'autant plus favorable pour nous que leur complexité algorithmique de calcul est faible. Nous avons vu Section 5.3.2 que la partie réelle des filtres de Gabor agissent comme des détecteur de régions [Mehrotra *et al.*, 1992], ce qui suggère que le contraste entre les neurones et le tissu non-marqué sont une information essentielle à la détection de ces derniers. Les contours peuvent être vus comme l'accumulation locale de motifs uniformes comme ceux détectés par les Motifs Binaires Locaux. En plus d'avoir une complexité algorithmique linéaire par rapport au rayon de l'élément structurant (voir Section 5.3.3), les MBL ont également été sélectionnés pour les deux marquages histologiques étudiés que ce soit à l'échelle des familles ou celle des caractéristiques. Nous avons donc l'alliance entre un détecteur de contour et de région. Par ailleurs, ces espaces de caractéristiques sont particulièrement adaptés pour notre objectif de minimisation du fait de leur cardinalité. Les espaces de Gabor comme les MBL sont plus efficaces avec l'augmentation de la taille de la structure étudiée, ce qui s'exprime par le rayon de l'élément structurant pour le MBL et par l'hyperparamètre λ pour les filtres de Gabor (Sections 12 et 13). Pour le MBL, on peut néanmoins remarquer que le plus grand rayon proposé pour le marqueur NeuN n'a pas été sélectionné durant l'étape de sélection des familles, ce qui suggère l'existence d'un rayon optimal pour le MBL proche du rayon maximal des neurones marqués à l'anticorps NeuN. Contrairement aux filtres de Gabor et au MBL, les espaces d'Haralick semblent être plus efficaces avec des rayons de petite dimension. Bien que non sélectionnée, la corrélation d'Haralick avec le rayon d'élément structurant minimum (5 pixels pour l'anti-Phox2B et 11 pixels pour le NeuN) est une famille de caractéristiques dont toutes les caractéristiques sont avantageuses (leurs VMP est inférieure à la moyenne des VMP) dans tous les cas.

Les espaces d'Haralick sont donc performants pour analyser des micro-textures, c'est à dire à l'échelle du plus petit élément étudié. De plus, la présence constante de cette famille d'Haralick dans les meilleures familles en termes de VMP suggère qu'elle pourrait être intéressante à intégrer dans la deuxième étape de sélection de caractéristiques. Néanmoins, les familles d'Haralick sont les plus complexes à calculer et la cardinalité de l'espace, duquel elles sont issues, est très important ($n=16$) ce qui nous pénaliserait très fortement du point de vue du coût de calcul. Nous n'avons pas intégré cet aspect computationnel dans le protocole global de sélection de caractéristiques que nous avons proposé mais il serait intéressant d'explorer cette voie sous la forme d'une pondération ou d'une pénalité à appliquer pour automatiser complètement le procédé. Cela nécessiterait probablement de définir empiriquement des règles pour ajuster ces paramètres en fonction des ressources informatiques et du temps disponible. De plus, la cardinalité de l'espace d'Haralick est un véritable obstacle à son exploration couplée à un objectif de minimisation. En effet, même en incluant un espace d'Haralick dans l'étape de sélection de caractéristiques, nous serions toujours limités par la profondeur d'exploration (que nous avons fixé à 4 pour les deux marqueurs et pour la station de travail utilisée). Nous pourrions en utilisant les ressources de calcul haute-performance augmenter cette limite, mais même dans ce cas, il sera difficile d'atteindre une profondeur d'exploration d'une dizaine d'élément (par exemple, il y a 1 961 256 combinaisons de 10 caractéristiques dans le vecteur de 24 caractéristiques issue de la première étape de sélection pour le marquage anti-Phox2B).

Pour améliorer l'exploration des espaces et des familles de caractéristiques par Force Brute, une des idées principales serait de déterminer un supremum (le plus petit des maximum)

de la taille du vecteur de caractéristiques optimisé nécessaire afin d'approcher la qualité en terme de F-Score du vecteur initial. De cette façon, la profondeur d'exploration pourrait être anticipée. Pour estimer ce supremum, le calcul du F-Score de plusieurs combinaisons de différentes tailles permettrait de déterminer à partir de quelle taille la majorité des combinaisons ont une qualité satisfaisante (la majorité des combinaisons tirées aléatoirement a un F-Score satisfaisant). Une fois ce supremum déterminé, une autre idée serait de créer une heuristique ou un échantillonnage pour explorer les combinaisons trop grandes pour nos ressources de calcul. La validation de cette approche pourrait être basée sur la convergence de la sélection vers un vecteur stable. Nous pourrions nous inspirer des méthodologies multi-échelles ou des algorithmes génétiques pour accomplir cette amélioration [Barker *et al.*, 2016][Fatima *et al.*, 2017]. Une autre piste d'investigation de nos travaux serait d'étendre l'étude à d'autres familles de caractéristiques (par exemple les autres caractéristiques statistiques d'Haralick ou des ondelettes), d'autres marqueurs et d'autres animaux. En effet, la méthode étant suffisamment générique, on pourrait comparer les familles privilégiées pour des marqueurs ayant des propriétés similaires ou très éloignées et établir des préconisations afin de proposer des vecteurs initiaux de caractéristiques primaires dans différentes configurations. Cela permettrait de gagner du temps sur les deux étapes de sélection de caractéristiques. Il n'existe pas, à notre connaissance, de procédure standardisée permettant d'identifier de manière fiable et reproductible des vecteurs de caractéristiques optimaux pour les données d'histologie dans la mesure où il existe une très grande diversité de problématiques (marqueurs histologiques, organes, espèces, pathologies, etc.). Les domaines pouvant bénéficier de mes développements méthodologiques iraient de l'histologie préclinique à l'anatomopathologie clinique, cette dernière étant probablement le premier producteur d'images histologiques actuellement. Par exemple, le diagnostic d'un cancer de la prostate pour un patient requiert d'étudier entre 12 et 20 sections de biopsies, soit approximativement entre 2,5 et 4 millions de pixels générés [Gurcan *et al.*, 2010]. Même si le cerveau de macaque marqué avec l'anticorps NeuN a approximativement 1 000 fois plus de pixels, le nombre de patients biopsiés par année dépasse les 100 000 en France [Gurcan *et al.*, 2010][Jéhannin-Ligier *et al.*, 2017].

Stabilité de la sélection

L'étape de sélection des caractéristiques à partir des familles sélectionnées est stable pour les deux marqueurs étudiés. Les valeurs moyennes et médianes de positions ont donné des sous-ensembles optimaux de caractéristiques différents pour les deux marquages. Au niveau des familles, ces deux valeurs nous ont donné des classements quasiment identiques. La représentation d'une caractéristique dans les meilleurs résultats s'est avérée être une information fondamentale et pertinente pour assurer la stabilité du procédé de sélection proposé. Notons enfin que la moyenne et la médiane ont donné des résultats de sélection très proches lors de l'exploitation de mon critère VMP. Néanmoins, la dynamique de la médiane étant plus grande d'environ 25% par rapport à la dynamique de la moyenne, nous préconisons d'utiliser la médiane qui, de plus, sélectionne des combinaisons de caractéristiques de cardinalité plus faible par rapport à la moyenne.

Le sous-ensemble optimal pour le marqueur NeuN a une taille de deux et un F-Score très proche du vecteur initial de 114 caractéristiques. Il est possible de réaliser la détection du marqueur NeuN avec un nombre réduit de caractéristiques. Ce n'est cependant pas le cas pour la détection de la classe "neurone" dans la base anti-Phox2B qui nécessiterait un plus grand nombre de caractéristiques ($n \geq 3$). En effet, l'algorithme ne rencontre pas sa condition d'arrêt et le F-Score du sous-ensemble stable de caractéristiques (0,45 en moyenne) n'est pas proche du F-Score du vecteur initial (0,61 en moyenne). Cela suggère que l'optimisation n'est peut-être pas finie et qu'il faudrait étendre la profondeur d'exploration avec des vecteurs de plus grandes tailles. Notre condition d'arrêt pourrait être améliorée en prenant en compte la stabilisation des

F-Scores que nous avons visualisée pour le marquage NeuN dans la Figure 37 (Section 12.2.2). En effet, les F-Scores calculés atteignent rapidement des valeurs supérieures à 0,8 et il pourrait être intéressant d'utiliser un critère prenant en compte cette convergence. L'algorithme se stopperait à l'atteinte du plateau s'il existe pour le problème de classification posé.

Pour mieux caractériser les vecteurs de caractéristiques optimaux, une autre piste intéressante serait de tester la séparabilité linéaire de l'ensemble de caractéristiques ainsi formé. Pour cela, il faudrait étendre l'exploration par Force Brute des combinaisons de caractéristiques pour le marqueur anti-Phox2B afin de trouver le sous-ensemble optimal ayant un F-Score associé proche de celui calculé avec le vecteur initial. Puis nous utiliserions ce sous-ensemble optimal pour entraîner la méthode d'apprentissage supervisé de machine à support de vecteurs (Section 5.2.1). L'obtention de bons résultats en termes de F-Score signifierait que notre méthodologie sélectionnerait des espaces de caractéristique linéairement séparable.

Pour les marqueurs NeuN et anti-Phox2B, les vecteurs jugés optimaux sont ceux sélectionnés avec la valeur médiane de positions. Ce vecteur est différent entre les différentes conditions de validation. En effet, la caractéristique de Gabor (deuxième sélectionnée pour NeuN, troisième pour anti-Phox2B) a une orientation différente entre les conditions de validation, néanmoins elles sont toutes les deux issues de la même famille de caractéristiques. Les neurones marqués avec l'anticorps NeuN et anti-Phox2B n'ont pas d'orientation privilégiées, l'angle a donc peu d'importance. Les Figures 34, 35, 40 et 41 étayaient cette conclusion, les familles de Gabor utilisées dans la deuxième étape de sélection sont compactes dans notre représentation. Tant que les autres hyperparamètres sont fixés, les différentes caractéristiques des familles de Gabor sélectionnées sont interchangeable. Lorsqu'il ne semble pas y avoir d'orientations privilégiées, nous pourrions calculer une nouvelle caractéristique représentant la moyenne des différentes caractéristiques sur tous les angles (on pourrait appliquer ce principe sur les caractéristiques d'Haralick comme le suggérait Haralick dans son article [Haralick *et al.*, 1973]). Nos vecteurs de caractéristiques seraient alors considérés comme stables. Nous avons observé que les différents vecteurs sélectionnés sont pour la plupart stables (sauf pour celui sélectionné par Force Brute seule). Cette propriété s'explique par la sélection préliminaire des familles que nous avons proposée et qui élimine les caractéristiques avec de faibles valeurs médianes de positions. Cela augmente la moyenne des F-Scores des combinaisons analysées (Figure 37 et 43) et la possibilité d'obtenir des combinaisons similaires entre les deux conditions de validation. Néanmoins, le facteur d'importance issu de la méthode de sélection incluse dans l'apprentissage des forêts aléatoires est biaisé vers les caractéristiques de Gabor pour les deux marquages. Dans le cas du marquage NeuN, les caractéristiques de Gabor seules permettent d'obtenir des segmentations de qualité (avec un F-Score supérieur à 0,8). Néanmoins, pour le marquage anti-Phox2B, les caractéristiques de Gabor seules ne permettent pas de détecter les neurones d'intérêts (le F-Score atteint au maximum 0,12 avec l'intégration de l'image moyenne, voir Section 13.2.2). Dans mon étude, la sélection de caractéristiques par les facteurs d'importances est biaisée envers les caractéristiques issues des familles de Gabor et ne permet donc pas une sélection pertinente pour tous les marquages (comme nous avons pu le constater pour le marquage anti-Phox2B). Nous ne préconisons donc pas l'utilisation de ce critère de sélection.

De manière plus générale, notre méthodologie est une méthode de sélection de caractéristiques intégrée et la recherche par Force Brute est la plus faible des heuristiques dans la mesure où elle est exhaustive pour un problème donné. Il serait donc intéressant de comparer des sous-ensembles sélectionnés par notre méthode avec les sous-ensembles sélectionnés par d'autres méthodes de sélection automatique de caractéristiques, comme les algorithmes génétiques, la méthode de minimum redondance maximum pertinence ou les méthodes à filtres. Cette comparaison permettrait de compléter les résultats obtenus avec la sélection par Force Brute (première combinaison en termes de F-Score) et par la sélection incluse dans

l'apprentissage des forêts aléatoire (facteur d'importance). Les critères de comparaison pourraient être : 1) la stabilité des sous-ensembles sélectionnés pour les validations directe et croisée, 2) la qualité de la segmentation résultante en terme de F-Score, 3) la complexité algorithmique de l'extraction du sous-ensemble et enfin 4) la cardinalité du sous-ensemble optimal. Cette étude comparative serait l'occasion d'évaluer la pertinence de ces méthodes dans le cadre de la segmentation de données histologiques et de faire des préconisations quant à leur cadre de validité.

Coût calculatoire et industrialisation

Même si la recherche par Force Brute que nous proposons peut sembler coûteuse en termes de ressources computationnelles, elle apporte un bénéfice rapidement lorsqu'une segmentation doit être réalisée sur des données massives. En effet, la réduction du nombre de caractéristiques a été rentabilisée dans notre cas dès le traitement de la première section de cerveau de macaque marquée avec NeuN (dès quatre sections pour anti-Phox2B). Dans ce contexte, le calcul des caractéristiques est l'étape la plus coûteuse dans le procédé global de segmentation. Dans le cadre de la segmentation des images histologiques numérisées à haute-résolution, la production sériée de sections rend indispensable la réduction du nombre de caractéristiques à utiliser. Plus généralement, la sélection de caractéristiques devrait être une étape fondamentale dans la conception des chaînes de traitement d'apprentissage automatique pour les données massives. En effet, la consommation énergétique des supercalculateurs et l'atteinte des limites physiques des semi-conducteurs ne permettent plus de baser la conception logicielle uniquement sur l'évolution future des technologies matérielles [Witkowski *et al.*, 2013][Thompson et Parthasarathy, 2016]. Il faut dorénavant penser à rationaliser l'exploitation des ressources de calculs et proposer des méthodologies adaptées. Notre recherche par Force Brute adaptée étant une preuve de concept pour la sélection stable des caractéristiques, notre futur travail sera de trouver des méthodologies complémentaires permettant d'atteindre les mêmes résultats avec un coût plus faible en ressources de calcul.

Cette méthode a été développée au sein de la plateforme logicielle BrainVISA et sera intégrée prochainement dans la future plateforme de micro-services 3DNeuroSecure, plateforme de calcul haute-performance de NEOXIA. Ce travail s'inscrit dans le cadre d'un projet d'investissement d'avenir qui a financé ma bourse CIFRE. Dès le début du doctorat, les codes ont été implémentés en ayant pour objectif la compatibilité avec des infrastructures de calcul haute-performance CPU. Cela a pu être démontré avec l'utilisation de mon exploration par Force Brute au TGCC du CEA. Les procédés d'extraction de caractéristiques et de segmentation par l'algorithme des forêts aléatoires sont en cours de test au TGCC. La prochaine étape sera de créer des micro-services, procédés indépendants remplissant une tâche spécifique, reproduisant les résultats présentés dans ce manuscrit. Ces micro-services créés, nous pourrions déployer mes procédés sur des ressources de calcul haute-performance virtualisées (*Cloud Computing*) ce qui devrait permettre de rendre accessible ces micro-services à des équipes ne disposant pas de ressources de calcul intensif.

Perspectives générales

D'autres chaînes de traitement comme 3D-HAPi [Vandenberghe *et al.*, 2016] et l'individualisation des neurones peuvent profiter directement des procédés proposés (Figure 47) [You *et al.*, 2016]. En effet, ces deux chaînes de traitement utilisent la segmentation des tissus histologiques marqués au niveau des pixels et peuvent bénéficier directement de l'optimisation réalisée et des plateformes de calcul haute-performance. Ceci permettrait de mener des études exhaustives à plus grande échelle avec plus de sections et sur d'autres d'organes à segmenter. 3D-HAPi et l'individualisation des neurones sont des procédés directement impliqués dans la

recherche préclinique dans le cadre de la maladie d'Alzheimer, le premier permettant la quantification des marqueurs multi-modalités dans des études de groupe (avec le marqueur BAM-10 pour les plaques amyloïdes et IBA1/CD68 pour l'étude de la microglie par exemple) et le deuxième la quantification de la mort neuronale (marquage NeuN). Un article pour lequel je suis co-auteur est en cours de rédaction concernant l'individualisation des neurones.

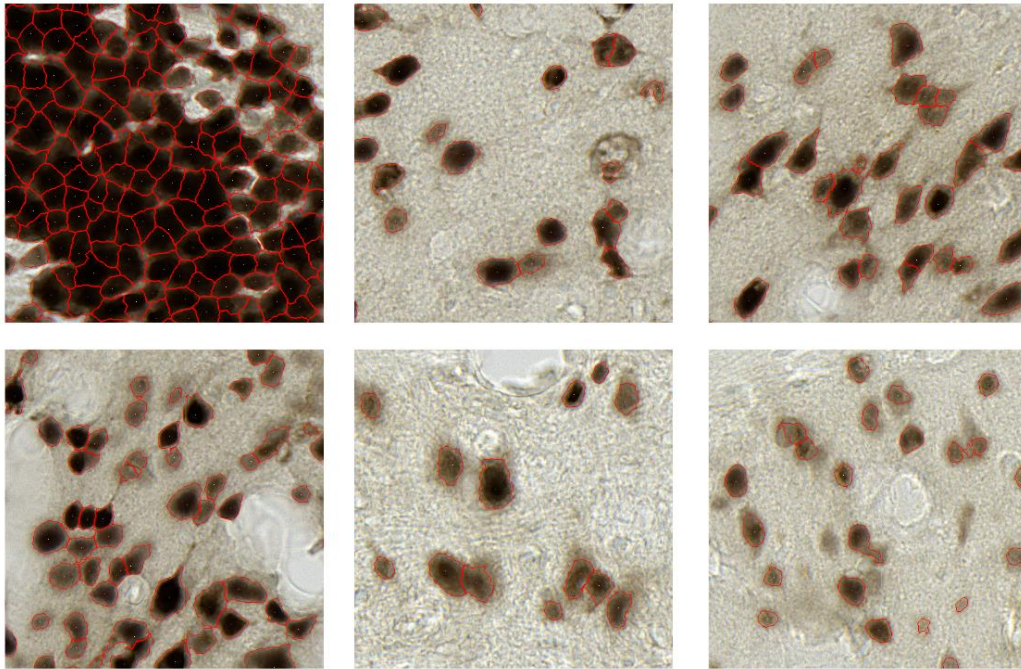


Figure 47 : Différents résultats d'individualisation de neurones marqués avec l'anticorps NeuN à partir des segmentations binaires réalisées à l'échelle du pixel [You *et al.*, 2016]. Les lignes rouges représentent les contours estimés pour chaque neurone [You, thèse, 2017].

L'histologie réalisée avec des marqueurs fluorescents n'a pas été étudiée dans le cadre de ce travail de thèse. Il n'y a pas de couleur à proprement parler dans ce type de données, les marqueurs réagissant à une longueur d'onde spécifique et produisant des images monochromes. Néanmoins, l'emploi des caractéristiques texturales peuvent être étendues aux tissus histologiques étudiés par fluorescence [Atlamazoglou *et al.*, 2001][Ericson *et al.*, 2005]. Une piste intéressante à considérer serait d'optimiser les caractéristiques texturales pour ce type d'images. Ce travail pourrait aussi être étendu aux numérisations de tissus clarifiés qui sont produites avec des microscopes à feuille de lumière (Section 2.2.1). Les images tridimensionnelles peuvent être considérées, à l'instar des organes reconstruits, comme une concaténation d'images planaires sur lesquelles la méthodologie proposée pourrait être appliquée. De plus, il serait intéressant de développer des méthodes et de nouvelles caractéristiques pour tirer avantage de la nature tridimensionnelle des images de tissus clarifiés. Le principal problème devrait venir des espaces de description de texture. En effet, il n'existe pas de définition formelle de la texture bidimensionnelle, chaque espace utilisant sa propre définition de la notion de texture (Section 5.3.2). La définition d'une texture tridimensionnelle devient dès lors complexe.

Pour conclure, notre méthodologie a été appliquée sur deux marqueurs révélant des neurones sur deux régions du système nerveux central de macaque. Les marquages résultants sont fondamentalement différents et nécessiteraient des études quantitatives au niveau des organes entiers pour confirmer nos résultats.

Annexe 1 : Régions sélectionnées pour la base NeuN

Numéro	Nom de l'image	Latéralité	Région	Base
1	CJ1301_slide81unsharp_BG_R	Droite	Fond (Background)	Apprentissage
2	CJ1301_slide81unsharp_CA1_R	Droite	<i>Cornus Ammonis</i> 1 (hippocampe)	Apprentissage
3	CJ1301_slide81unsharp_CA2_R	Droite	<i>Cornus Ammonis</i> 2 (hippocampe)	Apprentissage
4	CJ1301_slide81unsharp_CA3_R	Droite	<i>Cornus Ammonis</i> 3 (hippocampe)	Apprentissage
5	CJ1301_slide81unsharp_CA4_R	Droite	<i>Cornus Ammonis</i> 4 (hippocampe)	Apprentissage
6	CJ1301_slide81unsharp_caudate_L	Gauche	Noyau caudé	Test
7	CJ1301_slide81unsharp_caudate_R	Droite	Noyau caudé	Apprentissage
8	CJ1301_slide81unsharp_claustrum_L	Gauche	Clastrum	Test
9	CJ1301_slide81unsharp_claustrum_R	Droite	Clastrum	Apprentissage
10	CJ1301_slide81unsharp_ctx_L_layer1	Gauche	Couche corticale 1	Test
11	CJ1301_slide81unsharp_ctx_L_layer2	Gauche	Couche corticale 2	Apprentissage
12	CJ1301_slide81unsharp_ctx_L_layer3	Gauche	Couche corticale 3	Test
13	CJ1301_slide81unsharp_ctx_L_layer4	Gauche	Couche corticale 4	Test
14	CJ1301_slide81unsharp_ctx_L_layer5a	Gauche	Couche corticale 5a	Apprentissage
15	CJ1301_slide81unsharp_ctx_L_layer5b	Gauche	Couche corticale 5b	Apprentissage
16	CJ1301_slide81unsharp_ctx_L_layer6a	Gauche	Couche corticale 6a	Apprentissage
17	CJ1301_slide81unsharp_ctx_L_layer6b	Gauche	Couche corticale 6b	Test
18	CJ1301_slide81unsharp_ctx_R_layer1	Droite	Couche corticale 1	Apprentissage
19	CJ1301_slide81unsharp_ctx_R_layer2	Droite	Couche corticale 2	Test
20	CJ1301_slide81unsharp_ctx_R_layer3	Droite	Couche corticale 3	Apprentissage
21	CJ1301_slide81unsharp_ctx_R_layer4	Droite	Couche corticale 4	Apprentissage
22	CJ1301_slide81unsharp_ctx_R_layer5a	Droite	Couche corticale 5a	Test
23	CJ1301_slide81unsharp_ctx_R_layer5b	Droite	Couche corticale 5b	Test
24	CJ1301_slide81unsharp_ctx_R_layer6a	Droite	Couche corticale 6a	Test
25	CJ1301_slide81unsharp_ctx_R_layer6b	Droite	Couche corticale 6b	Apprentissage
26	CJ1301_slide81unsharp_DG_R	Droite	Gyrus denté	Apprentissage
27	CJ1301_slide81unsharp_globus_pallidus_ext_L	Gauche	<i>Globus pallidus</i>	Test

28	CJ1301_slide81unsharp_globus_pallidus_ext_R	Droite	<i>Globus pallidus</i>	Apprentissage
29	CJ1301_slide81unsharp_globus_pallidus_int_L	Gauche	<i>Globus pallidus</i>	Apprentissage
30	CJ1301_slide81unsharp_globus_pallidus_int_R	Droite	<i>Globus pallidus</i>	Test
31	CJ1301_slide81unsharp_hypotal	Centre	Hypotal	Apprentissage
32	CJ1301_slide81unsharp_mam_body_L	Gauche	Corps mamillaires	Apprentissage
33	CJ1301_slide81unsharp_mam_body_R	Droite	Corps mamillaires	Test
34	CJ1301_slide81unsharp_putamen_L	Gauche	Putamen	Apprentissage
35	CJ1301_slide81unsharp_putamen_R	Droite	Putamen	Test
36	CJ1301_slide81unsharp_subiculum_para_R	Droite	Para-subiculum	Apprentissage
37	CJ1301_slide81unsharp_subiculum_R	Droite	Subiculum	Apprentissage
38	CJ1301_slide81unsharp_thalamus_L	Gauche	Thalamus	Apprentissage
39	CJ1301_slide81unsharp_thalamus_R	Droite	Thalamus	Test
40	CJ1301_slide81unsharp_ventricle_R	Droite	Ventricule	Test
41	CJ1301_slide81unsharp_WM_L	Gauche	Matière blanche	Test
42	CJ1301_slide81unsharp_WM_R	Droite	Matière blanche	Apprentissage
43	CJ1301_slide101unsharp_BG_R	Droite	Fond (background)	Test
44	CJ1301_slide101unsharp_CA1_L	Gauche	<i>Cornus Ammonis</i> 1 (hippocampe)	Apprentissage
45	CJ1301_slide101unsharp_CA1_R	Droite	<i>Cornus Ammonis</i> 1 (hippocampe)	Test
46	CJ1301_slide101unsharp_CA2_L	Gauche	<i>Cornus Ammonis</i> 2 (hippocampe)	Apprentissage
47	CJ1301_slide101unsharp_CA2_R	Droite	<i>Cornus Ammonis</i> 2 (hippocampe)	Test
48	CJ1301_slide101unsharp_CA3a_L	Gauche	<i>Cornus Ammonis</i> 3a (hippocampe)	Apprentissage
49	CJ1301_slide101unsharp_CA3a_R	Droite	<i>Cornus Ammonis</i> 3a (hippocampe)	Test
50	CJ1301_slide101unsharp_CA3b_L	Gauche	<i>Cornus Ammonis</i> 3b (hippocampe)	Test
51	CJ1301_slide101unsharp_CA3b_R	Droite	<i>Cornus Ammonis</i> 3b (hippocampe)	Apprentissage
52	CJ1301_slide101unsharp_CA4_L	Gauche	<i>Cornus Ammonis</i> 4 (hippocampe)	Apprentissage
53	CJ1301_slide101unsharp_CA4_D	Droite	<i>Cornus Ammonis</i> 4 (hippocampe)	Test
54	CJ1301_slide101unsharp_caudate_L	Gauche	Noyau caudé	Apprentissage
55	CJ1301_slide101unsharp_caudate_R	Droite	Noyau caudé	Test
56	CJ1301_slide101unsharp_ctx_L_layer1	Gauche	Couche corticale 1	Apprentissage
57	CJ1301_slide101unsharp_ctx_L_layer2	Gauche	Couche corticale 2	Test
58	CJ1301_slide101unsharp_ctx_L_layer3	Gauche	Couche corticale 3	Apprentissage
59	CJ1301_slide101unsharp_ctx_L_layer4	Gauche	Couche corticale 4	Apprentissage

60	CJ1301_slide101unsharp_ctx_L_layer5a	Gauche	Couche corticale 5a	Test
61	CJ1301_slide101unsharp_ctx_L_layer5b	Gauche	Couche corticale 5b	Test
62	CJ1301_slide101unsharp_ctx_L_layer6a	Gauche	Couche corticale 6a	Test
63	CJ1301_slide101unsharp_ctx_L_layer6b	Gauche	Couche corticale 6b	Apprentissage
64	CJ1301_slide101unsharp_ctx_R_layer1	Droite	Couche corticale 1	Test
65	CJ1301_slide101unsharp_ctx_R_layer2	Droite	Couche corticale 2	Apprentissage
66	CJ1301_slide101unsharp_ctx_R_layer3	Droite	Couche corticale 3	Test
67	CJ1301_slide101unsharp_ctx_R_layer4	Droite	Couche corticale 4	Test
68	CJ1301_slide101unsharp_ctx_R_layer5a	Droite	Couche corticale 5a	Apprentissage
69	CJ1301_slide101unsharp_ctx_R_layer5b	Droite	Couche corticale 5b	Apprentissage
70	CJ1301_slide101unsharp_ctx_R_layer6a	Droite	Couche corticale 6a	Apprentissage
71	CJ1301_slide101unsharp_ctx_R_layer6b	Droite	Couche corticale 6b	Test
72	CJ1301_slide101unsharp_DG_L	Gauche	Gyrus denté	Apprentissage
73	CJ1301_slide101unsharp_DG_R	Droite	Gyrus denté	Test
74	CJ1301_slide101unsharp_habenula_L	Gauche	Habenula	Test
75	CJ1301_slide101unsharp_habenula_R	Droite	Habenula	Apprentissage
76	CJ1301_slide101unsharp_LG_L_1	Gauche	Corps genouillé latéral	Apprentissage
77	CJ1301_slide101unsharp_LG_L_2	Gauche	Corps genouillé latéral	Test
78	CJ1301_slide101unsharp_LG_L_3	Gauche	Corps genouillé latéral	Apprentissage
79	CJ1301_slide101unsharp_LG_L_4	Gauche	Corps genouillé latéral	Test
80	CJ1301_slide101unsharp_LG_L_5	Gauche	Corps genouillé latéral	Apprentissage
81	CJ1301_slide101unsharp_LG_L_6	Gauche	Corps genouillé latéral	Test
82	CJ1301_slide101unsharp_LG_R_1	Droite	Corps genouillé latéral	Test
83	CJ1301_slide101unsharp_LG_R_2	Droite	Corps genouillé latéral	Apprentissage
84	CJ1301_slide101unsharp_LG_R_3	Droite	Corps genouillé latéral	Test
85	CJ1301_slide101unsharp_LG_R_4	Droite	Corps genouillé latéral	Apprentissage
86	CJ1301_slide101unsharp_LG_R_5	Droite	Corps genouillé latéral	Apprentissage
87	CJ1301_slide101unsharp_LG_R_6	Droite	Corps genouillé latéral	Test
88	CJ1301_slide101unsharp_PAG_L	Gauche	Substance grise périaqueducule	Test
89	CJ1301_slide101unsharp_PAG_R	Droite	Substance grise périaqueducule	Apprentissage
90	CJ1301_slide101unsharp_pons_L	Gauche	Pons	Apprentissage
91	CJ1301_slide101unsharp_pons_R	Droite	Pons	Test

92	CJ1301_slide101unsharp_subiculum_para_L	Gauche	Para-subiculum	Apprentissage
93	CJ1301_slide101unsharp_subiculum_para_R	Droite	Para-subiculum	Test
94	CJ1301_slide101unsharp_subiculum_L	Gauche	Subiculum	Apprentissage
95	CJ1301_slide101unsharp_subiculum_R	Droite	Subiculum	Test
96	CJ1301_slide101unsharp_thalamus_L	Gauche	Thalamus	Test
97	CJ1301_slide101unsharp_thalamus_R	Droite	Thalamus	Apprentissage
98	CJ1301_slide101unsharp_ventricle_R	Droite	Ventricule	Apprentissage
99	CJ1301_slide101unsharp_WM_L	Gauche	Matière blanche	Apprentissage
100	CJ1301_slide101unsharp_WM_R	Droite	Matière blanche	Test

Tableau 17 : Liste des images de la base NeuN avec leur latéralité (D pour droite et G pour gauche), la région d'extraction et leur inclusion dans la base d'apprentissage (A) ou de test (T)

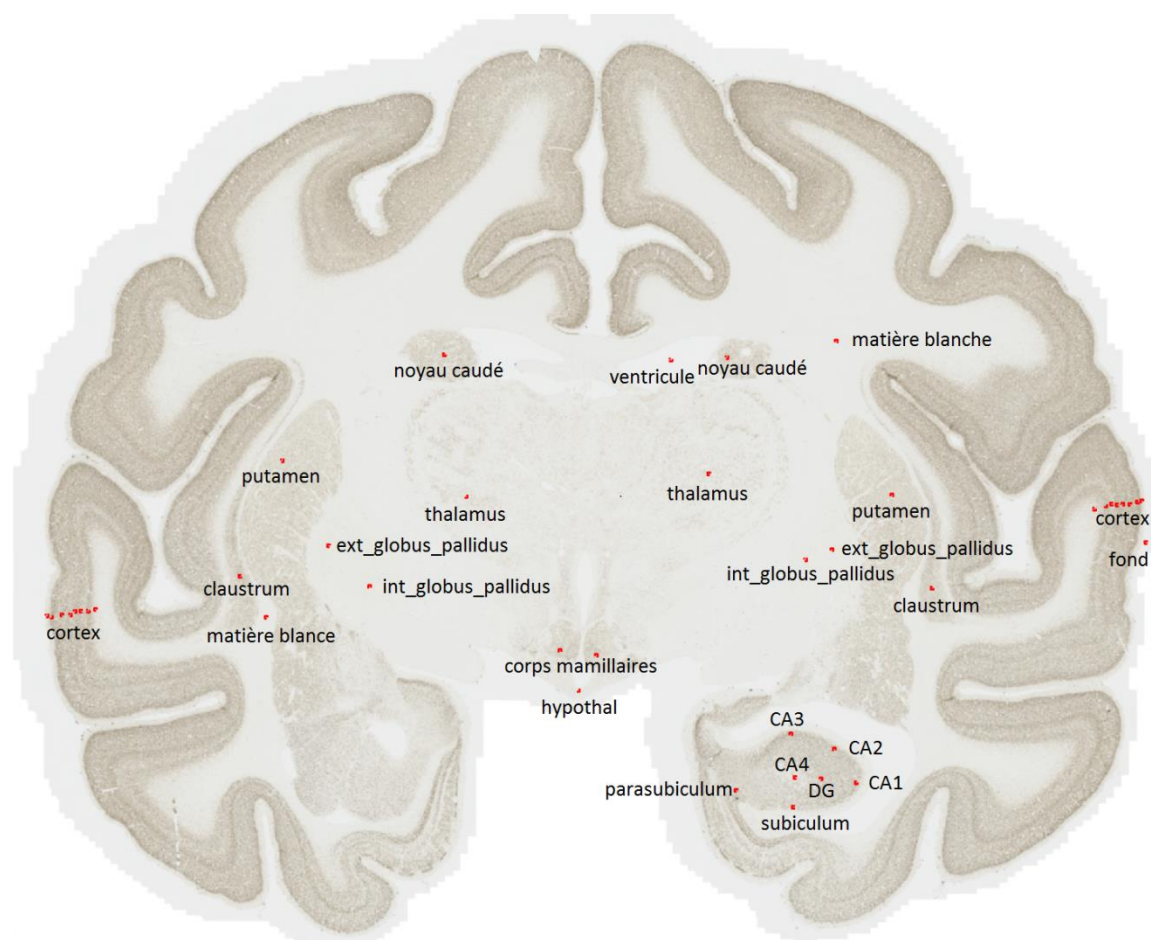


Figure 45 : Section 81 numérisée avec les emplacements des images extraites pour la base d'apprentissage et de test indiqués en rouge(tiré de la thèse de Zhenzhen You, 2017).

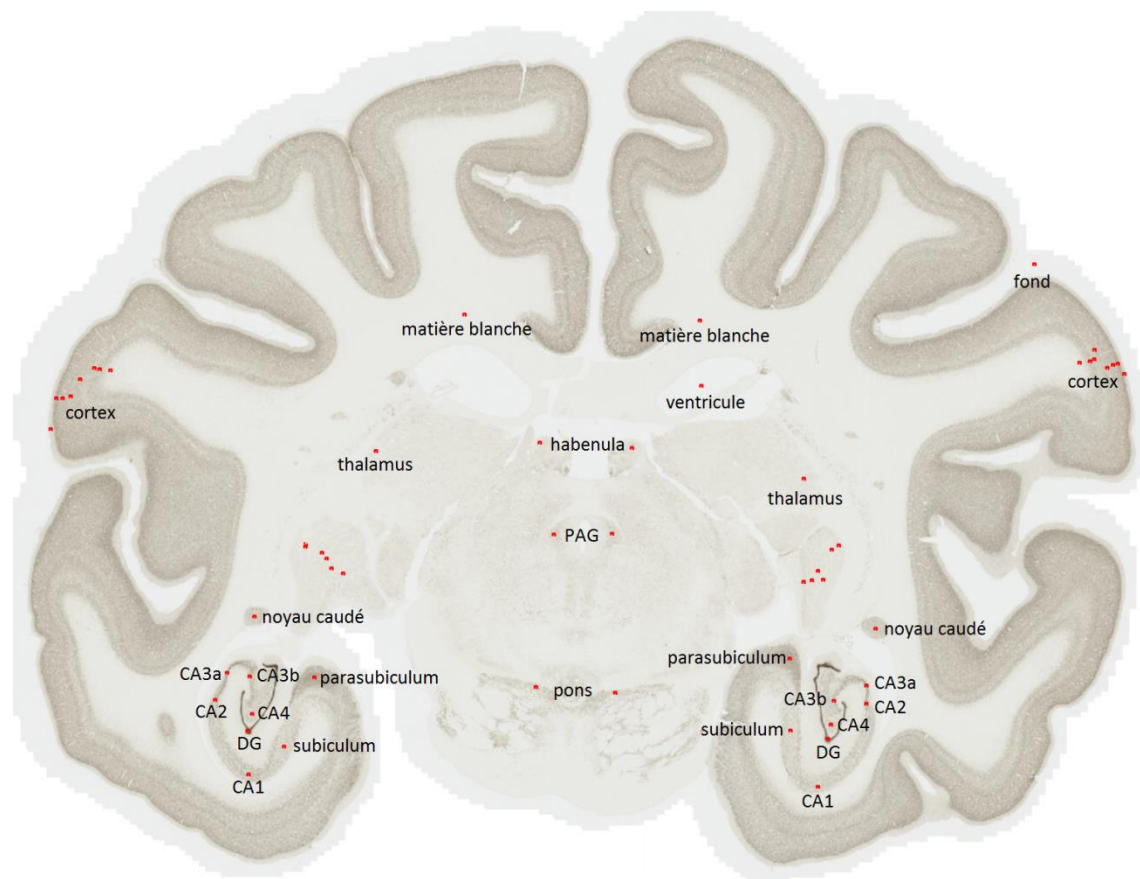


Figure 46 : Section 101 numérisée avec les emplacements des images extraites pour la base d'apprentissage et de test indiqués en rouge (tiré de la thèse de Zhenzhen You, 2017).

Annexe 2 : Tableaux des familles triées par les valeurs médianes et moyennes de positions

Dans les tableaux suivants, une nomenclature a été utilisée pour désigner les familles. La lettre "G" désigne une famille de Gabor, le chiffre qui suit sa longueur d'onde lambda et la lettre finale la partie du filtre ("R" pour la partie réelle et "I" pour la partie imaginaire). Le sigle "MBL" désigne les Motifs Binaires Locaux et le chiffre qui suit son rayon. Enfin, la lettre "H" désigne les familles d'Haralick. Le sigle qui suit désigne la caractéristique statistique "A" pour le second moment angulaire, "V" pour la variance, "Con" pour le contraste et "Corr" pour la corrélation. Le chiffre final est le rayon de son élément structurant.

La dispersion d'une famille de caractéristiques représente la différence entre la position de sa première caractéristique et de sa dernière caractéristique en termes de valeur médiane ou moyenne de positions.

Famille de caractéristiques en validation directe (VD)	Valeur médiane de positions moyenne de la famille en VD	Dispersion de la famille en VD	Famille de caractéristiques en validation croisée (VC)	Valeur médiane de positions moyenne de la famille en VC	Dispersion de la famille en VC
G 20 R	414,375	3	G 20 R	381,25	3
G 10 R	552,375	3	G 10 R	560,375	3
CIE XYZ	992,3333333	4	CIE XYZ	1041,3333333	4
RVB	1371	7	RVB	1407,166667	8
Image moyenne	1760,5	0	Image moyenne	1655,5	0
HSV	2166,833333	7	HSV	2180,666667	6
G 5 R	2328,875	18	G 5 R	2383,375	19
H Cor 11	2625,125	4	H Cor 11	2610,125	3
MBL 11	3030	0	MBL 40	3068	0
MBL 40	3178	0	MBL 11	3074	0
CIE L*a*b*	3229,666667	58	MBL 68	3204,5	0
H Con 11	3468,25	7	CIE L*a*b*	3299,666667	57
MBL 68	3558,5	0	H Con 11	3514,5	6
G 2,5 I	3692,75	89	G 2,5 R	3729,625	40
G 2,5 R	3699,875	39	G 2,5 I	3781,5	91
MBL 134	3868,5	0	H V 11	3917,5	4
H A 11	3978,625	4	Image de variance	3973	0
G 20 I	3991,125	91	G 20 I	3991,875	89
H V 11	4024,125	4	MBL 134	4048	0
Image de variance	4123,5	0	H A 11	4089,875	3

H Cor 40	4173,625	3	G 10 I	4332	90
G 10 I	4361,375	92	H V 40	4434,625	3
G 5 I	4419,75	78	G 5 I	4445,125	76
H V 40	4514,5	4	H Cor 40	4489,5	5
H Con 40	4683,25	46	H Con 40	4630,875	17
H Cor 68	4686,125	3	H V 68	4800,125	3
H V 68	4755,375	3	H V 134	4833,75	5
H V 134	4878,5	6	H Con 68	4908,375	18
H Cor 134	4952,625	11	H Cor 68	4914,75	7
H Con 134	5013,375	25	H A 40	4932	12
H A 40	5042,5	11	H Con 134	4939,75	18
H A 134	5064,625	10	H Cor 134	5020,125	15
H A 68	5066,875	6	H A 68	5240,125	5
H Con 68	5079,25	28	H A 134	5301	8

Tableau 18 : Classement des familles avec la valeur médiane de position pour le marqueur NeuN et dans les deux conditions de validation.

Famille de caractéristiques en validation directe (VD)	Valeur moyenne de positions moyenne de la famille en VD	Dispersion de la famille en VD	Famille de caractéristiques en validation croisée (VC)	Valeur moyenne de positions moyenne de la famille en VC	Dispersion de la famille en VC
G 20 R	532,3026316	3	G 20 R	493,8333333	3
G 10 R	623,3859649	3	G 10 R	592,6030702	3
CIE XYZ	1020,111111	4	CIE XYZ	1008,119883	5
RVB	1350,25731	7	RVB	1325,239766	8
Image moyenne	1644,763158	0	Image moyenne	1495,947368	0
HSV	2045,02924	6	HSV	2008,087719	6
G 5 R	2221,984649	21	G 5 R	2230,97807	23
H Cor 11	2451,390351	3	H Cor 11	2438,030702	3
MBL 11	2703,149123	0	MBL 11	2743,464912	0
MBL 40	2867,982456	0	MBL 40	2812,017544	0
H Con 11	2964,945175	9	MBL 68	2859,359649	0
MBL 68	3001,77193	0	H Con 11	3020,10307	8
CIE L*a*b*	3114,140351	76	CIE L*a*b*	3169,263158	81
G 2,5 I	3360,385965	91	G 2,5 I	3386,760965	91
H Cor 40	3409,063596	3	H V 11	3418,212719	5

G 2,5 R	3472,725877	67	G 10 I	3454,567982	91
H A 11	3516,79386	4	MBL 134	3457,964912	0
MBL 134	3518,482456	0	G 2,5 R	3479,616228	76
G 20 I	3604,754386	91	H A 11	3497,666667	4
H Cor 68	3691,357456	3	G 20 I	3625,006579	91
H V 11	3716,719298	3	H Cor 40	3635,675439	4
G 10 I	3770,317982	89	Image de variance	3636,421053	0
H Cor 134	3851,690789	3	H Con 40	3841,633772	8
Image de variance	3868,061404	0	H Cor 68	3910,254386	3
H Con 40	3868,929825	21	G 5 I	3982,881579	76
G 5 I	3959,372807	76	H Con 68	4041,217105	23
H Con 68	4077,333333	14	H Cor 134	4097,524123	5
H V 40	4129,074561	3	H Con 134	4099,539474	16
H Con 134	4140,785088	16	H V 40	4117,767544	5
H V 68	4276,254386	3	H A 40	4188,938596	8
H A 40	4299,609649	5	H V 68	4318,721491	3
H V 134	4350,247807	3	H V 134	4345,969298	3
H A 68	4393,449561	3	H A 68	4380,907895	5
H A 134	4446,953947	5	H A 134	4402,368421	12

Tableau 19 : Classement des familles avec la valeur moyenne de position pour le marqueur NeuN et dans les deux conditions de validation.

Famille de caractéristiques en validation directe (VD)	Valeur médiane de positions moyenne de la famille en VD	Dispersion de la famille en VD	Famille de caractéristiques en validation croisée (VC)	Valeur médiane de positions moyenne de la famille en VC	Dispersion de la famille en VC
MBL 40	586	0	MBL 40	210	0
MBL 20	1078	0	MBL 20	590	0
H Cor 5	1380,5	4	MBL 10	1478	0
G 20 R	1382	8	G 20 R	1602,75	3
G 10 R	1451,5	4	G 10 R	1740	7
HSV	1702,666667	27	RVB	1784	20
MBL 10	1799	0	H Cor 5	1784,75	3
RVB	1844	27	HSV	1825,333333	24
CIE XYZ	1886,666667	19	CIE XYZ	2153	20

H Cor 10	2615,75	4	CIE L*a*b*	2742,333333	5
Image moyenne	2811	0	Image moyenne	3248	0
H V 10	3117,25	15	Image de variance	3321	0
H V 5	3205	4	H Con 40	3945,75	15
CIE L*a*b*	3328,333333	36	H A 20	4020,5	24
Image de variance	3329	0	G 5 I	4172,5	38
H A 20	3507,5	22	G 5 R	4175	45
H Con 5	3657,75	31	G 2.5 I	4256,5	69
H Con 40	3915,25	30	G 2.5 R	4338,25	71
G 20 I	4113,25	65	H A 10	4369,25	29
G 5 R	4241,5	73	H V 10	4476	13
G 10 I	4275,5	61	G 10 I	4479,75	55
H Con 10	4299,25	69	H V 5	4482	9
H A 40	4307,25	37	G 20 I	4488,75	48
H Con 20	4482,25	41	H Con 20	4620,5	26
H A 10	4524,5	25	H Con 5	4632,5	42
H Cor 40	4598,75	12	H Cor 10	4838	10
H V 40	4676	17	MBL 5	4883	0
G 5 I	4815,25	35	H Con 10	4885,5	27
G 2.5 I	4864	50	H Cor 20	4993,5	7
H V 20	4878,25	11	H Cor 40	5066	5
G 2.5 R	4943,75	46	H V 20	5211,5	5
H Cor 20	5049	6	H A 5	5212	28
H A 5	5052,25	39	H V 40	5338,75	5
MBL 5	5236	0	H A 40	5354,75	4

Tableau 20 : Classement des familles avec la valeur médiane de position pour le marqueur anti-Phox2B et dans les deux conditions de validation.

Famille de caractéristiques en validation directe (VD)	Valeur moyenne de positions moyenne de la famille en VD	Dispersion de la famille en VD	Famille de caractéristiques en validation croisée (VC)	Valeur moyenne de positions moyenne de la famille en VC	Dispersion de la famille en VC
MBL 40	772,3076923	0	MBL 40	605,5726496	0
H Cor 5	1655,901709	4	MBL 20	1414,666667	0
MBL 20	1661,803419	0	RVB	1688,834758	17

G 20 R	1720,83547	4	HSV	1711,712251	24
HSV	1853,276353	27	G 10 R	1835,200855	11
G 10 R	1874,882479	7	G 20 R	1857,581197	6
RVB	1952,102564	23	CIE XYZ	1917,675214	18
CIE XYZ	2130,789174	19	H Cor 5	2467,925214	6
Image moyenne	2645,358974	0	Image moyenne	2521,290598	0
MBL 10	2696,752137	0	CIE L*a*b*	2582,512821	14
H Cor 10	2772,831197	4	MBL 10	2610,179487	0
Image de variance	2997,606838	0	Image de variance	2684,931624	0
CIE L*a*b*	3161,025641	24	G 5 R	3701,566239	69
H V 10	3516,818376	5	H A 20	3711,831197	5
H A 20	3732,858974	19	H Con 40	3786,606838	26
H V 5	3735,615385	8	G 20 I	3790,574786	53
H Cor 40	3780,25641	10	G 5 I	3807,698718	48
H Con 40	3868,373932	24	G 2.5 I	3813,482906	75
H Con 5	3869,651709	55	G 10 I	3839,982906	51
G 20 I	3903,726496	75	G 2.5 R	3892,557692	81
H Con 20	3966,525641	37	H Cor 40	3905,574786	14
G 5 R	3974,512821	77	H Con 20	3950,382479	39
G 10 I	4034,373932	73	H V 10	4023,380342	32
H A 40	4065,638889	25	H A 10	4071,318376	40
H Con 10	4075,418803	58	H Cor 20	4073,32906	7
H V 40	4119,230769	10	H Cor 10	4081,83547	20
H V 20	4184,489316	22	H Con 5	4124,24359	32
H A 10	4230,119658	27	H V 5	4204,418803	18
H Cor 20	4243,179487	12	H Con 10	4254,166667	32
MBL 5	4299,393162	0	H V 40	4278,90812	13
G 2.5 I	4404,818376	50	MBL 5	4287	0
G 5 I	4422,425214	40	H V 20	4389,549145	20
G 2.5 R	4514,358974	37	H A 40	4436,542735	5
H A 5	4579,196581	11	H A 5	4569,82265	5

Tableau 21 : Classement des familles avec la valeur moyenne de position pour le marqueur anti-Phox2B et dans les deux conditions de validation.

Annexe 3 : Publications

Conférence avec acte :

"Computational Optimization for Fast and Robust Automatic Segmentation in Virtual Microscopy using Brute-Force-based Feature Selection", **C. Bouvier** ; C. Clouchoux ; N. Souedet ; AS Herard ; Z. You ; C. Jan ; P. Hantraye ; G. Mergoil ; BH Rodriguez ; T. Delzescaux, ICPRAI 2018 Proceedings, p 407-413 (accepté)

Article de journaux :

"Study of textural and colorimetric feature space for histological images: selection, stability and efficiency", **C. Bouvier**, N. Souedet, J. Levy, C. Jan, Z. You, A-S Herard, G. Mergoil, BH Rodriguez, C. Clouchoux, T. Delzescaux (en préparation)

"Automated individualization of size-varying neurons in 2D microscopic images of macaque brain", Z. You ; Y. Balbastre ; **C. Bouvier** ; N. Souedet ; P. Gipchtein ; P. Hantraye ; C. Jan ; AS Herard ; T. Delzescaux (soumis à IEEE Transactions on Biomedical Engineering)

"Characterization of the Macaque retrotrapezoid nucleus (RTN), a critical structure in the neuronal network that regulates rhythmic respiration", J. Levy ; P. Facchinetti ; C. Jan ; M. Achour ; **C. Bouvier** ; JF Brunet ; T. Delzescaux ; F. Giuliano (en préparation)

Remarque : les trois articles devraient être soumis avant la fin de l'année 2018

Références bibliographiques

Abe, Naoki, et al. "Cross channel optimized marketing by reinforcement learning." *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2004.

Ale, Angelique, et al. "FMT-XCT: in vivo animal studies with hybrid fluorescence molecular tomography–X-ray computed tomography." *Nature methods* 9.6 (2012): 615.

Amiel, Jeanne, et al. "Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central hypoventilation syndrome." *Nature genetics* 33.4 (2003): 459.

Andersen, Per, et al., eds. *The hippocampus book*. Oxford university press, 2007.

Annese, J., et al. "Postmortem high-resolution 3-dimensional imaging of the primate brain: blockface imaging of perfusion stained tissue." *Neuroimage* 30.1 (2006): 61-69.

Atlamazoglou, V., et al. "Texture analysis of fluorescence microscopic images of colonic tissue sections." *Medical and Biological Engineering and Computing* 39.2 (2001): 145-151.

Avenel, Christophe, Pierre Fortin, and Dominique Béréziat. "Parallel birth and death process for cell nuclei extraction in histopathology images." *Parallel Processing (ICPP), 2013 42nd International Conference on*. IEEE, 2013.

Awada, Wael et al. "A Review of the Stability of Feature Selection Techniques for Bioinformatics Data." *Proceedings of the 2012 IEEE 13th International Conference on Information Reuse and Integration, IRI 2012* (2012): 356–363. Web.

Azevedo, Frederico AC, et al. "Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain." *Journal of Comparative Neurology* 513.5 (2009): 532-541.

Bacus, S. S., et al. "Biological grading of breast cancer using antibodies to proliferating cells and other markers." *The American journal of pathology* 135.5 (1989): 783.

Bagnell, J. Andrew, and Jeff G. Schneider. "Autonomous helicopter control using reinforcement learning policy search methods." *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164)*. Vol. 2. IEEE, 2001.

Balbastre, Y., et al. "Parallel computing in image analysis using BrainVisa software: application to histopathological staining segmentation in whole slide images." *European Congress on Digital Pathology*. 2014.

Balbastre, Yaël. Développement et validation d'outils pour l'analyse morphologique du cerveau de Macaque: Morphometry Analysis Tools for the Macaque Brain; Thèse. Université de Paris-Saclay, 2017.

Barat, Cécile, and Christophe Ducottet. "String representations and distances in deep Convolutional Neural Networks for image classification." *Pattern recognition* 54 (2016): 104-115.

Barker, Jocelyn et al. "Automated Classification of Brain Tumor Type in Whole-Slide Digital Pathology Images Using Local Representative Tiles." *Medical Image Analysis* 30 (2016): 60–71. Web.

Basavanhally, Ajay et al. "Multi-Field-of-View Framework for Distinguishing Tumor Grade in ER+ Breast Cancer from Entire Histopathology Slides." *IEEE Transactions on Biomedical Engineering* 60.8 (2013): 2089–2099. Web.

Bell, Gordon, Tony Hey, and Alex Szalay. "Computer Science. Beyond the Data Deluge." *Science* (New York, N.Y.) 323.5919 (2009): 1297–1298. Web.

Bellman, Richard. "A Markovian decision process." *Journal of Mathematics and Mechanics* (1957): 679-684.

Bellman, Richard. *Dynamic programming*. Courier Corporation, 2013.

Berrios, German E. "Alzheimer's disease: a conceptual history." *International Journal of Geriatric Psychiatry* 5.6 (1990): 355-365.

Boser, Bernhard E., Isabelle M. Guyon, and Vladimir N. Vapnik. "A training algorithm for optimal margin classifiers." *Proceedings of the fifth annual workshop on Computational learning theory*. ACM, 1992.

Bouvier, Clément et al. "Computational Optimization for Fast and Robust Automatic Segmentation in Virtual Microscopy using Brute-Force-based Feature Selection." *ICPRAI 2018 Proceedings*, p 407-413. (2018)

Braak, Heiko, and Eva Braak. "Neuropathological staging of Alzheimer-related changes." *Acta neuropathologica* 82.4 (1991): 239-259.

Braun, Allen R., et al. "Regional cerebral blood flow throughout the sleep-wake cycle. An H2 (15) O PET study." *Brain: a journal of neurology* 120.7 (1997): 1173-1197.

Breiman, Leo. "Bagging predictors." *Machine learning* 24.2 (1996): 123-140.

Breiman, L. "Random Forests." *Machine learning* 45.1 (2001): 5–32. Web.

Brureau, Anthony et al. "NF-L in Cerebrospinal Fluid and Serum Is a Biomarker of Neuronal Damage in an Inducible Mouse Model of Neurodegeneration." *Neurobiology of Disease* 104 (2017): 73–84. Web.

Chandrashekar, Girish, and Ferat Sahin. "A Survey on Feature Selection Methods." *Computers and Electrical Engineering* 40.1 (2014): 16–28. Web.

Chen, Chao, Andy Liaw, and Leo Breiman. "Using Random Forest to Learn Imbalanced Data." *University of California, Berkeley* 1999 (2004): 1–12. Web.

Cheng, H. D. et al. "Color Image Segmentation: Advances and Prospects." *Pattern Recognition* 34.12 (2001): 2259–2281. Web.

Cichocki, Andrzej. "Era of Big Data Processing: A New Approach via Tensor Networks and Tensor Decompositions." *arXiv* (2014): 1–30. Web.

Chubb, Charles, et al. "BioVision: an application for the automated image analysis of histological sections." *Neurobiology of aging* 27.10 (2006): 1462-1476.

Chung, Kwanghun et al. "Structural and Molecular Interrogation of Intact Biological Systems." *Nature* 497.7449 (2013): 332–337. Web.

Ciresan, D C et al. "Mitosis Detection in Breast Cancer Histology Images Using Deep Neural Networks." *Proc Medical Image Computing Computer Assisted Intervention (MICCAI)* (2013): 411–418. Print.

Cruz-Roa, Angel, Juan C. Caicedo, and Fabio a. Gonzalez. "Visual Pattern Mining in Histology Image Collections Using Bag of Features." *Artificial Intelligence in Medicine* 52.2 (2011): 91–106. Web.

Cruz-Roa, Angel, et al. "High-throughput adaptive sampling for whole-slide histopathology image analysis (HASHI) via convolutional neural networks: Application to invasive breast cancer detection." *PloS one* 13.5 (2018): e0196828.

Damadian, Raymond. "Tumor detection by nuclear magnetic resonance." *Science* 171.3976 (1971): 1151-1153.

Darwin, Charles. *On the origin of species*, 1859. Routledge, 2004.

Dash, M., and H. Liu. "Feature Selection for Classification." *Intelligent Data Analysis* 1.3 (1997): 131–156. Web.

Daugman, John G. "Image analysis and compact coding by oriented 2D Gabor primitives." *Image Understanding and the Man-Machine Interface*. Vol. 758. International Society for Optics and Photonics, 1987.

Dauguet, Julien et al. "Three-Dimensional Reconstruction of Stained Histological Slices and 3D Non-Linear Registration with in-Vivo MRI for Whole Baboon Brain." *Journal of Neuroscience Methods* 164.1 (2007): 191–204. Web.

Dempster, Arthur P., Nan M. Laird, and Donald B. Rubin. "Maximum likelihood from incomplete data via the EM algorithm." *Journal of the royal statistical society. Series B (methodological)* (1977): 1-38.

Denoix, P. F. "TNM classification." *Bull Inst Nat Hyg* 1.69 (1944): 5282.

Department of Biochemistry and Molecular Biophysics Thomas Jessell, Steven Siegelbaum, and A. J. Hudspeth. *Principles of neural science*. Eds. Eric R. Kandel, James H. Schwartz, and Thomas M. Jessell. Vol. 4. New York: McGraw-hill, 2000.

DiFranco, Matthew D. et al. "Ensemble Based System for Whole-Slide Prostate Cancer Probability Mapping Using Color Texture Features." *Computerized Medical Imaging and Graphics* 35.7–8 (2011): 629–645. Web.

Ding, Yuchun et al. "Novel Methods for Microglia Segmentation, Feature Extraction and Classification." *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 5963.c (2016): 1–1. Web.

Doyle, Scott et al. "Automated Grading of Breast Cancer Histopathology Using Spectral Clustering with Textural and Architectural Image Features." 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Proceedings, ISBI (2008): 496–499. Web.

Dubois, a et al. "Detection by Voxel-Wise Statistical Analysis of Significant Changes in Regional Cerebral Glucose Uptake in an APP/PS1 Transgenic Mouse Model of Alzheimer's Disease." *Neuroimage* 51.2 (2010): 586–598. Web.

Ericson, Marica B., et al. "Bispectral fluorescence imaging combined with texture analysis and linear discrimination for correlation with histopathologic extent of basal cell carcinoma." *Journal of biomedical optics* 10.3 (2005): 034009.

Fatima, Kiran, Hammad Majeed, and Humayun Irshad. "Nuclear Spatial and Spectral Features Based Evolutionary Method for Meningioma Subtypes Classification in Histopathology." *Microscopy Research and Technique* February (2017): 1–11. Web.

Feki, Abdelmonem, et al. "Fully automated and adaptive detection of amyloid plaques in stained brain sections of Alzheimer transgenic mice." *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Berlin, Heidelberg, 2007.

Fogel, I., and D. Sagi. "Gabor Filters as Texture Discriminator." *Biological Cybernetics* 61.2 (1989): 103–113. Web.

Förstl, H, and A Kurz. "Clinical Features of Alzheimer's Disease." *European archives of psychiatry and clinical neuroscience* 249.6 (1999): 288–290. Web.

Gabor, Dennis. "Theory of communication. Part 1: The analysis of information." *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering* 93.26 (1946): 429-441.

Gakuba, Clement, et al. "General Anesthesia Inhibits the Activity of the "Glymphatic System"." *Theranostics* 8.3 (2018): 710.

Gartner, Leslie P. *Textbook of Histology E-Book*. Elsevier Health Sciences, 2015.

Gaultier, Claude, et al. "Pediatric disorders with autonomic dysfunction: what role for PHOX2B?." *Pediatric research* 58.1 (2005): 1.

Geman, Stuart, and Donald Geman. "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images." *IEEE Transactions on pattern analysis and machine intelligence* 6 (1984): 721-741.

Gheyas, Iffat a., and Leslie S. Smith. "Feature Subset Selection in Large Dimensionality Domains." *Pattern Recognition* 43.1 (2010): 5–13. Web.

Gospodarowicz, Mary K., James D. Brierley, and Christian Wittekind, eds. *TNM classification of malignant tumours*. John Wiley & Sons, 2017.

Goya-Outi, Jessica, et al. "Computation of reliable textural indices from multimodal brain MRI: suggestions based on a study of patients with diffuse intrinsic pontine glioma." *Physics in Medicine & Biology* 63.10 (2018): 105003.

Grassmann, Hermann. "Zur theorie der farbenmischung." *Annalen der Physik* 165.5 (1853): 69-84.

Guild, John. "The Colorimetric Properties of the Spectrum." The Royal Society (1931): n. pag. Print.

Gurcan, Metin N., et al. "Histopathological image analysis: A review." *IEEE reviews in biomedical engineering* 2 (2009): 147.

Guyon, Isabelle, and André Elisseeff. "An Introduction to Variable and Feature Selection." *Journal of Machine Learning Research (JMLR)* 3.3 (2003): 1157–1182. Web.

Hamon, Julie. *Optimisation combinatoire pour la sélection de variables en régression en grande dimension: Application en génétique animale*. Diss. Université des Sciences et Technologie de Lille-Lille I, 2013.

Haralick, Robert, K. Shanmugan, and I. Dinstein. "Textural Features for Image Classification." *IEEE Transactions on Systems, Man and Cybernetics* 1973: 610–621. Web.

Hay, Joel W., and Richard L. Ernst. "The economic costs of Alzheimer's disease." *American Journal of Public Health* 77.9 (1987): 1169-1175.

Hernandez, F., and J. Avila. "Tauopathies." *Cellular and Molecular Life Sciences* 64.17 (2007): 2219-2233.

Heule, Marijn J. H., and Oliver Kullmann. "The Science of Brute Force." *Communications of the ACM* 60.8 (2017): 70–79. Web.

Ho, Tin Kam, and Murray Hill. "Random Decision Forests." (1995): 278–. Web.

Jaber, M. "The cerebellum as a major player in motor disturbances related to Autistic Syndrome Disorders." *L'Encephale* 43.2 (2017): 170-175.

Jack Jr, Clifford R., et al. "In vivo visualization of Alzheimer's amyloid plaques by magnetic resonance imaging in transgenic mice without a contrast agent." *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* 52.6 (2004): 1263-1271.

- Jacob, François. *La logique du vivant: une histoire de l'hérédité*. Gallimard, 1987.
- Jéhannin-Ligier, K., et al. "Projection de l'incidence et de la mortalité par cancer en France métropolitaine en 2017." *Rapport technique. Saint-Maurice: Santé publique France* (2017): 1-80.
- Joblove, George H., and Donald Greenberg. "Color spaces for computer graphics." *ACM siggraph computer graphics*. Vol. 12. No. 3. ACM, 1978.
- Jucker, Mathias. "The Benefits and Limitations of Animal Models for Translational Research in Neurodegenerative Diseases." *Nature Medicine* 16.11 (2010): 1210–1214. Web.
- Kallenbach-Thieltges, Angela et al. "Immunohistochemistry, Histopathology and Infrared Spectral Histopathology of Colon Cancer Tissue Sections." *Journal of Biophotonics* 6.1 (2013): 88–100. Web.
- Kalousis, a., J. Prados, and M. Hilario. "Stability of Feature Selection Algorithms." Fifth IEEE International Conference on Data Mining (ICDM'05) (2005): 218–225. Web.
- Kather, Jakob Nikolas et al. "Multi-Class Texture Analysis in Colorectal Cancer Histology." *Scientific Reports* 6 (2016): 27988. Web.
- Kermany, Daniel S et al. "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning Resource Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning." *Cell* 172.5 (2018): 1122–1124.e9. Web.
- Kim, Kee K., Robert S. Adelstein, and Sachiyo Kawamoto. "Identification of neuronal nuclei (NeuN) as Fox-3, a new member of the Fox-1 gene family of splicing factors." *Journal of Biological Chemistry* (2009): jbc-M109.
- Knight, Michael J., et al. "Quantitative T 1 and T 2 MRI signal characteristics in the human brain: different patterns of MR contrasts in normal ageing." *Magnetic Resonance Materials in Physics, Biology and Medicine* 29.6 (2016): 833-842.
- Kooi, Thijs et al. "Large Scale Deep Learning for Computer Aided Detection of Mammographic Lesions." *Medical Image Analysis* 35 (2016): 303–312. Web.
- Kudo, Mineichi, and Jack Sklansky. "Comparison of Algorithms That Select Features for Pattern Classifiers." *Pattern Recognition* 33.1 (2000): 25–41. Web.
- Kuncheva, Ludmila I., and Juan J. Rodríguez. "On Feature Selection Protocols for Very Low-Sample-Size Data." *Pattern Recognition* (2018): n. pag. Web.
- Laguitton, S et al. "Soma-Workflow: A Unified and Simple Interface to Parallel Computing Resources." *MICCAI Workshop on High Performance and Distributed Computing for Medical Imaging* (2011): n. pag. Print.
- Laurén, Juha, et al. "Cellular prion protein mediates impairment of synaptic plasticity by amyloid- β oligomers." *Nature* 457.7233 (2009): 1128.

Lazar, Cosmin, et al. "A survey on filter techniques for feature selection in gene expression microarray analysis." *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)* 9.4 (2012): 1106-1119.

Le Bihan, Denis, et al. "Diffusion tensor imaging: concepts and applications." *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* 13.4 (2001): 534-546.

LeCun, Yann, et al. "A theoretical framework for back-propagation." *Proceedings of the 1988 connectionist models summer school*. Vol. 1. CMU, Pittsburgh, Pa: Morgan Kaufmann, 1988.

Lecun Y., Bengio Y., and Hinton G. "Deep Learning." *Nature* 521.7553 (2015): 436–444. Web.

Leenhardt, Romain, et al. "A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy." *Gastrointestinal endoscopy* (2018).

Li, Xingyu, and Konstantinos N. Plataniotis. "A Complete Color Normalization Approach to Histopathology Images Using Color Cues Computed From Saturation-Weighted Statistics." *IEEE Transactions on Biomedical Engineering* 62.7 (2015): 1862–1873. Web.

Liu, Haoyue. "Decision Tree Rule-Based Feature Selection for Large-Scale Imbalanced Data." *IEEE* (2017): 1–6. Print.

Llewellyn-Smith, Ida J., and Anthony JM Verberne, eds. *Central regulation of autonomic functions*. Oxford University Press, 2011.

Lucasius, Carlos B., and Gerrit Kateman. "Understanding and using genetic algorithms Part 1. Concepts, properties and context." *Chemometrics and intelligent laboratory systems* 19.1 (1993): 1-33.

Malandain, Grégoire et al. "Fusion of Autoradiographs with an MR Volume Using 2-D and 3-D Linear Transformations." *NeuroImage* 23.1 (2004): 111–127. Web.

Marill, Thomas, and D. Green. "On the effectiveness of receptors in recognition systems." *IEEE transactions on Information Theory* 9.1 (1963): 11-17.

Mehrotra, H, Kr Namuduri, and N Ranganathan. "Gabor Filter-Based Edge Detection." *Pattern Recognition* 25.12 (1992): 1479–1494. Print.

Meinshausen, Nicolai, and Peter Bühlmann. "Stability Selection." *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 72.4 (2010): 417–473. Web.

Membres du « Groupe du dictionnaire » de l'ANM : C. Buffet, V. Delmas, J. Frottier, C-P. Giudicelli(Vice-président), D. Hauteville, J. Hureau (Président), F-X. Maquart, J-L. Michaux, G. Morvan, F. Rodhain, J. Rouëssé, R. Veitia, J-L. Wemeau.
Consultants réguliers : R. Ardaillou, J-L. Arné, B. Dréno, B ; Falissard, C. Géraut, JL. Haw, J-P. Michel, J. Milliez, J-P. Olier, Y. Touitou, R. Trèves. "Dictionnaire de l'académie de médecine." lien: <http://dictionnaire.academie-medecine.fr/>. (2018)

Moore, Andrew W. "Very fast EM-based mixture model clustering using multiresolution kd-trees." *Advances in Neural information processing systems*. 1999.

Muthu Rama Krishnan, M. et al. "Textural Characterization of Histopathological Images for Oral Sub-Mucous Fibrosis Detection." *Tissue and Cell* 43.5 (2011): 318–330. Web.

Naik, Shivang et al. "Automated Gland and Nuclei Segmentation for Grading of Prostate and Breast Cancer Histopathology." 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Proceedings, ISBI c (2008): 284–287. Web.

Newton, Isaac. "Opticks: or, a treatise of the reflexions, refractions, inflexions and colours oflight, Also two treatises of the species and magnitude ofcurvilinearfigures." *London* 1704: 1887-1925.

Nguyen Quang, Nam, et al. "Time-lapse imaging of molecular evolution by high-throughput sequencing." *Nucleic acids research* 46.15 (2018): 7480-7494.

Ogawa, Seiji, et al. "Brain magnetic resonance imaging with contrast dependent on blood oxygenation." *Proceedings of the National Academy of Sciences* 87.24 (1990): 9868-9872.

Ojala, T., M. Pietikainen, and D. Harwood. "Performance Evaluation of Texture Measures with Classification Based on Kullback Discrimination of Distributions." *Proceedings of the 12th IAPR International Conference on Pattern Recognition (ICPR)* 1 (1994): 582–585. Web.

Ojala, Timo, Matti Pietikainen, and Topi Maenpaa. "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24.7 (2002): 971–987. Web.

Otsu, Nobuyuki. "Otsu1975." 20.1 (1979): 62–66. Print.

Pääkkönen, Pekka, and Daniel Pakkala. "Reference Architecture and Classification of Technologies, Products and Services for Big Data Systems." *Big Data Research* 2.4 (2015): 166–186. Web.

Papernot, Nicolas et al. "The Limitations of Deep Learning in Adversarial Settings." *IEEE* (2015): n. pag. Web.

Pascanu, Razvan, Tomas Mikolov, and Yoshua Bengio. "On the Difficulty of Training Recurrent Neural Networks." 1994 (2012): n. pag. Web.

Pavlov, Ivan Petrovich. "Conditional reflexes: an investigation of the physiological activity of the cerebral cortex." (1927).

Paxinos, George, Xu-Feng Huang, and Arthur W. Toga. "The rhesus monkey brain in stereotaxic coordinates." (2000).

Pedregosa, Fabian et al. "Scikit-Learn: Machine Learning in Python." *Journal of Machine Learning Research* 12 (2012): 2825–2830. Web.

Peng, Hanchuan, Fuhui Long, and Chris Ding. "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy." *IEEE Transactions on pattern analysis and machine intelligence* 27.8 (2005): 1226-1238.

Peng, Yonghong, Zhiqing Wu, and Jianmin Jiang. "A Novel Feature Selection Approach for Biomedical Data Classification." *Journal of Biomedical Informatics* 43.1 (2010): 15–23. Web.

Powers, David Martin. "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." (2011).

Priller, Christina, et al. "Synapse formation and function is modulated by the amyloid precursor protein." *Journal of Neuroscience* 26.27 (2006): 7212-7221.

Pudil, Pavel, Jana Novovičová, and Josef Kittler. "Floating search methods in feature selection." *Pattern recognition letters* 15.11 (1994): 1119-1125.

Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks." arXiv (2015): 1–15. Web.

Rakic, Pasko. "Specification of cerebral cortical areas." *Science* 241.4862 (1988): 170-176.

Renier, Nicolas, et al. "iDISCO: a simple, rapid method to immunolabel large tissue samples for volume imaging." *Cell* 159.4 (2014): 896-910.

REPORT FROM THE COMMISSION TO THE COUNCIL AND THE EUROPEAN PARLIAMENT Sixth Report on the Statistics on the Number of Animals used for Experimental and other Scientific Purposes in the Member States of the European Union SEC(2010) 1107

Rivière, D., et al. "BrainVISA: an extensible software environment for sharing multimodal neuroimaging data and processing tools." *Neuroimage* 47 (2009): S163.

Rivière, D., et al. "Anatomist: a python framework for interactive 3D visualization of neuroimaging data." *Anatomist: A Python Framework for Interactive 3D Visualization of Neuroimaging Data, (Editor (Ed)^(Eds) Book)* (2011).

Rosenblatt, Frank. *The perceptron, a perceiving and recognizing automaton Project Para*. Cornell Aeronautical Laboratory, 1957.

Ross, Michael H., and Wojciech Pawlina. *Histology*. Lippincott Williams & Wilkins, 2006.

Roullier, Vincent et al. "Multi-Resolution Graph-Based Analysis of Histopathological Whole Slide Images: Application to Mitotic Cell Extraction and Visualization." *Computerized Medical Imaging and Graphics* 35.7–8 (2011): 603–615. Web.

Rudy, Carolyn C., et al. "The role of the tripartite glutamatergic synapse in the pathophysiology of Alzheimer's disease." *Aging and disease* 6.2 (2015): 131.

Salakhutdinov, Ruslan, and Hugo Larochelle. "Efficient learning of deep Boltzmann machines." *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. 2010.

Schafer, Dorothy P., Emily K. Lehrman, and Beth Stevens. "The "quad-partite" synapse: Microglia-synapse interactions in the developing and mature CNS." *Glia* 61.1 (2013): 24-36.

Schmitt, Oliver et al. "Parallelized Segmentation of a Serially Sectioned Whole Human Brain." *Image and Vision Computing* 26.2 (2008): 289–301. Web.

Sertel, O. et al. "Computer-Aided Prognosis of Neuroblastoma on Whole-Slide Images: Classification of Stromal Development." *Pattern Recognition* 42.6 (2009): 1093–1103. Web.

Shah, Rajen, and Richard J Samworth. "Variable Selection with Error Control : Another Look at Stability Selection." *Royal Statistical Society* (2012): 1–33. Web.

Silver, David et al. "Mastering the Game of Go with Deep Neural Networks and Tree Search." *Nature* 529.7587 (2016): 484–489. Web.

Sima, Chao, and Edward R. Dougherty. "The peaking phenomenon in the presence of feature-selection." *Pattern Recognition Letters* 29.11 (2008): 1667-1674.

Squire, Larry, et al., eds. *Fundamental neuroscience*. Academic Press, 2012.

Sweet, William H. "The uses of nuclear disintegration in the diagnosis and treatment of brain tumor." *New England Journal of Medicine* 245.23 (1951): 875-878.

Taha, Abdel Aziz, and Allan Hanbury. "Metrics for Evaluating 3D Medical Image Segmentation: Analysis, Selection, and Tool." *BMC Medical Imaging* 15.1 (2015): 29. Web.

Tambasco, Mauro, and Anthony M. Magliocco. "Relationship between Tumor Grade and Computed Architectural Complexity in Breast Cancer Specimens." *Human Pathology* 39.5 (2008): 740–746. Web.

Tang, Jiliang, Salem Alelyani, and Huan Liu. "Feature Selection for Classification: A Review." *Data Classification: Algorithms and Applications* (2014): 37–64. Web.

Tesauro, Gerald. "Practical issues in temporal difference learning." *Advances in neural information processing systems*. 1992.

Thompson, Scott E., and Srivatsan Parthasarathy. "Moore's law: the future of Si microelectronics." *Materials today* 9.6 (2006): 20-25.

Tuerk, Craig, and Larry Gold. "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase." *Science* 249.4968 (1990): 505-510.

Turner, Paul R., et al. "Roles of amyloid precursor protein and its fragments in regulating neural activity, plasticity and memory." *Progress in neurobiology* 70.1 (2003): 1-32.

Vågberg, William et al. "Cellular-Resolution 3D Virtual Histology of Human Coronary Arteries Using x-Ray Phase Tomography." *Scientific Reports* October 2017 (2018): 1–7. Web.

Vandenberghe, Michel E. et al. "Robust Supervised Segmentation of Neuropathology Whole-Slide Microscopy Images." *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2015–Novem* (2015): 3851–3854. Web.

Vandenberghe, Michel E. et al. "High-Throughput 3D Whole-Brain Quantitative Histopathology in Rodents." *Scientific reports* January 2015 (2016): 1–12. Web.

Van Eycke, Yves-Rémi, et al. "Segmentation of glandular epithelium in colorectal tumours to automatically compartmentalise IHC biomarker quantification: A deep learning approach." *Medical image analysis* 49 (2018): 35-45.

Von Goethe, Johann Wolfgang. *Zur Farbenlehre*. Jazzybee Verlag, 2015.

West, M. J., L. H. J. G. Slomianka, and H. Jx G. Gundersen. "Unbiased stereological estimation of the total number of neurons in the subdivisions of the rat hippocampus using the optical fractionator." *The Anatomical Record* 231.4 (1991): 482-497.

Whitney, A. Wayne. "A direct method of nonparametric measurement selection." *IEEE Transactions on Computers* 100.9 (1971): 1100-1103.

Whone, Alan L., et al. "Slower progression of Parkinson's disease with ropinirole versus levodopa: the REAL-PET study." *Annals of neurology* 54.1 (2003): 93-101.

Witowski, Stanley R., and Cecil H. Brown. "An Explanation of NFC." *American Anthropologist* 1969 (1977): 1–9. Print.

Witkowski, Michal, et al. "Practical power consumption estimation for real life HPC applications." *Future Generation Computer Systems* 29.1 (2013): 208-217.

Witthaus, E et al. "Burden of Mortality and Morbidity from Dementia." *Alzheimer Disease and Associated Disorders* 13.3 (1999): 176–181. Web.

World Health Organization. "Fact Sheet: Dementia." link:<http://www.who.int/en/news-room/fact-sheets/detail/dementia>. (12 December 2017)

Wright, William David. "A re-determination of the trichromatic coefficients of the spectral colours." *Transactions of the Optical Society* 30.4 (1929): 141.

Wu, Chialin, David A. Landgrebe, and P. H. Swain. "The decision tree approach to classification." (1975).

Wu, Xindong et al. "Data Mining with Big Data." *IEEE Transactions on Knowledge and Data Engineering* 26.1 (2014): 97–107. Web.

Xie, Lulu, et al. "Sleep drives metabolite clearance from the adult brain." *science* 342.6156 (2013): 373-377.

Yankner, Bruce A., Lawrence K. Duffy, and Daniel A. Kirschner. "Neurotrophic and neurotoxic effects of amyloid beta protein: reversal by tachykinin neuropeptides." *Science* 250.4978 (1990): 279-282.

Yaqub, M. et al. "Investigation of the Role of Feature Selection and Weighted Voting in Random Forests for 3-D Volumetric Segmentation." *IEEE Transactions on Medical Imaging* 33.2 (2014): 258–271. Web.

Yin, Liuzhi et al. "Feature Selection for High-Dimensional Imbalanced Data." *Neurocomputing* 105 (2013): 3–11. Web.

You, Zhenzhen, et al. "Automated cell individualization and counting in cerebral microscopic images." *Image Processing (ICIP), 2016 IEEE International Conference on*. IEEE, 2016.

You, Zhenzhen. "Etude de la morphologie et de la distribution des neurones dans le cerveau de macaque par microscopie optique." Université Pierre et Marie Curie. 2017.

Young, Anne B., et al. "PET scan investigations of Huntington's disease: cerebral metabolic correlates of neurological features and functional decline." *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* 20.3 (1986): 296-303.

Zilles, Karl, et al. "Quantitative analysis of sulci in the human cerebral cortex: development, regional heterogeneity, gender difference, asymmetry, intersubject variability and cortical architecture." *Human brain mapping* 5.4 (1997): 218-221.

Zwitter, Andrej. "Big Data Ethics." *Big Data & Society* 1.2 (2014): 205395171455925. Web.

Résumé

En recherche préclinique et plus particulièrement en neurobiologie, l'histologie est utilisée et produit des images à l'échelle cellulaire grâce à des microscopes optiques toujours plus performants. La quantification du tissu marqué comme, par exemple, les neurones s'appuie de plus en plus sur des segmentations par apprentissage automatique. Cependant, l'apprentissage automatique nécessite une grande quantité d'informations intermédiaires, ou caractéristiques, extraites de la donnée brute multipliant d'autant la quantité de données à traiter. Ainsi, le nombre important de ces caractéristiques est un obstacle au traitement robuste et rapide de grandes séries d'images histologiques. Les algorithmes de sélection de caractéristiques pourraient réduire la quantité d'informations nécessaires mais les ensembles de caractéristiques sélectionnés sont peu reproductibles.

Nous proposons une méthodologie originale fonctionnant sur des infrastructures de calcul haute-performance (CHP) visant à sélectionner des petits ensembles de caractéristiques stables afin de permettre des segmentations rapides et robustes sur des images histologiques acquises à très haute-résolution. Cette sélection se déroule en deux étapes : la première à l'échelle des familles de caractéristiques (regroupement intermédiaire entre les espaces et les caractéristiques). La deuxième est appliquée directement sur les caractéristiques issues de ces familles. Dans ce travail, nous avons obtenu des ensembles généralisables et stables pour deux marquages neuronaux différents. Ces ensembles permettent des réductions significatives des temps de traitement et de la mémoire vive utilisée. Cette méthodologie rendra possible des études histologiques exhaustives à haute-résolution sur des infrastructures CHP que ce soit en recherche préclinique et possiblement clinique.

Abstract

In preclinical research and more specifically in neurobiology, histology uses images produced by increasingly powerful optical microscopes digitizing entire sections at cell scale. Quantification of stained tissue such as neurons relies on machine learning driven segmentation. However such methods need a lot of additional information, or features, which are extracted from raw data multiplying the quantity of data to process. As a result, the quantity of features is becoming a drawback to process large series of histological images in a fast and robust manner. Feature selection methods could reduce the amount of required information but selected subsets lack of stability.

We propose a novel methodology operating on high performance computing (HPC) infrastructures and aiming at finding small and stable sets of features for fast and robust segmentation on high-resolution histological whole sections. This selection has two selection steps: first at feature families scale (an intermediate pool of features, between space and individual feature). Second, feature selection is performed on pre-selected feature families. In this work, the selected sets of features are stables for two different neurons staining. Furthermore the feature selection results in a significant reduction of computation time and memory cost. This methodology can potentially enable exhaustive histological studies at a high-resolution scale on HPC infrastructures for both preclinical and clinical research settings.