



Etude de la morphologie et de la distribution des neurones dans le cerveau de macaque par microscopie optique

Zhenzhen You

► To cite this version:

Zhenzhen You. Etude de la morphologie et de la distribution des neurones dans le cerveau de macaque par microscopie optique. Bio-informatique [q-bio.QM]. Université Pierre et Marie Curie - Paris VI, 2017. Français. NNT : 2017PA066278 . tel-02311480

HAL Id: tel-02311480

<https://theses.hal.science/tel-02311480>

Submitted on 11 Oct 2019

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

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

Université Pierre et Marie Curie

École Doctorale

Informatique, Télécommunications et Électronique

**Etude de la morphologie et de la distribution des neurones
dans le cerveau de macaque par microscopie optique**

Présentée par

Mlle. Zhenzhen YOU

Pour obtenir le grade de

DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE

Dirigée par Thierry Delzescaux du directeur de thèse

Date de soutenance : 09/10/2017

Devant le jury composé de :

| | | |
|------|--------------------|--------------------|
| Mme. | Frédérique Frouin | Rapporteur |
| M. | Philippe Andrey | Rapporteur |
| M. | Dominique Béréziat | Examineur |
| M. | Eric Petit | Examineur |
| M. | Marc Dhenain | Examineur |
| M. | Thierry Delzescaux | Directeur de thèse |

Table des matières

| | |
|---|----|
| Glossaire | 1 |
| Introduction..... | 3 |
| 1. Contexte | 7 |
| 1.1. Le système nerveux central | 7 |
| 1.1.1. Introduction générale | 7 |
| 1.1.2. Les principales structures et fonctions du cerveau | 8 |
| 1.1.3. Composition cellulaire du cerveau | 10 |
| 1.1.3.1. Les neurones | 10 |
| 1.1.3.2. Les cellules gliales | 11 |
| 1.1.4. Etudes relatives aux neurones..... | 12 |
| 1.1.4.1. Etude du cerveau sain..... | 12 |
| 1.1.4.2. Cas des maladies neurodégénératives | 14 |
| 1.2. Modèles animaux en neurosciences..... | 17 |
| 1.2.1. La recherche translationnelle | 17 |
| 1.2.2. Modèles animaux..... | 17 |
| 1.2.3. Intérêt des modèles animaux | 19 |
| 1.3. Imagerie du cerveau <i>in vivo/post mortem</i> | 20 |
| 1.3.1. L'imagerie <i>in vivo</i> | 20 |
| 1.3.1.1. Imagerie par résonance magnétique | 20 |
| 1.3.1.2. Imagerie par tomographie à émission de positons | 21 |
| 1.3.2. L'imagerie <i>post mortem</i> | 22 |
| 1.3.2.1. Principe de l'histologie..... | 22 |
| 1.3.2.2. Production de coupes histologiques..... | 24 |
| 1.3.2.3. Imagerie des coupes histologiques..... | 27 |
| 1.3.2.4. Analyse des coupes histologiques | 29 |
| 1.3.3. Complémentarité des techniques d'imagerie du cerveau..... | 30 |
| 1.4. Problématique liée à l'étude des neurones en microscopie optique | 31 |
| 2. Méthode de segmentation des neurones | 35 |
| 2.1. Etat de l'art | 35 |
| 2.1.1. Méthodes de segmentation..... | 35 |
| 2.1.2. Validation des méthodes de segmentation d'objets..... | 39 |
| 2.2. Données biologiques de l'étude | 41 |
| 2.2.1. Production des coupes histologiques de cerveau de macaque..... | 41 |
| 2.2.2. Bases de données de l'étude | 41 |

| | | |
|----------|---|-----|
| 2.2.2.1. | Base d'images de coupes histologiques marquées par NeuN | 42 |
| 2.2.2.2. | Base d'images pour la segmentation du marquage NeuN | 42 |
| 2.2.2.3. | Base d'images pour l'individualisation des neurones..... | 45 |
| 2.2.2.4. | Base d'images pour le comptage stéréologique de référence des neurones | 46 |
| 2.2.2.5. | Résumé des bases d'images créées | 47 |
| 2.3. | Segmentation du marquage neuronal par apprentissage/classification..... | 47 |
| 2.3.1. | Principe de la méthode des forêts aléatoires (RF)..... | 48 |
| 2.3.2. | Segmentation de référence des neurones | 48 |
| 2.3.3. | Création du modèle d'apprentissage optimal par RF | 49 |
| 2.3.4. | Classification des pixels de la base de segmentation | 52 |
| 2.3.5. | Débruitage des résultats de segmentation du marquage NeuN | 56 |
| 2.4. | Méthode d'individualisation des neurones | 57 |
| 2.4.1. | Principe de la méthode – Hypothèses de travail | 58 |
| 2.4.2. | Segmentation du marquage NeuN par RF sur la base de données d'individualisation ... | 60 |
| 2.4.3. | Détection des centroïdes des neurones par calcul d'une cartographie des extrema | 60 |
| 2.4.4. | Individualisation des neurones par croissance de région..... | 62 |
| 2.4.5. | Optimisation multi-échelle de l'individualisation des neurones | 64 |
| 2.4.6. | Validation de la méthode d'individualisation des neurones | 73 |
| 2.4.6.1. | Validation par comptage des neurones par RDI | 73 |
| 2.4.6.2. | Validation basée sur des caractéristiques de populations des neurones..... | 74 |
| 2.4.6.3. | Validation par comparaison avec le comptage stéréologique..... | 76 |
| 2.5. | Résultats et discussion | 78 |
| 2.5.1. | Optimisation des paramètres de la méthode..... | 78 |
| 2.5.2. | Données computationnelles..... | 81 |
| 2.5.3. | Visualisation des résultats de l'individualisation | 82 |
| 2.5.4. | Comptage des neurones par RDI | 85 |
| 2.5.5. | Validation des méthodes d'individualisation basée sur des caractéristiques des neurones | 96 |
| 2.5.6. | Comptage stéréologique de référence | 98 |
| 2.6. | Conclusion..... | 103 |
| 3. | Cartographies mésoscopiques quantitatives relatives aux neurones..... | 105 |
| 3.1. | Principe et intérêt de l'intégration à l'échelle mésoscopique | 105 |
| 3.2. | Calcul des paramètres descripteurs des neurones | 106 |
| 3.3. | Génération de cartographies quantitatives à l'échelle mésoscopique | 108 |
| 3.4. | Exemples de cartographies obtenues pour des neurones..... | 110 |
| 3.4.1. | Cartographies du nombre de neurones..... | 110 |

| | |
|--|-----|
| 3.4.2. Cartographies de la morphologie des neurones..... | 111 |
| 3.4.3. Cartographies d'organisation des neurones..... | 114 |
| 3.5. Conclusion..... | 117 |
| 4. Discussion et perspectives | 119 |
| Annexe..... | 125 |
| Publications | 145 |
| Bibliographie..... | 147 |
| Résumé..... | 155 |
| Abstract | 156 |

Glossaire

| | |
|--------|--|
| ADN | Acide Désoxyribonucléique |
| CA | Corne d'Ammon |
| CAG | Cytosine, Adénine et Guanine |
| CIE | Commission Internationale de l'Eclairage |
| CZI | Carl Zeiss Imaging |
| DAB | Diaminobenzidine |
| DG | Gyrus Denté |
| EDF | Extended Depth of Focus (profondeur de champ étendu) |
| FDG | Fluorodesoxyglucose |
| GVF | Gradient Vector Flow |
| HSV | Hue Saturation Value (teinte, saturation et valeur) |
| IHC | Immunohistochimie |
| IRM | Imagerie par Résonance Magnétique |
| LBP | Local Binary Patterns |
| IG | Gain d'Information |
| LG | Lateral Geniculate (région du corps genouillé latéral) |
| LoG | Laplacian-of-Gaussian |
| MA | Maladie d'Alzheimer |
| MH | Maladie de Huntington |
| MP | Maladie de Parkinson |
| MIRCen | Molecular Imaging Research Center |
| PNH | Primate Non-Humains |
| RDI | Régions d'Intérêt |
| REL | Réticulum Endoplasmique Lisse |
| RER | Réticulum Endoplasmique Rugueux |
| RF | Random Forest (forêts aléatoires) |
| RGB | Red Green Blue (rouge, vert et bleu) |
| TEP | Tomographie par Emission de Positons |
| TDM | Imagerie par Tomodensitométrie |
| WSI | Whole Slide Imaging |

Introduction

Les neurones jouent un rôle majeur dans les fonctions avancées réalisées par le cerveau, la partie principale du système nerveux central qui contrôle le corps humain. Leurs caractéristiques et fonctionnement fascinent depuis longtemps les scientifiques. Surtout, le nombre de neurones est un élément essentiel pour l'étude de la neurogenèse ou de la neurodégénérescence dans le domaine des neurosciences. Afin de mieux comprendre les mécanismes du cerveau sain, d'identifier les mécanismes impliqués dans les maladies neurodégénératives et de développer de nouvelles approches thérapeutiques, de nombreuses études ont été réalisées pour déterminer le nombre et les caractéristiques des neurones depuis la fin des années 1980. Dans nombre de ces recherches, la stéréologie constitue un outil de référence qui permet d'estimer des informations quantitatives (ex. le nombre de neurones) et non biaisées sur des régions d'intérêt (RDIs). Cette technique est robuste mais repose sur une analyse manuelle, longue et fastidieuse à réaliser. De plus, la précision de la mesure dépend de la complexité des régions anatomiques, de plusieurs paramètres de la méthode qu'il est nécessaire d'ajuster et de l'expérience des biologistes. Aujourd'hui, des études plus poussées et plus précises sur les neurones sont en cours de développement en utilisant de nouvelles méthodes automatiques qui devraient permettre de mieux exploiter tant les modèles animaux que les images produites par des techniques d'imagerie en microscopie.

Les modèles animaux, rongeurs et primates non-humains (PNHs), ont des cerveaux de petite taille en comparaison de celui de l'Homme. Ils partagent cependant un pourcentage important de leur génome et également des caractéristiques cérébrales structurelles et fonctionnelles similaires en particulier entre les PNHs et les humains. Ces modèles sont d'une aide précieuse pour la recherche car ils aident à mieux explorer le cerveau sain ou pathologique et à tester plus efficacement de nouvelles approches thérapeutiques.

L'imagerie par résonance magnétique (IRM) et la tomographie par émission de positons (TEP) sont deux techniques d'imagerie *in vivo*. Elles sont utilisées dans les études précliniques parce qu'elles autorisent un suivi longitudinal qui permet de voir la structure du cerveau directement en 3D ainsi que le fonctionnement du cerveau vivant. Cependant, ces études sont limitées par la faible résolution spatiale qui est de l'ordre d'un à plusieurs millimètres ce qui empêche de mener les études sur le cerveau à l'échelle cellulaire. L'imagerie *post mortem* de données histologiques présente quant à elle un intérêt majeur pour l'étude des neurones. En effet, l'histologie permet d'accéder à une résolution spatiale très élevée jusqu'à quelques dixièmes de micromètres. Grâce à cette technique, les neurones peuvent être visualisés directement par microscopie optique. A l'aide de l'immunohistochimie (IHC) en utilisant l'anticorps NeuN et des méthodes d'acquisition de microscopie optique, les neurones peuvent ainsi être imagés à très haute résolution.

En contrepartie, cette très haute qualité d'image génère des données très massives allant de quelques dizaines de gigaoctets à plusieurs téraoctets pour une étude préclinique incluant plusieurs dizaines de sujets. De plus, certaines régions anatomiques du cerveau sont composées d'amas de neurones très denses (ex. gyrus denté dans l'hippocampe). De ce fait, il est nécessaire de mettre en œuvre des méthodes de traitement d'image capables de segmenter les neurones à partir du tissu et d'individualiser les neurones accolés. Il est de plus nécessaire de développer des méthodes d'analyse qui permettent de quantifier et d'analyser les informations d'intérêt produites pour des populations de neurones – leur morphologie et leur distribution dans les différentes régions anatomiques du cerveau.

Ce travail de thèse commence par l'étude menée sur trois coupes histologiques de cerveau de macaque sain. L'objectif est de développer et d'appliquer une méthode d'analyse d'images histologiques couleur afin de segmenter et de synthétiser l'information relative aux neurones marqués à l'aide de l'anticorps NeuN sur ces trois coupes histologiques. Ce travail doit permettre de quantifier des paramètres d'intérêt caractérisant les neurones dans le cerveau (nombre, taille, localisation, distribution, etc.). De plus, des stratégies d'intégration multi-échelles doivent être développées pour mener à bien ces explorations biologiques des populations de neurones de l'échelle cellulaire à l'échelle de régions anatomiques, de coupes voire, à terme, de cerveaux entiers.

Ce projet de thèse, dirigé par M. Thierry DELZESCAUX, a été réalisé au sein de l'équipe traitement de l'image du laboratoire des maladies neurodégénératives (CEA-CNRS-UMR 9199) à MIRCen (Molecular Imaging Research Center). MIRCen est un centre d'imagerie moléculaire préclinique, travaillant principalement dans le domaine des maladies neurodégénératives. Il possède un scanner de lames histologiques Axio ScanZ.1 (Carl Zeiss Microscopy GmbH Jena, Allemagne) qui permet d'effectuer des numérisations des coupes histologiques en champ clair et en fluorescence avec un grossissement allant jusqu'à X40 ce qui correspond à une résolution spatiale (taille du pixel/voxel dans l'image) de 0,11 μm dans le plan d'acquisition. MIRCen possède de plus deux aimants pour l'IRM du PNH (7 T) et du rongeur (11,7 T), deux appareils de TEP et des laboratoires d'histologie, de biochimie et de virologie.

Le manuscrit de cette thèse est subdivisé en quatre parties principales.

Le chapitre 1 décrit l'organisation du cerveau et introduit les principaux travaux menés pour étudier les neurones. Ces derniers aident à mieux comprendre les caractéristiques tant morphologiques que colorimétriques ainsi que le fonctionnement des différentes structures et les mécanismes impliqués dans les maladies neurodégénératives. Dans ce contexte, l'utilisation de modèles expérimentaux pertinents et de techniques d'imagerie adaptées est primordiale. L'imagerie *in vivo* permet de voir la structure et le fonctionnement du cerveau au niveau macroscopique tandis que l'imagerie *post mortem* – combinant histologie et microscopie optique – permet d'observer finement l'anatomie du cerveau *post mortem* à l'échelle cellulaire. Ce chapitre présente également les problématiques abordées dans ce travail de thèse.

Le chapitre 2 se concentre quant à lui sur la méthode de segmentation et d'individualisation des neurones dans les images histologiques couleur. Pour étudier les neurones, il faut d'abord les segmenter pour les isoler du tissu et du fond. Ensuite, il faut individualiser les neurones qui se touchent. Plusieurs travaux de validation comprenant la stéréologie ont été menés pour évaluer objectivement les performances de notre méthode. Au final, nous discutons les résultats obtenus, présentons des analyses préliminaires et concluons sur les avantages et les inconvénients de notre méthode.

Le chapitre 3 présente la méthode d'analyse d'images histologiques qui a pour objectif de synthétiser l'information obtenue après avoir individualisé un très grand nombre de neurones. Pour cela, nous quantifions des caractéristiques spécifiques (morphologiques et colorimétriques) des neurones et synthétisons ces informations sous forme de cartographies présentant des résolutions spatiales plus faibles. Cela devrait permettre à terme d'offrir de nouveaux outils afin d'analyser statistiquement de très grands volumes de données.

Le chapitre 4 présente les perspectives de nos travaux. Nous présentons les pistes envisageables afin d'améliorer notre méthode d'individualisation des neurones. Les défis restants pour passer de l'analyse de régions anatomiques à une coupe entière de cerveau, voire à un cerveau entier sont

également discutés. Finalement, nous terminons ce document de thèse par la présentation de plusieurs applications biologiques qui pourraient tirer avantage de l'utilisation de notre méthode.

1. Contexte

1.1. Le système nerveux central

1.1.1. Introduction générale

Le **système nerveux** joue un rôle important dans le système de régulation de la fonction physiologique du corps humain. Les humains vivent dans un environnement en évolution dont les changements impactent les fonctions physiologiques du corps humain. Cela nécessite un ajustement rapide dans le corps de sorte que ce dernier s'adapte aux changements de l'environnement tant internes qu'externes. Le système qui met en œuvre cette fonction de régulation est le système nerveux.

Le système nerveux (Figure 1) est principalement constitué par le **système nerveux central** et le **système nerveux périphérique**. Le système nerveux central, comprenant le cerveau et la moelle épinière, est la partie principale du système nerveux humain. Le système nerveux périphérique se compose quant à lui des nerfs dont certains recueillent des informations tandis que d'autres transmettent les ordres.

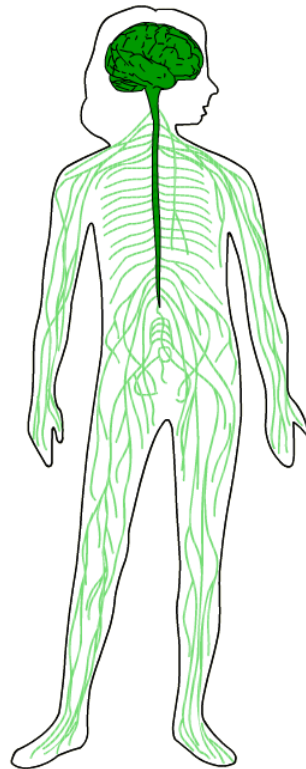


Figure 1 : Le système nerveux. Vert foncé : le système nerveux central, vert clair : le système nerveux périphérique. D'après <http://thebrain.mcgill.ca/>.

Le **cerveau** est la partie la plus importante du système nerveux central qui contrôle le corps humain. Il se compose de tissus mous. Les neuroanatomistes divisent généralement le cerveau en six parties principales (Figure 2) : le télencéphale, le diencephale, le mésencéphale, le cervelet, le pons et le bulbe rachidien. Chaque partie peut être décomposée en un grand nombre de structures. Des milliers des régions peuvent ainsi être identifiées dans le cerveau en distinguant les structures neuronales, la composition chimique voire les connectivités.

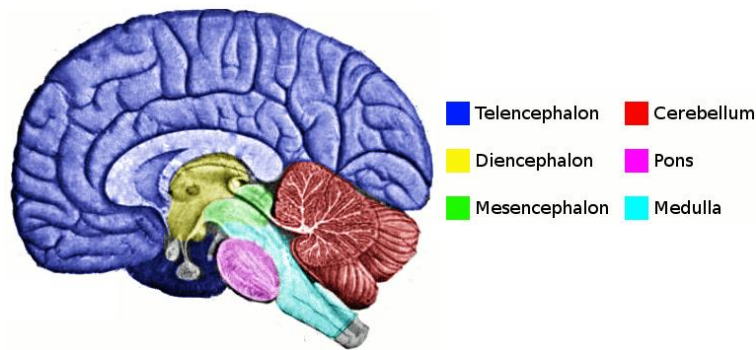


Figure 2 : La division du cerveau (Ranson, 1920).

1.1.2. Les principales structures et fonctions du cerveau

Le cerveau est divisé au niveau macroscopique en deux hémisphères cérébraux (Figure 3) qui sont reliés par le corps calleux composé de fibres nerveuses. La surface de l'hémisphère cérébral a de nombreuses fissures d'enroulement, appelées sillons, la partie apparente est appelée gyrus.

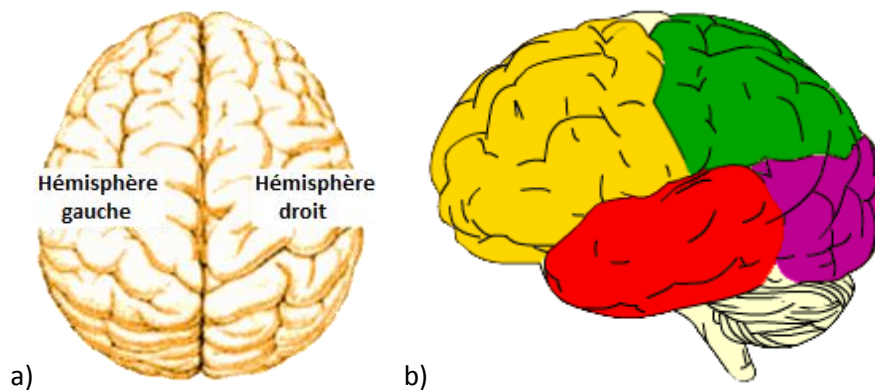


Figure 3 : Cerveau humain. a) Vue de dessus du cerveau. b) Quatre lobes principaux à la surface du cerveau humain. Jaune : lobe frontal, vert : lobe pariétal, rouge : lobe temporal, violet : lobe occipital. D'après <http://thebrain.mcgill.ca/>.

Le **cortex cérébral** est une région anatomique importante du cerveau, il couvre la partie externe de ce dernier. Il est lui-même recouvert par les méninges. Quatre lobes principaux sont identifiés à la surface de chaque hémisphère.

- Le **lobe frontal** (Figure 3b jaune) est la plus grande partie du cerveau humain par rapport aux autres lobes du cerveau. Cette structure est liée à la planification, la formation et l'expression du langage, l'autonomie et le contrôle volontaire des muscles.
- Le **lobe pariétal** (Figure 3b vert) traite divers types d'informations sensorielles (y compris la vision, l'audition, le toucher, le goût, la douleur, la température, etc.), mais est également lié au langage et à la mémoire. Sans cette partie, la compréhension de la parole et de l'écriture serait impossible.
- Le **lobe temporal** (Figure 3b rouge) est le centre auditif. Les parties supérieures des lobes temporaux permettent de comprendre la signification de la parole. Les parties inférieures participent quant à elles à la reconnaissance visuelle des objets et des visages. Les lobes temporaux aident également à mémoriser. Le lobe temporal droit est plus impliqué dans la mémoire visuelle, tandis que le gauche concerne plutôt la mémoire verbale.
- La fonction du **lobe occipital** (Figure 3b violet) est de traiter l'information visuelle. A ce moment précis, vos lobes occipitaux sont activés parce que vous êtes en train de lire cette thèse. Les lobes

occipitaux analysent aussi la forme, la couleur et le mouvement des objets pour rendre possible l'identification des objets.

Contrairement aux quatre lobes externes du cerveau, le lobe limbique (incluant notamment le gyrus cingulaire) n'est visible que sur une vue sagittale du cerveau. Il aide le système nerveux autonome à ajuster la pression sanguine, la fréquence cardiaque, la taille de la pupille et d'autres réponses psychosomatiques qui accompagnent les réactions émotionnelles. La Figure 4 montre les principales structures anatomiques en vue sagittale (cortex cérébral, thalamus, hippocampe, cervelet, noyau caudé, putamen, claustrum, pallidum, etc.).

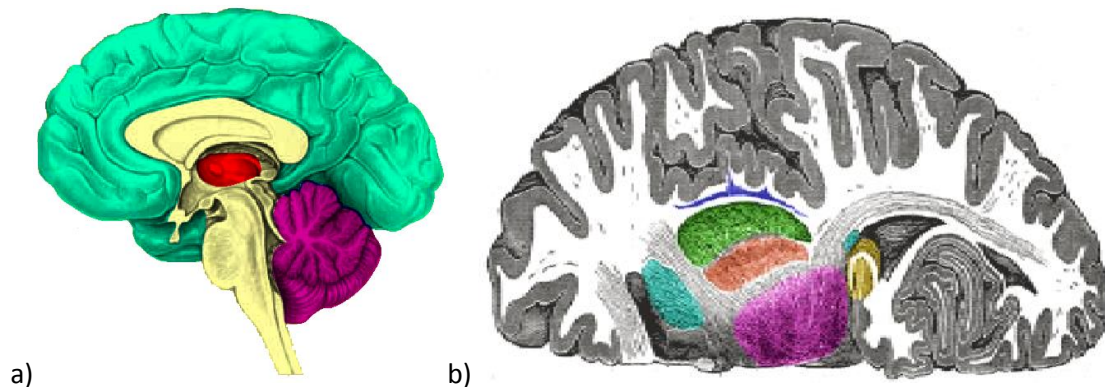


Figure 4 : a) Vue sagittale du cerveau humain à l'échelle macroscopique. Vert : cortex, rouge : thalamus, violet : cervelet. D'après <http://thebrain.mcgill.ca/>. b) Hémisphère cérébral en vue axiale (Gray, 1918). Cyan : noyau caudé, bleu : claustrum, vert : putamen, rouge : pallidum, violet : thalamus, jaune : hippocampe.

Les structures les plus importantes du cerveau sont présentées ci-dessous avec une brève description de leur fonction.

L'**hippocampe**, de même origine que le cortex (structure télencéphalique), est une structure importante du cerveau humain et des Mammifères qui est composée de deux parties, une de chaque côté du cerveau. Il est composé de trois sous-régions principales : le subiculum, la corne d'Ammon (composée des aires CA1, CA2, CA3 et CA4) et le gyrus denté. Il fait partie du système limbique du cerveau et joue un rôle dans la mémoire à court terme, la mémoire à long terme ainsi que la localisation spatiale.

Le **cervelet** joue un rôle important dans la perception sensorielle, la coordination et le contrôle moteur. Il est également associé à l'attention, au langage et à bien d'autres fonctions cognitives, et régit aussi des réactions telles que la peur et la joie. Le cervelet ne déclenche pas de mouvement, mais il reçoit des signaux du système sensoriel de la moelle épinière et d'autres régions du cerveau, affectant la coordination motrice, la précision et la synchronisation (Fine et al., 2002).

Le **thalamus** est la structure anatomique principale du diencephale. En plus de l'odeur, le reste des diverses informations sensorielles reçues sont transmises au cortex cérébral à travers le thalamus. De ce fait, le thalamus est parfois appelé le centre du cerveau.

Le **noyau caudé** et le **putamen** font partie du striatum. Ils sont responsables de la régulation des mouvements et influencent également divers types d'apprentissage (Nestler et al., 2008).

Le **globus pallidus** ou **pallidum** est une structure qui joue un rôle dans la régulation du mouvement volontaire (Dodson et al., 2015).

1.1.3. Composition cellulaire du cerveau

Le système nerveux humain est composé principalement de deux types de cellules à l'échelle microscopique: les **cellules neuronales (neurones)** et les **cellules gliales**. Le cerveau comprend à peu près 86 milliards de neurones et autant de cellules gliales (Azevedo et al., 2009). Les neurones jouent un rôle central majeur dans les fonctions avancées réalisées par le cerveau. Les cellules gliales, quant à elles, sont essentielles dans la mesure où elles permettent aux neurones de fonctionner correctement.

1.1.3.1. Les neurones

Les cellules neuronales, comme toutes les cellules de l'organisme, sont constituées d'un noyau contenant les gènes, d'un cytoplasme et de petits organites (réticulum endoplasmique rugueux, réticulum endoplasmique lisse, appareil de Golgi, mitochondries, etc.) qui permettent de produire de l'énergie et de fabriquer des protéines. Ils sont entourés par une membrane qui régule les échanges avec l'espace extracellulaire. Cette partie des neurones est appelée corps cellulaire, soma, ou encore périkaryon. Les neurones ont deux types de prolongements spéciaux : l'axone et les dendrites qui permettent la transmission de l'information sous forme d'influx nerveux. Les axones sont généralement très longs, et chaque neurone n'en a qu'un. Les axones transportent l'information du corps cellulaire du neurone vers d'autres neurones, avec lesquels ils établissent des connexions grâce aux synapses. Les axones peuvent également stimuler directement d'autres types de cellules, par exemple, les cellules musculaires. Les dendrites sont nombreuses et ont une structure arborescente ramifiée. Elles recueillent des informations et les transmettent aux corps cellulaires des neurones auxquels elles sont connectées. Un schéma complet d'un neurone est montré dans la Figure 5.

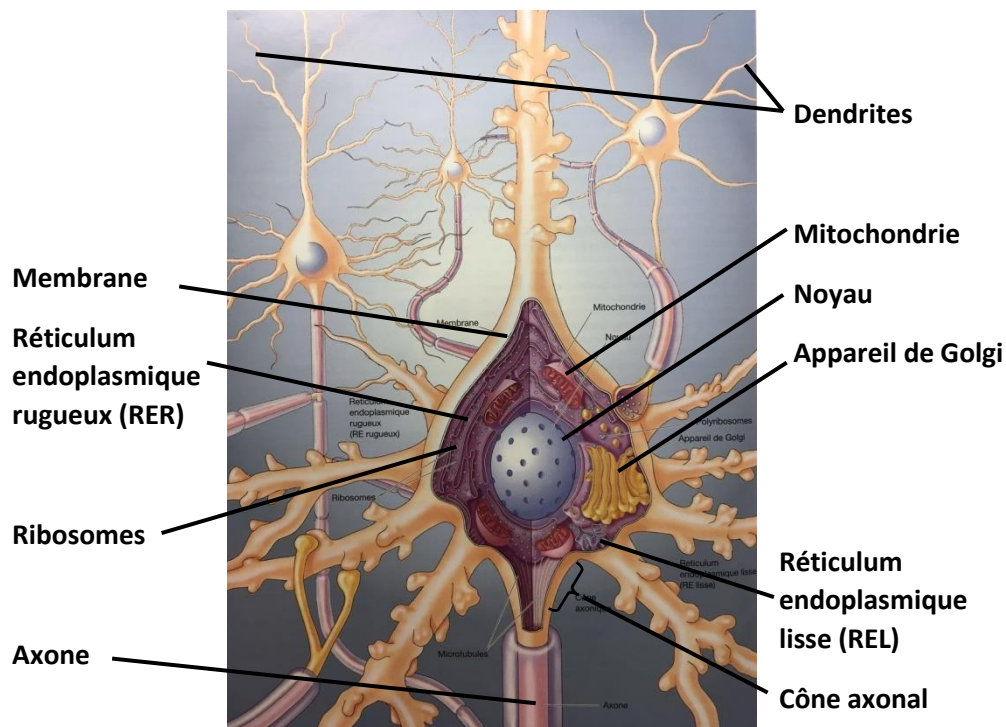


Figure 5 : Schéma de la structure du neurone (Bear et al., 2016).

Les neurones sont classés de différentes façons selon : leur nombre de neurites (tous les prolongements du corps cellulaire) un neurone pouvant être unipolaire, bipolaire, multipolaire (Figure 6); leur fonction (neurone sensitif, moteur ou interneurone, Figure 6); leur localisation (corticale, spinale, etc., Figure 7) ; l'identité de la substance (neurotransmetteur) qu'ils synthétisent

et libèrent (glutaminergique, cholinergique, etc.) ; leur forme (neurone pyramidal, étoilé, granulaire, Figure 7) ; leur taille (petit, moyen, grand voire géant) qui varie de 5 à 30 μm de diamètre (Andersen et al., 2016).

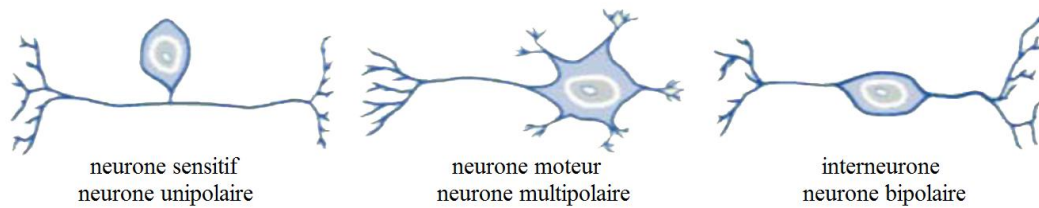


Figure 6 : Neurones classés en fonction de leur fonction ou du nombre de neurites. Les neurones sensitifs (à gauche) reçoivent des signaux des organes sensitifs et les transmettent par des axones courts au système nerveux central. Ils sont dits unipolaires car ils ne possèdent qu'un seul neurite (une courte extension qui se divise rapidement en deux branches). Les neurones moteurs (au milieu) conduisent les commandes motrices du cortex à la moelle épinière ou de la moelle épinière aux muscles. Ils sont dits multipolaires car ils comprennent plus de deux neurites. Les interneurons (à droite) interconnectent divers neurones dans le cerveau ou la moelle épinière. Ils sont dits bipolaires car ils possèdent deux neurites. D'après <http://thebrain.mcgill.ca/>.

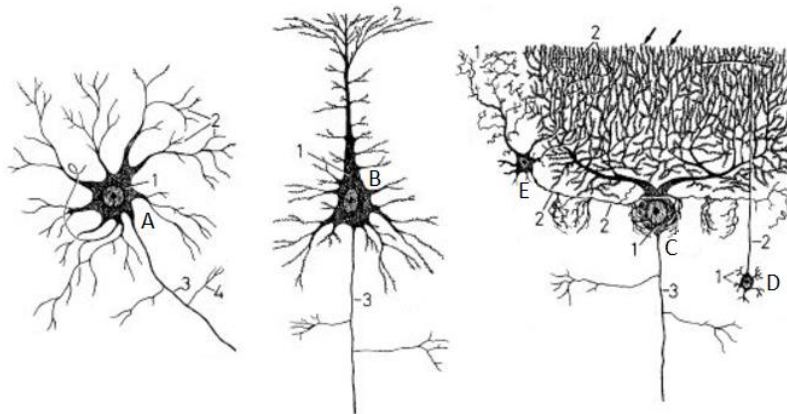


Figure 7 : Quelques exemples descriptifs de neurones. (A) Cellule de la moelle épinière qui possède un soma étoilé A1, un seul axone A3 émettant des collatérales A4 et de nombreuses dendrites A2. (B) Cellule pyramidale de la substance grise du cerveau qui possède un soma triangulaire B1, un axone B3 et de nombreuses dendrites B2. (C) Cellule de Purkinje du cervelet. (D) Cellules granulaires et (E) cellules en corbeille, qui possèdent un grand nombre de dendrites (D1, E1) mais dont les axones sont courts (D2, E2), elles s'articulent avec la cellule de Purkinje. D'après [http://www.db-gersite.com/HISTOLOGIE/HISTGENE/histgen1/histgen7/histgen7.htm#II.2. LES 5 TYPES](http://www.db-gersite.com/HISTOLOGIE/HISTGENE/histgen1/histgen7/histgen7.htm#II.2.LES%205%20TYPES).

1.1.3.2. Les cellules gliales

Les cellules gliales supportent les neurones en fournissant l'alimentation, le soutien physique et la protection. Lorsque les neurones meurent, les cellules gliales permettent d'éliminer les déchets. Enfin, elles accélèrent la conduction neurale en se comportant comme une gaine isolante autour de certains axones.

Les cellules gliales peuvent être séparées en trois catégories : les astrocytes, les cellules microgliales et les oligodendrocytes (Figure 8).

- Les **astrocytes** en étoile, qui forment la barrière hémato-encéphalique, fournissent aux neurones un support biochimique, leur apportent des nutriments, maintiennent l'équilibre

des ions extracellulaires et jouent un rôle immunitaire, de réparation et de cicatrisation. Ils digèrent et éliminent toutes sortes de débris.

- Les **cellules microgliales** constituent la première ligne de défense contre les « envahisseurs » étrangers. Elles détectent et phagocytent les débris et les agents infectieux présents dans le cerveau. Ce sont les macrophages du cerveau.
- Les **oligodendrocytes** forment quant à eux la gaine de myéline qui entoure les axones de nombreux neurones. La façon très particulière dont les oligodendrocytes enveloppent les axones accélère la conduction neurale. Les cellules de Schwann ont les mêmes fonctions dans le système nerveux périphérique.

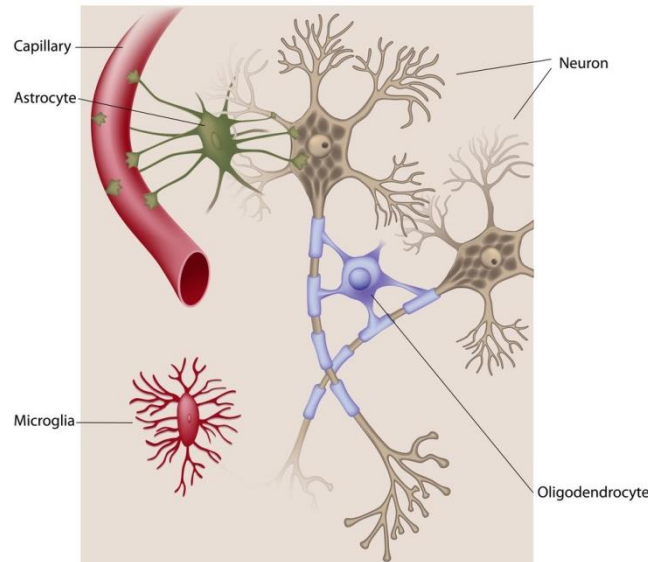


Figure 8 : Schéma de cellules gliales. D'après <http://brainpictures.org/>.

1.1.4. Etudes relatives aux neurones

Malgré les connaissances fondamentales actuelles de la structure et de la fonction des zones cérébrales, le cerveau humain reste encore aujourd'hui un mystère et de nombreuses études s'efforcent d'en dévoiler peu à peu les secrets. Les scientifiques étudient les différences cérébrales entre hommes et femmes, jeunes et vieux, droitiers et gauchers, etc. Ils s'intéressent notamment aux différences notables à l'échelle cellulaire. Dans les cas pathologiques, les mécanismes impliqués dans la maladie représentent un point important à comprendre. La comparaison entre des cerveaux sains et pathologiques ainsi que la possibilité de mener des recherches à l'échelle cellulaire offrent la possibilité de mieux comprendre les mécanismes et devraient fournir des outils très pertinents pour développer et valider de nouvelles thérapies.

1.1.4.1. Etude du cerveau sain

Age, sexe et latéralisation

Le cerveau humain est divisé en deux hémisphères. Un grand nombre d'études ont montré que le cerveau humain est asymétrique tant au niveau des structures anatomiques que de ses fonctions. Ceci est possiblement influencé par l'âge, le sexe, la latéralisation (gaucher, droitier), etc. Récemment, une grande étude (Guadalupe et al., 2016) sur l'asymétrie du cerveau humain a été menée sur 15 847 individus (7524 hommes et 8323 femmes, dont 959 gauchers et 11236 droitiers) dans le monde. L'asymétrie est décrite par un indice de l'asymétrie, défini comme la différence de volume relatif entre la structure gauche et droite par rapport à son volume bilatéral total: (Gauche -

Droite) / (Gauche + Droite). Il est constaté des différences liées au sexe dans les asymétries du globus pallidus et du putamen. Ces deux structures avec l'hippocampe et le thalamus présentent les asymétries les plus fortement héritées parmi les sept structures analysées (noyau accumbens, amygdale, noyau caudé, globus pallidus, hippocampe, putamen et thalamus). L'asymétrie entre droitiers et gauchers n'est pas encore démontrée mais au cours du vieillissement, il y a des changements dans l'asymétrie moyenne du putamen. D'ailleurs, une latéralisation liée à une asymétrie sous-corticale n'a pas été détectée. Joliot et al. (2016) ont étudié la relation entre l'asymétrie dans la connectivité intrinsèque intra-hémisphérique et la dominance hémisphérique du langage. Ils ont de plus considéré un effet potentiel de la latéralisation sur cette relation chez 290 individus sains dont 52,7% étaient gauchers. Maingault et al. (2016) ont étudié l'asymétrie liée à latéralisation en fonction de paramètres morphologiques (l'épaisseur, la surface, le volume de cortex et l'épaisseur des sillons, etc.) chez 250 individus sains incluant 120 gauchers. Afin de mieux comprendre les changements morphologiques observés à une échelle macroscopique, il serait intéressant d'étudier leur impact directement sur les neurones à l'échelle cellulaire.

Morphologie de neurones

De nombreuses études ont été menées sur des cerveaux sains pour étudier le développement et vieillissement des neurones en étudiant le nombre et la distribution des neurones. Il est estimé que le cerveau humain contient approximativement 100 milliards de neurones et dix fois plus de cellules gliales (Williams et Herrup, 1988; Ullian et al., 2001; Doetsch, 2003; Nishiyama et al., 2005; Noctor et al., 2007; Herculano-Houzel, 2009). Cependant, ces chiffres sont des estimations qui n'ont pas été validées par des études expérimentales du fait de la très grande taille du cerveau humain et de l'hétérogénéité des structures anatomiques qui empêchent le comptage des cellules par des méthodes de stéréologie (méthode de comptage qui sera présentée dans la section 2.1.2). En effet, le comptage des neurones dans le cerveau humain a été réalisé par stéréologie dans quelques structures anatomiques plutôt homogènes, par exemple dans le cortex cérébral (Pakkenberg et Gundersen, 1997; Larsen et al., 2006; Pelvig et al., 2008; Karlsen et Pakkenberg, 2011; Walløe et al., 2014). Le Tableau 1 présente des estimations du nombre de neurones et de cellules gliales dans le cortex humain. Le nombre absolu de neurones dans le cerveau humain reste cependant approximatif. Par exemple, Azevedo et al. (2009) ont estimé le nombre de neurones dans des cerveaux entiers à $86,1 \pm 8,1$ milliards et celui des cellules gliales à $84,6 \pm 9,8$ milliards à partir de cerveaux sains de quatre hommes âgés de 50, 51, 54 et 71 ans.

Tableau 1 : Nombres de neurones et de cellules gliales dans le cortex cérébral humain estimés par stéréologie.

| Références | Sexe | Nombre d'individu | Age | Nombre ($\times 10^9$) | |
|-------------------------------|------|-------------------|-------------------------------|--------------------------|------------------|
| | | | | Neurones | Cellules gliales |
| Pakkenberg et Gundersen, 1997 | M | 62 | 19-87 ans | 23 | / |
| | F | 32 | 18-93 ans | 19 | / |
| Larsen et al., 2006 | M/F | 5/5 | Nouveau-nés 38-42 semaines | 19,8 | 5,48 |
| Pelvig et al., 2008 | M | 13 | 19-87 ans | 26,3 | 38,9 |
| | F | 18 | 18-93 ans | 21,4 | 27,9 |
| Azevedo et al., 2009 | M | 4 | 50-71 ans | 16,3 | 24,2 |
| Karlsen et Pakkenberg, 2011 | F | 6 | 60-80 ans | 17,8 | 18,2 |

Il existe également une grande variabilité de la morphologie des neurones dans le cerveau humain. Ceci constitue un travail de recherche important à mener pour obtenir des mesures morphologiques fiables des neurones. Flood et Coleman (1988) ont montré qu'il pourrait y avoir une corrélation entre la diminution du nombre et la taille des neurones. Les études ayant permis de détecter une perte neuronale ont également montré une réduction de la taille des neurones au cours du vieillissement. Korbinian (Brodman, 1909) a étudié la structure et la distribution des neurones dans le cortex

cérébral, ce qui lui a permis d'établir un partitionnement du cortex cérébral. Alberts et al. (2002) ont montré que le noyau occupe en moyenne 10% du corps cellulaire dans les neurones. (Davies, 2002) a montré quant à lui que le noyau des neurones peut varier de 3 à 18 μm de diamètre ce qui met en évidence un manque de précision et de consistance dans les estimations proposées. Gittins et Harrison (2004) ont estimé que la surface moyenne des neurones dans la zone 24b (Vogt et al., 1995) du cerveau humain est comprise entre 170 et 373 μm^2 . Dans le livre Neurosciences (Bear et al., 2016), il est indiqué que le corps cellulaire d'un neurone typique présente un diamètre d'environ 20 μm . Andersen et al. (2016) ont quant à eux estimé que la taille des neurones de l'hippocampe varie de 5 à 30 μm de diamètre.

Si beaucoup d'études concernent l'Homme, de nombreux travaux ont également été menés sur des modèles animaux. Ishizuka et al. (1995) et Pyapali et al. (1998) ont étudié les arbres dendritiques des neurones pyramidaux dans l'hippocampe de rongeurs. Scorcioni et al. (2004) se sont intéressés aux neurones pyramidaux dans l'hippocampe des rongeurs à l'aide de trente paramètres morphologiques (surface de soma, nombre de bifurcation, nombre de branches, nombre d'arbres dendritiques issus de soma, distance d'extension totale de l'arborisation dendritique, distance maximale entre soma et la terminaison la plus lointaine, etc.). Kolb et al. (1995) ont étudié les couches individuelles du cortex cérébral d'un macaque (Figure 9) à l'aide de deux marqueurs : coloration des corps de Nissl (noyaux de toutes les cellules) et Cytochrome oxydase (niveaux d'activité métabolique neuronale).

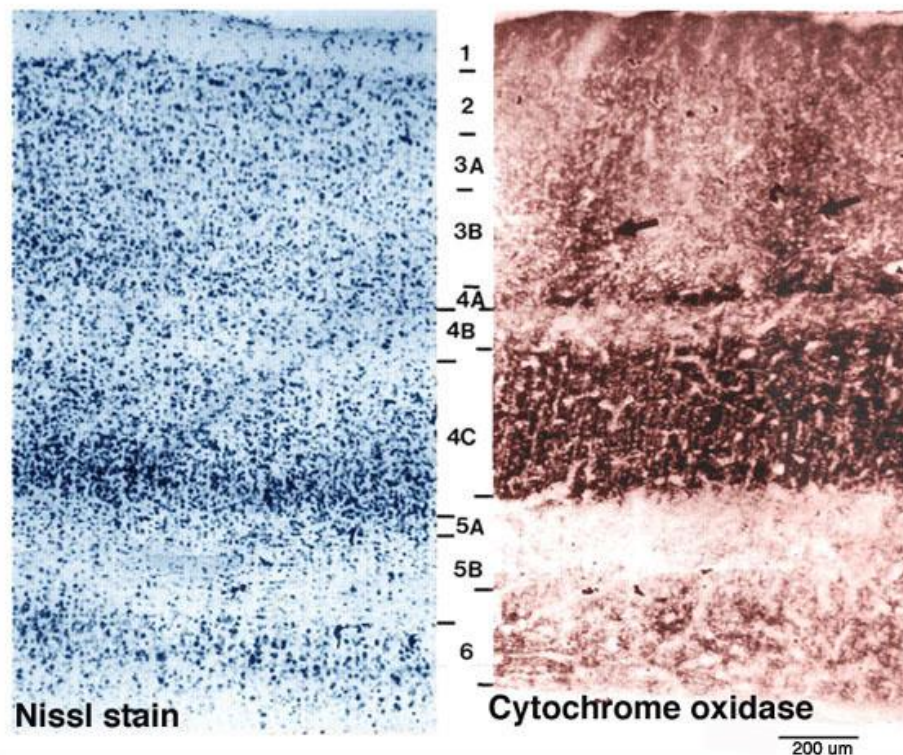


Figure 9 : Coupes transversales marquées par Nissl (à gauche) et par Cytochrome oxydase (à droite) du cortex visuel d'un macaque montrant différentes couches cellulaires.

Si l'étude des neurones chez des sujets sains est un domaine de recherche très important pour comprendre le fonctionnement du cerveau ainsi que son évolution au cours du développement et du vieillissement, elle constitue également un élément majeur d'étude dans le cadre des maladies neurodégénératives dans la mesure où les neurones y jouent un rôle central.

1.1.4.2. Cas des maladies neurodégénératives

Une maladie neurodégénérative se traduit entre autre par une perte neuronale et/ou une destruction des gaines de myéline dans le cerveau et la moelle épinière. Les neurones cérébraux et

spinaux ont des fonctions différentes, que ce soit pour le contrôle du mouvement, le traitement de l'information sensorielle mais aussi pour la prise de décisions. Les neurones du cerveau et de la moelle épinière ne sont généralement pas renouvelés, même si des études (Spalding et al., 2013) ont montré que le renouvellement des neurones au cours de la vie peut se produire dans l'hippocampe. Cependant, la dégénérescence est plus rapide que cette neurogenèse. Par conséquent, des dommages excessifs peuvent être dévastateurs voire irréversibles. Les maladies neurodégénératives s'aggravent avec le temps conduisant à d'importants dysfonctionnements (la perte des fonctions mentales et surtout de la mémoire, une bradykinésie, une rigidité, des tremblements de repos, des troubles moteurs et cognitifs, la perte de l'autonomie, etc.).

Les maladies neurodégénératives ont donc un impact très important sur la santé publique et dans la plupart des cas, il n'y a que peu voire pas de remède à ce jour. Elles constituent un défi majeur d'autant plus qu'une augmentation importante du nombre de cas est d'ores et déjà prévue dans les années à venir. La plateforme d'imagerie préclinique MIRCen s'intéresse tout particulièrement à trois maladies neurodégénératives : la maladie d'Alzheimer, la maladie de Parkinson et la maladie de Huntington.

La maladie d'Alzheimer

La maladie d'Alzheimer (MA) est un trouble dégénératif neurologique progressif. Il se caractérise par la présence d'agrégats pathologiques : les plaques amyloïdes et les dégénérescences neurofibrillaires. Les signes cliniques se caractérisent par la démence, comme l'altération de la mémoire, l'aphasie, la perte de conscience, l'atteinte de compétences du système visuel (désorientation), le dysfonctionnement moteur et les changements de comportement (Burns et Iliffe, 2009; Bick, 1987). La maladie peut être liée à plusieurs facteurs de risque, comme les antécédents familiaux (environ 70% du risque est considéré génétique avec de nombreux gènes impliqués), les traumatismes crâniens, la dépression, l'hypertension, etc. (Burns et Iliffe, 2009; Ballard et al., 2011). En 2015, à peu près 45 millions de personnes dans le monde souffraient de la MA (GBD 2015 Disease and Injury Incidence and Prevalence Collaborators, 2016) et elle a entraîné environ 1,9 million de décès (GBD 2015 Mortality and Causes of Death, Collaborators, 2016).

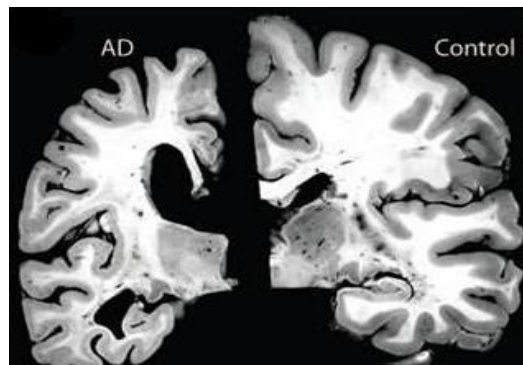


Figure 10 : Atrophie sévère du cerveau dans la maladie d'Alzheimer (Holtzman et al., 2011). Deux coupes de cerveaux humains acquises *post mortem* à l'échelle macroscopique: coupe d'un patient décédé de la MA à gauche et coupe d'un sujet sain à droite.

L'atrophie cérébrale est évidente et très importante dans le cas de la MA (Figure 10). L'étude à l'échelle microscopique montre que la maladie se caractérise par la mort de neurones et/ou des changements morphologiques de ces derniers (Bick, 1987). Outre les problèmes posés par l'éthique, l'accès à ce type d'information reste aujourd'hui difficile à obtenir du fait de la complexité des formes de cette maladie et des quantités colossales d'information qu'il faudrait produire et pouvoir

analyser pour les exploiter pleinement. L'étude de données *post mortem* (présentée dans la section 1.3.2) constitue un élément de recherche majeur dans ce contexte.

La maladie de Parkinson

La maladie de Parkinson (MP) est une maladie neurologique chronique dégénérative, la deuxième des maladies neurodégénératives la plus fréquente en terme du nombre de patients, après la MA. Elle se caractérise principalement par la bradykinésie, des tremblements, une rigidité musculaire, une instabilité posturale, etc. ("Parkinson's Disease Information Page," 2016). En 2015, la MP a affecté 6,2 millions de personnes et a entraîné environ cent mille décès dans le monde (GBD 2015 Disease and Injury Incidence and Prevalence Collaborators, 2016) (GBD 2015 Mortality and Causes of Death, Collaborators, 2016).

La MP affecte le système nerveux central. Ses causes sont aujourd'hui encore mal connues, mais les pathologistes pensent qu'elle est liée à divers facteurs : une mutation d'origine génétique, une maladie cérébro-vasculaire, une artériosclérose cérébrale, une infection, une intoxication, des traumatismes, la drogue (Kalia and Lang, 2015; Barranco Quintana et al., 2009; Mao et al., 2016). Les études menées ont montré que dans la MP, un nombre important de neurones dégénèrent dans la substance noire, une région du mésencéphale, ainsi que dans les faisceaux nigro-striés (reliant la substance noire au striatum) qui sont atteints. Il en résulte une quantité insuffisante de dopamine produite dans ces régions. Les altérations cérébrales se propagent ensuite à la sphère dopaminergique, mais d'autres systèmes de neurotransmetteurs sont également affectés (Figure 11).

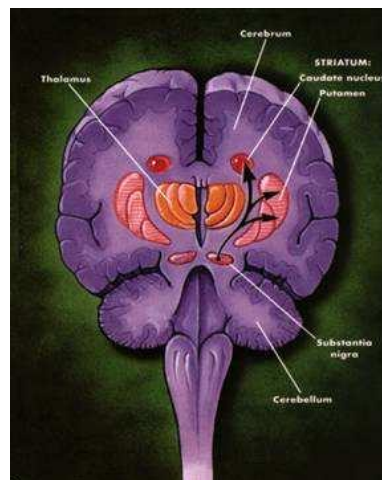


Figure 11 : Les symptômes de la maladie de Parkinson sont le résultat de la dégénérescence des neurones produisant la dopamine dans la substance noire. D'après www.sfn.org.

La maladie de Huntington

La maladie de Huntington (MH), aussi connue sous le nom de Chorée de Huntington, est une maladie héréditaire dominante. A cause de mutations génétiques ou de l'augmentation excessive du nombre de répétitions des tri-nucléotides CAG (cytosine, adénine et guanine) au sein du gène de la huntingtine, les cellules nerveuses sont altérées et leur lien avec des cellules en charge du contrôle musculaire est rompu. Ceci conduit à une dégradation progressive du système nerveux central, à l'apparition de gestes incohérents et anormaux (mouvements choréiques ou danse de Saint-Guy), des léthargies, des troubles de l'équilibre, des contractions incontrôlables (Warby et al., 1993; Frank, 2014; Dayalu et Albin, 2015). Dans les cas extrêmes, une démence peut se développer voire même

mener à la mort. La MH affecte notamment les personnes de descendance européenne, environ 4 à 15 personnes pour cent mille souffrent de cette maladie.



Figure 12 : Cerveaux humains imagés *post mortem*, sujet sain à gauche et un patient souffrant de la maladie de Huntington à droite. Source: Medical Genetics, Second Edition, Ed. Mosby.

L'atrophie cérébrale (Figure 12) se présente essentiellement dans le striatum, formé du noyau caudé et du putamen, ce qui est caractéristique de la MH.

Dans le cadre des maladies neurodégénératives, accéder à une information sur le nombre, la forme voire la distribution des neurones serait d'un grand intérêt pour mieux comprendre l'évolution de ces pathologies : décours temporel, zones touchées, description fine des remaniements à l'échelle cellulaire, etc. Il serait également possible d'utiliser ces informations pour développer et valider de nouvelles approches thérapeutiques.

1.2. Modèles animaux en neurosciences

1.2.1. La recherche translationnelle

Dans le domaine de la biologie, la recherche translationnelle occupe une place de plus en plus importante. Elle crée des ponts entre la recherche fondamentale et la recherche clinique. La recherche fondamentale se concentre sur la compréhension du fonctionnement des organes sains et des mécanismes qui aboutissent aux maladies tandis que la recherche clinique se concentre sur le traitement des patients. La recherche translationnelle permet de tester de nouvelles approches thérapeutiques sur des modèles animaux qui miment la pathologie observée chez l'Homme. En cas de succès lors d'un test thérapeutique mené sur les modèles animaux, la recherche sera étendue à l'homme à l'occasion d'un essai clinique. La recherche translationnelle rend possible un échange des connaissances dans les deux sens. Elle est capable, dans certains cas, d'identifier les pistes qui ont peu de chances d'obtenir des résultats positifs en recherche clinique et elle favorise l'exploration de nouvelles théories et hypothèses de travail. Cependant, il est important de noter que la grande différence inter-espèces qui existe entre le rongeur et l'Homme en passant par le PNH constitue une source importante d'échecs dans ce processus translationnel mais reste une voie de recherche primordiale pour espérer développer de nouvelles approches et apporter des solutions thérapeutiques à toutes ces maladies.

1.2.2. Modèles animaux

La recherche sur les modèles animaux aide les scientifiques à comprendre comment le cerveau se développe et fonctionne chez un sujet sain et à voir ce qui pose problème dans le cas d'une maladie. La Figure 13 montre la grande diversité en termes de taille des cerveaux pour différentes espèces.

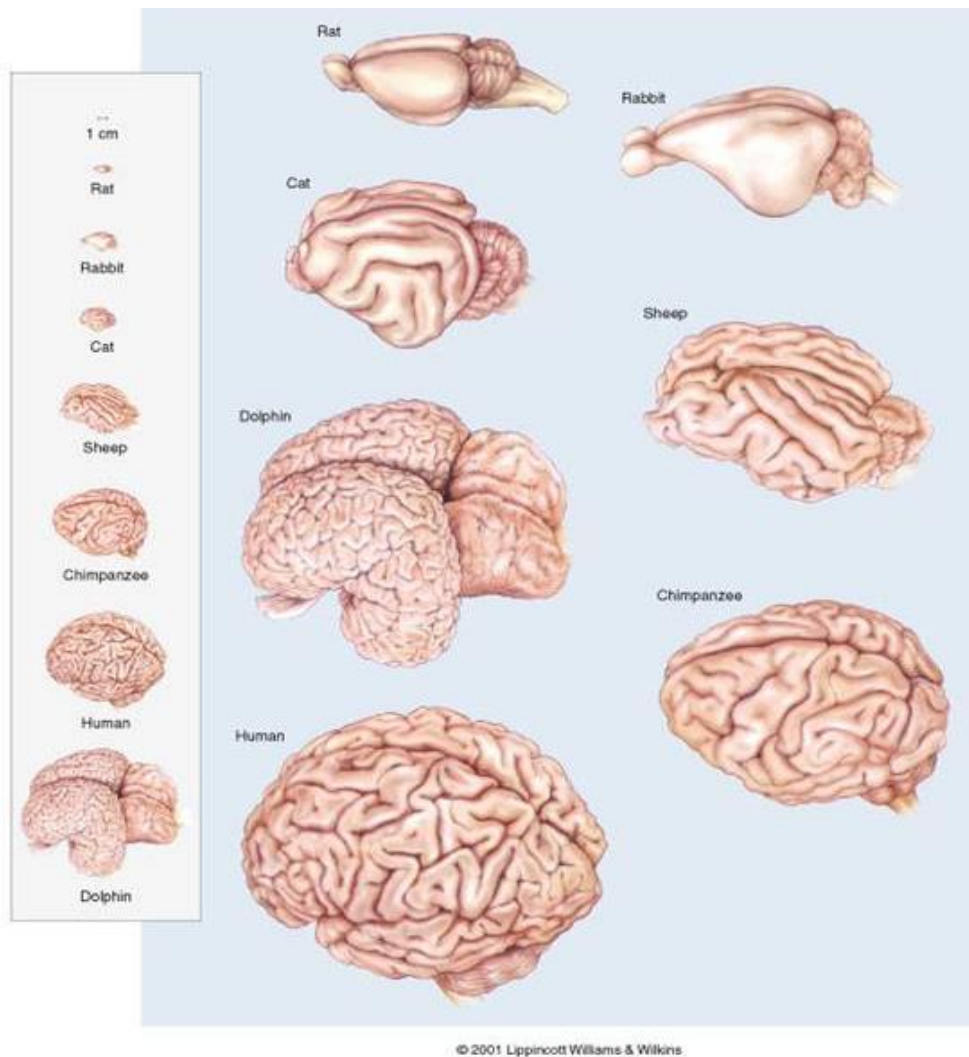


Figure 13 : Comparaison des cerveaux de différentes espèces (Lippincott Williams et Wilkins, 2001).

A MIRCen, nous travaillons en recherche préclinique sur les maladies neurodégénératives en utilisant les modèles rongeurs (souris et rat) et primates (microcèbe et macaque) (Figure 14).

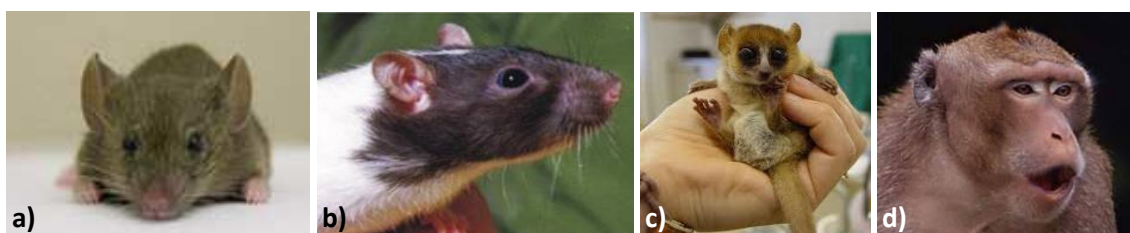


Figure 14 : Espèces étudiées à MIRCen. a) souris, b) rat, c) microcèbe et d) macaque.

Une des raisons pour lesquelles les rongeurs sont si utiles pour la recherche scientifique est la très bonne connaissance de leur génome ainsi que sa proximité avec celui de l'Homme. Les souris ne ressemblent pas du tout à l'humain, mais en fait partagent jusqu'à 80% de leur ADN (acide désoxyribonucléique) avec celui-ci (<https://canvas.brown.edu>). En ce qui concerne le cerveau des primates, celui-ci ressemble davantage encore au cerveau humain. Les primates partageant plus de caractéristiques cérébrales structurelles et fonctionnelles avec les humains, ils permettent de mieux comprendre les mécanismes de fonctionnement du cerveau. Comme les hommes ont des capacités cognitives non trouvées chez les macaques et comme leurs cerveaux ne sont pas strictement

identiques, Passingham (2009) a fait remarquer qu'il faut être prudent sur les résultats obtenus sur ces modèles et propose de confirmer les résultats directement chez les hommes quand c'est possible.

Dans le cadre des maladies neurodégénératives, certaines espèces sont ciblées pour créer des modèles animaux *ad hoc*. Pour cela, différentes stratégies sont possibles : manipulation génétique chez la souris pour intégrer dans son génome des gènes caractéristiques de la maladie, utilisation de vecteurs viraux pour insérer localement des gènes d'intérêt, lésions chimiques excitotoxiques pour reproduire une lésion des tissus cérébraux (Onos et al., 2016; Platt et al., 2013; Dong et al., 2009, etc.). L'obtention d'un modèle animal pertinent met en jeu une bonne connaissance des processus pathologiques et sa proximité avec la vraie pathologie s'avère d'une importance capitale pour valider l'efficacité des nouvelles approches thérapeutiques testées. La possession de modèles animaux très réalistes constitue un avantage compétitif majeur pour proposer et valider de nouvelles approches thérapeutiques. La validité de ces modèles est évaluée par des critères morphologiques et comportementaux que l'on compare à ceux observés chez les patients humains.

Les modèles animaux sont d'une importance primordiale pour la recherche cérébrale physiopathologique, cependant, ils présentent des avantages et des inconvénients. Le petit animal est facilement accessible car il a une capacité de reproduction importante et sa durée de vie courte réduit d'autant la durée des études, toutefois son cerveau est de petite taille et de morphologie différente en comparaison de celui de l'Homme. Le PNH est plus difficile d'accès du fait des règles éthiques, de son taux de reproduction faible et de sa durée de vie de plusieurs années qui réduisent les effectifs et allongent la durée des études. En recherche préclinique fondamentale, les études sont en première intention réalisées sur les rongeurs, puis chez les PNHs si le succès de l'étude permet d'envisager un passage en clinique.

En raison de problèmes éthiques, les recherches menées sur des cultures de cellules *in vitro* en laboratoire et le développement de modèles informatiques (simulations numériques) de plus en plus complexes et performants ont été proposées pour limiter voire supprimer l'utilisation des modèles animaux en recherche préclinique (Doke et Dhawale, 2015). Malheureusement, force est de constater que ces alternatives s'avèrent aujourd'hui insuffisantes pour se passer complètement de l'utilisation de tels modèles car ils ne miment qu'imparfaitement ce qui se passe *in vivo* dans les modèles animaux (complexité des systèmes biologiques, dynamique des processus, etc.). Si ces alternatives occupent une place croissante et sont pertinents dans certaines études, le recours à des modèles animaux reste encore aujourd'hui indispensable pour mener une recherche scientifique au plus haut niveau. Il est également très important de souligner que toutes ces recherches sont menées dans un cadre très contrôlé et sont soumises à l'approbation de comités d'éthiques vigilants ainsi qu'à de bonnes pratiques de laboratoire pour limiter la douleur et le nombre des animaux inclus dans ces études.

1.2.3. Intérêt des modèles animaux

L'avantage de ces modèles est la réalisation de ces études *in vivo* ce qui permet la mise en place d'un suivi longitudinal ce qui apporte une dimension supplémentaire cruciale (évolution temporelle, définition de biomarqueurs caractéristiques de la pathologie, ordre d'apparition et localisation anatomique des événements, etc.). Les progrès de l'instrumentation ont de plus permis de développer des systèmes d'imagerie (présentés dans la section 1.3.1) qui reproduisent de façon similaire chez l'animal les images du cerveau qui peuvent être obtenues chez l'Homme. Ceci renforce l'aspect translationnel de la recherche préclinique et est aujourd'hui considéré comme un vecteur d'accélération pour le développement et la validation de nouvelles thérapies. Enfin, au terme de

l'étude *in vivo*, l'accès direct aux tissus prélevés dans des conditions optimales rend possible de pousser l'analyse à l'échelle microscopique en profitant pleinement de la grande diversité des marqueurs histologiques existants (présentés dans la section 1.3.2) et des systèmes de microscopie. Les cycles de vie relativement courts des différents modèles animaux utilisés peuvent aller de quelques mois à quelques années en fonction des espèces ce qui est compatible avec la durée des recherches précliniques menées.

1.3. Imagerie du cerveau *in vivo/post mortem*

Afin d'étudier le cerveau, différentes techniques d'imagerie nous permettent de voir la structure et le fonctionnement du cerveau *in vivo* et d'observer l'anatomie du cerveau *ex vivo* à l'échelle cellulaire. Cependant, chaque méthode présente naturellement ses propres limites. Nous allons introduire les principales techniques d'imagerie *in vivo* dans la section 1.3.1 et d'imagerie *post mortem* dans la section 1.3.2.

1.3.1. L'imagerie *in vivo*

Les études du cerveau en neurosciences sont basées sur l'observation par des techniques d'imageries *in vivo* : imagerie par tomodensitométrie (TDM), imagerie par résonance magnétique (IRM), imagerie par tomographie à émission de positons (TEP). Ces techniques permettent de réaliser de façon répétée des études chez le même sujet vivant, comme pour l'étude de la structure du cerveau, l'identification de la localisation d'une lésion cérébrale, le suivi thérapeutique de pathologies, le suivi longitudinal du métabolisme de l'oxygène ou du glucose, le suivi du débit sanguin cérébral, l'étude de fonctionnement des zones cérébrales en exécutant des actions cognitives, etc. Ici, nous allons présenter deux techniques, l'IRM et la TEP, qui sont largement utilisées en neurosciences aujourd'hui.

1.3.1.1. Imagerie par résonance magnétique

L'**IRM** est une technique d'imagerie qui utilise des phénomènes de résonance magnétique pour obtenir des signaux électromagnétiques du corps humain et reconstituer une information anatomique ou fonctionnelle humaine.

Comme environ 70 % du corps humain est composé d'eau, l'IRM peut être générée en fonction des atomes d'hydrogène contenus dans l'eau. Lorsqu'un objet est placé dans un champ magnétique produit par un aimant et est soumis à une onde électromagnétique appropriée, le sens de rotation des atomes d'hydrogène est alors changé pour résonner. Les atomes sont activés, absorbent l'énergie de l'onde électromagnétique et passent en état de haute énergie. Quand l'onde électromagnétique est coupée, les atomes activés vont revenir à un état d'équilibre en libérant des ondes électromagnétiques qui sont enregistrées par des détecteurs spécialisés. Comme les différents tissus produisent différents signaux d'ondes électromagnétiques, la position et le type des noyaux qui composent l'objet vont ainsi pouvoir être analysés. Grâce à ces informations, l'IRM de l'objet étudié pourra être reconstituée en trois dimensions.

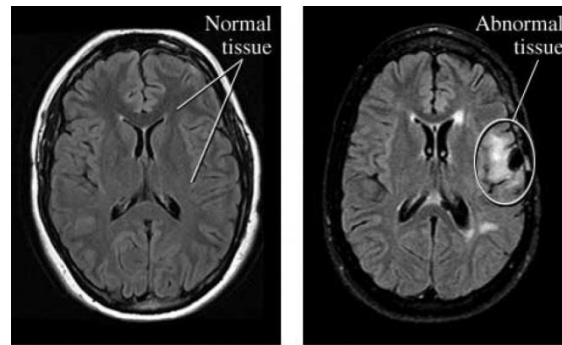


Figure 15 : Exemple d'IRM de cerveau humain normal et anormal. D'après <http://brainpictures.org/>.

L'IRM est non invasive et n'irradie pas le sujet. Elle permet d'imager la structure du cerveau vivant et d'analyser son organisation avec une résolution spatiale de l'ordre du millimètre, notamment lorsque des tumeurs ou des réorganisations anatomiques après des traumatismes cérébraux apparaissent. La résolution spatiale accessible par cette technique ne donne pas accès à une résolution cellulaire ni à des indications quantitatives précises sur le nombre ou les types de cellules présentes dans les images. Les principales informations anatomiques fournies sont macroscopiques et concernent le volume des régions, l'épaisseur corticale, la géométrie des sillons, etc. (Balbastre et al., 2017). A travers ces mesures macroscopiques, il est possible, de façon indirecte, de suivre le développement ou la neurodégénérescence des neurones dans le cerveau. Le contraste des tissus dans des examens IRM du cerveau dépend de l'intensité du champ magnétique et des séquences mises en œuvre (T1, T2, diffusion, etc.). De manière globale, les contrastes obtenus par IRM permettent de discerner dans le cerveau la matière grise (population de cellules neuronales et gliales), la matière blanche (principalement composée des axones) et le liquide céphalo-rachidien.

1.3.1.2. *Imagerie par tomographie à émission de positons*

La **TEP** est une technique d'imagerie en médecine nucléaire qui permet de générer l'image tridimensionnelle fonctionnelle de la biodistribution d'un traceur radioactif fixé sur une molécule d'intérêt (par exemple avec la molécule de glucose) pour l'ensemble ou des parties du corps (Bailey et al., 2005).

Le métabolisme imagé par la TEP permet de mesurer l'activité des neurones. Lorsque les neurones sont activés, ils ont besoin de plus d'oxygène, de plus de glucose, etc. (Bear et al., 2016). Ces molécules peuvent être marquées avec des radionucléides de courte durée de demi-vie (de quelques minutes à quelques heures) appelés radiotraceurs (carbone, fluor, azote, oxygène, etc., Carlson, 2012) injectés par voie intraveineuse et qui vont se distribuer jusqu'au cerveau lorsqu'ils peuvent passer la barrière hémato-encéphalique. Les neurones les plus actifs qui utilisent normalement ces molécules (oxygène, glucose, etc.) vont capter spécifiquement ces radionucléides dont la désintégration au niveau des neurones est un index de l'activité métabolique neuronale. Ces radionucléides vont libérer des positons (des électrons chargés positivement), puis lorsqu'un positon rencontre un électron, ils vont s'annihiler. Il en résulte la production d'une paire de photons qui sont émis dans des directions opposées. Ces photons peuvent être détectés par des caméras très sensibles qui permettent alors de calculer le niveau d'activité en 3D dans chaque région cérébrale et de produire une image qui révèle les régions qui sont les plus actives dans le cerveau.

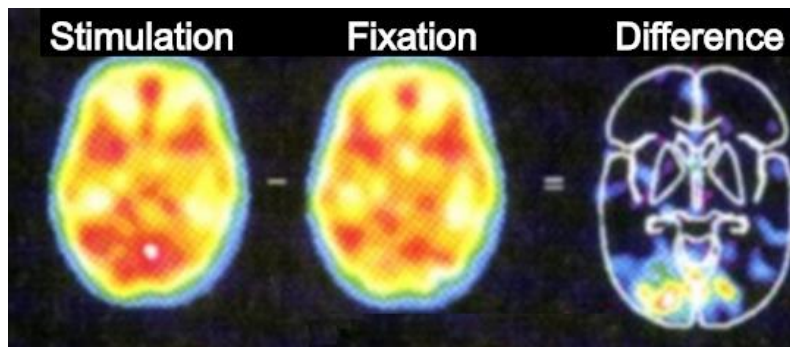


Figure 16 : Un exemple d'image TEP des régions cérébrales activées pendant qu'un sujet détaille un tableau (Bear et al., 2016).

La TEP est une technique d'imagerie capable de mettre en évidence le métabolisme des biomolécules et d'accéder au fonctionnement d'un cerveau qui est en train de penser, de planifier ou d'exécuter une action. Elle est actuellement très largement utilisée pour étudier l'activité du cerveau, diagnostiquer et suivre l'efficacité de traitements principalement en neurosciences (Lida et al., 2000 ; Logan et al., 1987 ; Hume et al., 1996). Cependant, l'acquisition de l'image TEP est relativement lente, de l'ordre de plusieurs dizaines de minutes pour obtenir un seul scan. De plus, la résolution spatiale reste faible, de l'ordre de quelque 5 à 10 mm³, ce qui représente l'activité de plusieurs milliers de cellules (Bear et al., 2016). Cette technique d'imagerie est reconnue pour être quantitative (possibilité de mesurer un nombre de désintégration absolu) et reflète l'état fonctionnel des neurones imagés. Dans le cadre de la MA, par exemple, l'acquisition d'images du métabolisme du glucose par Fluorodesoxyglucose (FDG) donne des indications sur la maladie car un hypométabolisme dans les cas pathologiques de la maladie a été démontré (Mosconi, 2005).

1.3.2. L'imagerie *post mortem*

Les techniques d'imagerie *in vivo* permettent aux scientifiques d'étudier le cerveau humain vivant et de voir sa structure à une échelle millimétrique mais il est impossible d'étudier la structure cellulaire des tissus avec ces techniques. C'est l'histologie ainsi que les techniques d'immunohistochimie qui, à l'aide des microscopes, rendent les études à l'échelle cellulaire possibles.

1.3.2.1. Principe de l'histologie

L'**histologie** est l'étude des tissus biologiques à l'échelle microscopique en s'intéressant à leur formation, leur organisation et à leur fonction. Elle constitue une branche importante tant de la biologie que de la médecine. Elle rend également possible le diagnostic, l'analyse de la pathologie, l'élucidation des mécanismes impliqués dans des maladies ainsi que l'évaluation de nouvelles approches thérapeutiques. L'histologie repose sur la production de coupes histologiques qui sont généralement obtenues au cours de six étapes : **extraction** de l'échantillon, **préparation**, **enrobage**, **coupe**, **coloration** et **numérisation**. Chaque étape peut conduire à la déformation de l'échantillon, elles doivent donc être effectuées avec le plus grand soin.

L'étude des structures anatomiques à l'œil nu en utilisant des équipements simples, comme le couteau et les ciseaux, commence dès le début du II^e siècle. On considère que l'histologie a été fondée au XVII^e siècle par Marcello Malpighi, qui étudiait de minuscules entités biologiques avec les premières techniques de microscopies. Avec l'amélioration des techniques de microscopie, de nombreux organes ont été explorés. En 1799, Xavier Bichat a proposé la notion de tissu biologique (Bichat, 1799). Il a proposé l'hypothèse selon laquelle le corps humain serait composé d'une vingtaine de tissus bien qu'il n'ait fait aucune observation par microscopie. En 1819, Mayer et

Heusinger ont utilisé pour la première fois le terme d'histologie. En 1838, Matthias Jakob Schleiden et Theodor Schwann ont proposé la théorie cellulaire en s'inspirant des travaux de Robert (Hooke, 1667), afin de clarifier le fait que la cellule constitue l'unité fondamentale de la structure et de la fonction des animaux et des végétaux. A partir de ce moment-là, les études basées sur des cellules évoluent progressivement vers la **cytologie**. Les spécificités de la structure et de la fonction des divers tissus sont liés à leur composition cellulaire ce qui rend primordial leur étude par l'histologie.. La recherche histologique fournit des images des organes normaux, tissus et cellules. Ces images histologiques sont la base nécessaire de l'histopathologie. La connaissance histologique des tissus normaux permet de mieux comprendre les anomalies dans les pathologies. En 1873, Camillo Golgi invente la technique de coloration qui rend visible les cellules nerveuses à l'aide d'une coloration au nitrate d'argent. Cette technique permet d'étudier la morphologie et la distribution des cellules nerveuses. Grâce à cette technique, Santiago Ramon y Cajal proposa sa théorie du neurone. A l'aide d'une autre coloration à l'argent utilisée par Max Bielschowsky, Aloïs Alzheimer découvrit les lésions de la MA qui sont les plaques amyloïdes et les dégénérescences neurofibrillaires de protéines Tau. Ces colorations sont appelées méthodes histochimiques.

En 1941, l'**immunohistochimie** (IHC) est proposée par Coons et al. (1941). L'IHC est une technique qui permet la localisation, l'analyse qualitative et semi-quantitative des antigènes dans les cellules d'une coupe de tissu grâce à une réaction antigène-anticorps. L'IHC combine la spécificité de la réaction immunologique et la visualisation par l'histochimie pour détecter divers antigènes (telles que les protéines, polypeptides, enzymes, hormones, etc.) au niveau cellulaire au moyen de l'imagerie et du grossissement des microscopes. La mise en évidence de la réaction antigènes-anticorps peut être réalisée de plusieurs manières. L'anticorps peut être conjugué à une enzyme, telle que la peroxydase, qui peut catalyser une réaction colorée. L'anticorps peut aussi être marqué grâce à un fluorochrome, tel que la fluorescéine ou la rhodamine, qui sont visibles à l'aide d'un microscope à **fluorescence**.

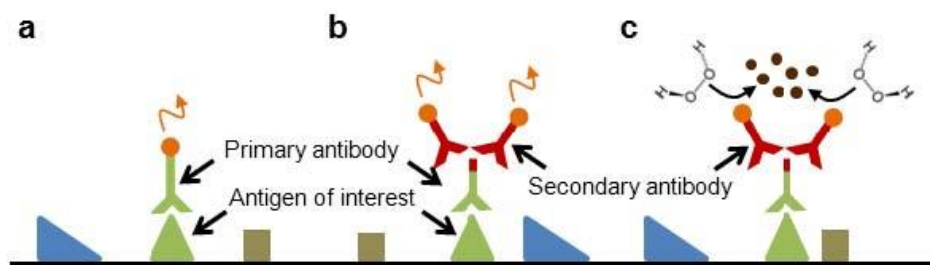


Figure 17 : Principe de l'immunohistochimie. a) Immunohistofluorescence directe. Le premier anticorps a pour le but de détecter l'antigène b-c) Immunohistochimie indirecte. Le deuxième anticorps reconnaît le premier anticorps et amplifie le signal (figure inspirée de la thèse de (Vandenberghe, 2015)).

Afin de détecter l'antigène cible, il existe deux méthodes : la méthode directe et la méthode indirecte (Figure 17). La méthode directe ne concerne qu'un seul anticorps, qui est utilisé pour détecter l'antigène qui émet de la lumière lorsque la fluorescéine est activée. Cette méthode est simple, rapide mais elle est sensible à la lumière. Dans le cas général, l'IHC a besoin de deux anticorps différents (méthode indirecte) pour détecter l'antigène. La méthode indirecte utilise un deuxième anticorps pour reconnaître le premier anticorps dans le but d'amplifier le signal. Ce deuxième anticorps est souvent lié à un colorant fluorescent ou à de la peroxydase de raifort qui génère un produit de teinte marron, par oxydation avec la diaminobenzidine (DAB) et le peroxyde d'hydrogène, qui permet de détecter la molécule d'intérêt. L'IHC est très performante et très utilisée dans la recherche au niveau du cerveau. Parce que chaque type de cellule exprime des protéines spécifiques, l'IHC peut être utilisée pour distinguer efficacement les différents types cellulaires (Figure 18). L'IHC

peut être également utilisée par les pathologistes pour tester la présence de biomarqueurs cliniquement importants dans les maladies, par exemple, les plaques amyloïdes et les dégénérescences neurofibrillaires de protéines Tau dans la MA.

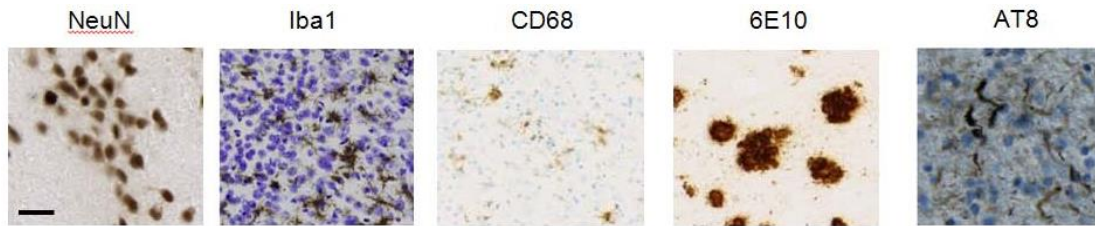


Figure 18 : IHC dans le cerveau. De gauche à droite : marquage neuronal par NeuN, marquage des cellules microgliales par Iba1 et contre-coloration par violet de crésyl (Nissl), marquage de cellules phagocytaires par CD68 et contre-coloration légère par Nissl, marquage de plaques A β par 6E10, marquage des enchevêtrements neurofibrillaires par AT8 et contre-coloration par Nissl. (Vandenberghe, 2015).

Dans le projet de ma thèse, nous nous intéressons tout particulièrement aux neurones, qui peuvent être marqués par l'anticorps NeuN (Neuronal Nuclei). En 1992, Mullen et al. (1992) ont observé l'expression de NeuN dans la plupart des neurones dans toutes les régions du cerveau, y compris le cortex cérébral, l'hippocampe, le thalamus, le noyau caudé, le putamen, etc. Il convient de mentionner que l'expression de NeuN n'est pas seulement observée dans le noyau mais aussi dans le cytoplasme qui est coloré dans une moindre mesure. Cependant, dans des cellules pyramidales du cortex cérébral, le noyau est intensément coloré à l'exception du nucléole, tandis que le cytoplasme n'est, quant à lui, que légèrement coloré (Figure 19). Nous allons par la suite étudier les neurones sur des coupes histologiques de cerveau macaque qui est un modèle plus proche du cerveau humain par rapport aux modèles murins.

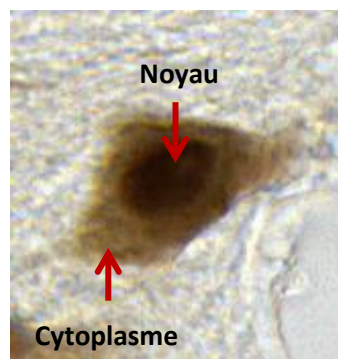


Figure 19 : Un neurone marqué par l'anticorps NeuN. Son noyau est coloré plus intensément que son cytoplasme.

1.3.2.2. Production de coupes histologiques

Pour produire les coupes histologiques, à MIRCen, le protocole général réalisé *post mortem* intègre le prélèvement des échantillons (ex. extraction du cerveau), la congélation ou la déshydratation, l'enrobage (milieu pour congélation ou paraffine) et la coupe. Après ces premières étapes, les coupes de tissus vont être colorées. L'analyse de ces coupes pourra être réalisée avec un logiciel soit directement connecté avec un microscope soit après numérisation. Les étapes suivantes décrivent plus en détail le protocole utilisé pour produire les données traitées dans cette thèse à savoir la production de coupes histologiques de cerveau de macaque marquées par NeuN pour étudier les neurones.

Extraction du cerveau

Après euthanasie de l'animal, les tissus sont fixés par perfusion intracardiaque de fixateur (paraformaldéhyde 4%). Son cerveau est ensuite extrait par dissection du crâne (Figure 20). L'extraction doit être réalisée avec le plus grand soin pour ne pas abîmer le cerveau.



Figure 20 : Vues d'un cerveau de macaque après extraction.

Congélation de l'échantillon

Une fois le cerveau extrait, il doit être coupé en fines tranches de quelques dizaines de micromètres d'épaisseur afin que les tissus puissent être observés au microscope. Pour réaliser ces coupes, le cerveau doit être congelé. Il est au préalable immergé dans une série de solutions graduées de saccharose et à froid pendant 4 jours afin d'enlever l'eau de l'échantillon et ainsi d'éviter l'apparition de cristaux qui pourraient détériorer les tissus lors de la congélation. A ce stade, la coupe de l'échantillon peut être mise en attente même si un traitement rapide permet de conserver une meilleure qualité des coupes à produire.

Enrobage de l'échantillon

Afin de fixer le prélèvement biologique sur son support et de réaliser les coupes du cerveau, l'échantillon congelé doit tout d'abord être enrobé par un produit spécifique, appelé milieu d'enrobage. Les milieux d'enrobage les plus utilisés pour la congélation sont l'OCT®, le Cryomatrix® et le Tissue-Tek®. Généralement, ces produits sont de couleur blanche après congélation. Comme des photographies sont acquises pendant le processus de coupe afin de pouvoir réaliser des reconstructions 3D à partir d'images 2D, un produit colorant vert (Fast Green, Sigma-Aldrich, Lyon, France) est ajouté au milieu d'enrobage (M1 embedding matrix, Shandon, Pittsburg, PA, USA) pour augmenter le contraste avec le tissu (Figure 21). Ce protocole a été optimisé au sein de notre laboratoire au cours des années et permet de faciliter certains traitements de l'image (segmentation couleur, reconstruction 3D).

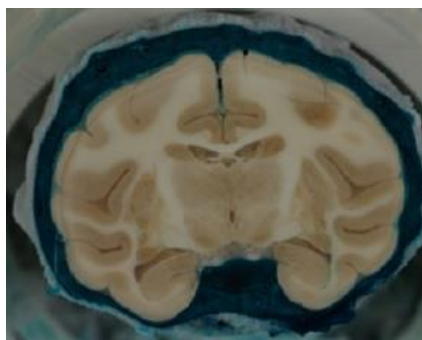


Figure 21 : Cerveau de macaque enrobé avec un milieu d'enrobage coloré en vert.

Coupe de l'échantillon

La coupe de cerveau de macaque est généralement réalisée selon une incidence coronale sur un microtome à glissière (Figure 22). Ce dernier est muni d'un système de refroidissement qui maintient le cerveau congelé (environ -20°C) et permet la réalisation de coupes d'épaisseur de l'ordre de $40\text{ }\mu\text{m}$ pour le macaque. Les coupes sont récupérées avec un pinceau et placées dans une solution tampon sous la forme de coupes « flottantes ». Cette étape de la coupe engendre la perte de la cohérence tridimensionnelle du cerveau et produit une série de coupes en 2D.

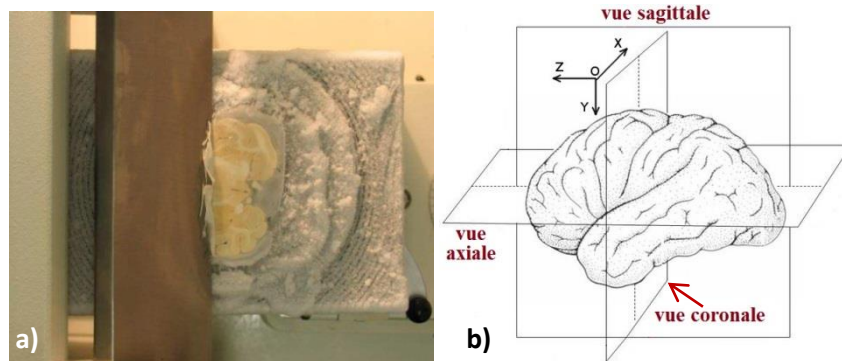


Figure 22 : a) Coupe d'un cerveau en cours selon une incidence coronale avec un microtome (Leica SM2400, Microsystems, Wetzlar, Germany). b) Schémas des trois plans de coupe possibles – vue axiale, sagittale et coronale de cerveau.

Coloration des tissus

Comme un cerveau de macaque mesure approximativement 5 cm de longueur selon la direction antéro-postérieure (perpendiculaire à l'incidence coronale), plusieurs milliers de coupes de $40\text{ }\mu\text{m}$ peuvent être produites. Les coupes successives et ordonnées obtenues lors de la coupe peuvent être divisées en plusieurs séries entrelacées ce qui rend possible la réalisation de plusieurs marquages d'intérêt sur un même cerveau (NeuN, Nissl, A β , Iba1, etc.). Ceci permet de marquer différents types d'objets (respectivement les neurones, les noyaux, les plaques amyloïdes, les cellules microgliales, etc.). Le nombre de séries dépend du nombre de marquages d'intérêt ou de l'échantillonnage jugé nécessaire pour répondre à la question biologique posée. Par exemple, pour N marquages il est possible de diviser le nombre total de coupes en N séries. La première série sera constituée des coupes 1, 1+N, et ainsi de suite. La taille du cerveau de macaque est plus grande que celle des rongeurs, une coupe colorée doit alors être montée sur une grande lame de verre (76 x 52 mm, double lame) pour la visualisation par microscopie. La coloration des coupes flottantes est réalisée de façon manuelle pour de nombreux marquages histologiques, ce qui peut entraîner une variabilité de la coloration. Les coupes de tissu colorées sont montées manuellement sur une lame de verre. En fin de traitement, une lamelle de verre est fixée sur la lame à l'aide d'un milieu de montage spécifique (sorte de colle) afin de protéger la coupe de tissu et de faciliter sa manipulation et sa conservation.

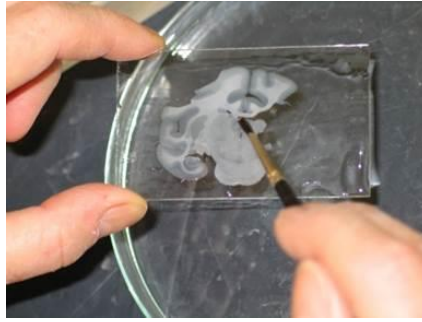


Figure 23 : Coupe de cerveau de macaque montée sur une double lame de verre.

1.3.2.3. *Imagerie des coupes histologiques*

Les coupes histologiques une fois produites sont soit observées directement sous un microscope soit numérisées de différentes manières en fonction des informations d'intérêt recherchées. Si une résolution de l'ordre de quelques micromètres à quelques dizaines de micromètres est nécessaire, l'utilisation d'un scanner à plat est suffisante. Par contre, si une résolution inférieure ou égale au micromètre est nécessaire, il est alors obligatoire d'utiliser un microscope ou un scanner de lame qui est proche par construction d'un microscope standard mais qui intègre en plus des mécanismes permettant d'automatiser la numérisation de série de coupes (déplacement des lames, numérisation de mosaïques, etc.). Pour numériser des coupes entières (WSI – Whole Slide Imaging), MIRCen s'est équipé d'un scanner de lames WSI Axio Scan Z.1 (Carl Zeiss Microscopy GmbH Jéna, Allemagne) (Figure 24). La plupart des grandes campagnes de numérisation de coupes à MIRCen sont réalisées avec ce système car il dispose d'un automate qui permet de lancer la numérisation de 50 grandes lames de verre au sein d'une même session. Ce système de microscopie automatisé fournit aux chercheurs des numérisations des coupes en champ clair et en fluorescence. Dans notre cas, l'Axio Scan Z.1 peut numériser des coupes aux grossissements x2.5, x5, x10, x20, x40 ce qui correspond à des résolutions de respectivement 1,76 μm , 0,88 μm , 0,44 μm , 0,22 μm et 0,11 μm en deux dimensions dans le plan d'acquisition. L'augmentation de la résolution augmente la durée de numérisation et la taille des images. En considérant la taille des neurones du cerveau de macaque, nous avons choisis un grossissement x20 qui est utilisé couramment dans ce type d'étude. Le temps de numérisation d'une coupe à la résolution de 0,22 μm est de quelques heures.



Figure 24 : Scanner de lames Axio Scan Z.1 à gauche, interface graphique du logiciel de microscopie à droite.

Les coupes numérisées sont ensuite gérées et traitées par le logiciel ZEN de Zeiss qui permet de piloter le scanner et de sauvegarder les images au format propriétaire CZI (Carl Zeiss Imaging). Il est également possible d'exporter des images de plus basse résolution aux formats TIFF, JPEG, PNG et GIF. Différentes stratégies d'acquisition sont possibles avec ce système : acquisition d'un plan unique (mise au point de la netteté sur un plan pré-calculé à partir de points de mesure), acquisition en plan unique avec l'algorithme *unsharp masking* pour augmenter la netteté de l'image, acquisition en pile

appelée « *Z-stack* » et acquisition en pile Z avec l'utilisation *a posteriori* d'un algorithme permettant de calculer la mise au point optimale entre le jeu de coupes acquises ce qui permet d'améliorer localement la netteté des images produites avec en contrepartie un temps d'acquisition, un volume de données et des temps de calcul beaucoup plus importants (profondeur de champ étendu: EDF, *extended depth of focus*)

Source:

<https://www.zeiss.com/microscopy/int/products/microscope-software/zen/image-acquisition.html>

Description des principaux modes de numérisation à disposition :

➤ **acquisition en plan unique** (Figure 25a)

C'est la méthode d'acquisition la plus simple basée sur une mosaïque de tuiles. L'épaisseur de coupe que nous avons utilisée est de 40 µm ce qui est standard pour l'histologie produite chez le primate. L'Axio Scan Z.1 fait tout d'abord une mise au point grossière, à un grossissement plus faible que celui utilisé pour la numérisation, selon une stratégie d'échantillonnage définie par l'utilisateur en fonction de la pièce de tissu à numériser (positionnement automatique de points de focus répartis sur l'ensemble du tissu, ici stratégie en « peau d'oignon » avec un maximum de 25 points de focus). Ceci permet de trouver un plan optimal dans le tissu qui servira de point de départ pour une mise au point plus fine. Cette mise au point fine est réalisée au grossissement utilisé pour la numérisation sur un échantillonnage adapté de tuiles (ici, 1 tuile sur 5, le plan focal optimal de la première tuile étant appliqué aux 4 tuiles suivantes). Ces deux étapes de mise au point permettent d'établir automatiquement des cartes de mise au point utilisées par le scanner pour la numérisation finale.

➤ **acquisition en plan unique avec l'algorithme *unsharp masking*** (Figure 25b)

Cette méthode applique la technique de filtrage passe-haut (masque *unsharp*) qui amplifie les informations de haute fréquence dans une image afin d'améliorer la netteté de l'image. L'image originale est d'abord dupliquée et ensuite filtrée par un noyau gaussien. Le masque *unsharp* est défini par la différence entre les valeurs de l'image originale et celles du duplicata filtré. L'addition du masque *unsharp* et de l'image originale construit la nouvelle image dans laquelle les détails (les informations de haute fréquence) sont renforcés.

➤ **acquisition en mode *Z-stack*** (Figure 25d-f)

L'acquisition en mode *Z-stack* permet d'acquérir des piles d'images axiales avec le microscope Zeiss selon la direction de la profondeur de la coupe. Le programme d'acquisition d'image commence par le plan central, ensuite autour de ce plan, des acquisitions sont réalisées de part et d'autre avec un pas d'échantillonnage fixé par l'opérateur (généralement de l'ordre de 1 micromètre). Pour chaque plan, une mise au point spécifique est réalisée, le nombre de coupes peut atteindre 20, le résultat est un pseudo-volume.

➤ **acquisition en pile Z combinée avec l'algorithme EDF** (Figure 25c)

Le logiciel ZEN peut traiter la pile de coupes acquises en mode *Z-stack* après la numérisation avec un algorithme de calcul d'une profondeur de champ étendue. Ce module extrait des détails nets sur un plan à partir d'images existantes à différentes positions de focalisation en Z et les combine pour créer une nouvelle image. Avec cette méthode, nous pouvons avoir une image dans un plan unique avec un focus optimal calculé à travers la profondeur numérisée de la coupe même pour les échantillons épais.

En utilisant le logiciel ZEN, de grandes données peuvent être numérisées et analysées. Il propose de plus un outil de visualisation performant qui permet de naviguer dans les images à différentes résolutions ce qui permet de zoomer facilement dans toutes les régions d'intérêt de la coupe du cerveau. Il permet également d'ajuster le contraste et d'accéder à des caractéristiques morphologiques basiques (diamètre, surface, etc.) des neurones.

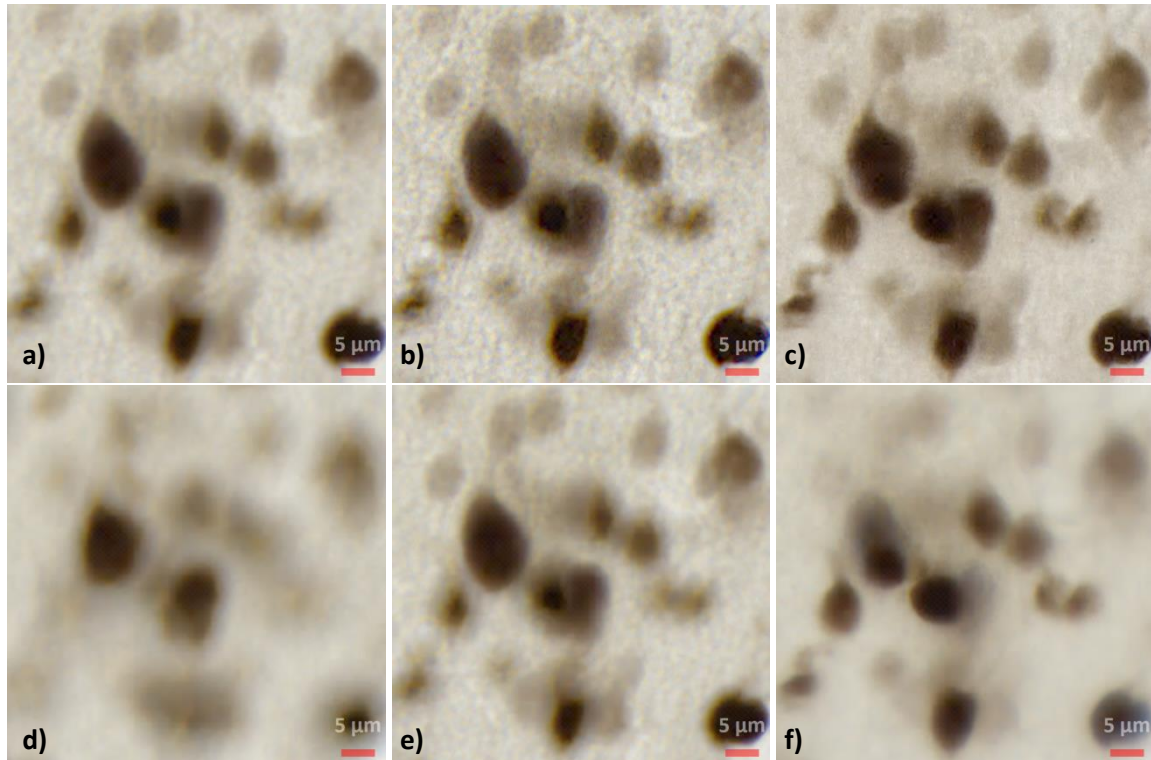


Figure 25 : Visualisation des images acquises, au grossissement x20, en plan unique (a), en plan unique avec l'algorithme *unsharp masking* (b), en pile Z puis traité avec l'algorithme d'EDF (c) et en pile Z (d-f). La coupe épaisse acquise en pile Z est divisée en 11 plans, les images montrées dans cet exemple sont les 1^{ère} (d), 6^{ème} (e) et 11^{ème} (f).

La résolution des images numérisées est de 0,22 µm/pixel dans le plan d'acquisition X/Y au grossissement x20. Pour mon travail, les coupes histologiques de cerveau de macaque ont été numérisées avec cette résolution, la coupe la plus grande correspond à peu près à 150 Go pour une dimension de centaines de milliers par des centaines de milliers pixels.

1.3.2.4. Analyse des coupes histologiques

L'image histologique numérisée à haute résolution spatiale nous offre la possibilité unique d'analyser les différentes régions anatomiques du cerveau à un niveau cellulaire. La Figure 26 montre une coupe histologique d'un cerveau de macaque sain colorée par NeuN. Dans cette vue globale, nous pouvons voir des régions anatomiques telles que le cortex, l'hippocampe, le noyau caudé, le claustrum, le putamen, le thalamus, les ventricules, etc. Visuellement, l'intensité dans le cortex cérébral est plus sombre par rapport à celle du thalamus. Ceci s'explique en raison de la densité et/ou de la forme variable des neurones. La densité du signal dans l'image donne une information relative sur ces derniers paramètres.

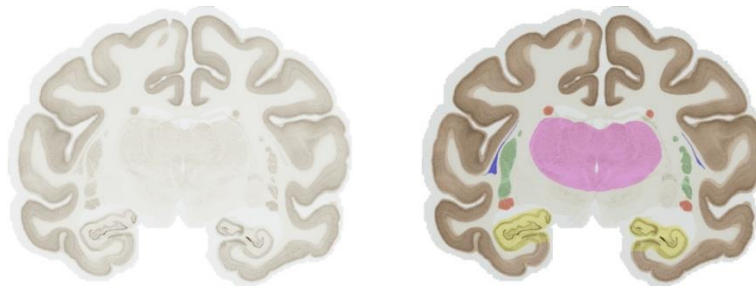


Figure 26 : Coupe histologique de cerveau de macaque sain marquée par NeuN (à gauche), les différentes régions anatomiques (à droite), marron - cortex cérébral, jaune - hippocampe, pourpre - thalamus, rouge - noyau caudé, vert - putamen, bleu - claustrum.

Comme introduit dans la partie 1.1.4, de nombreuses études relatives aux neurones ont déjà été réalisées. Les scientifiques s'intéressent au nombre, à la morphologie et à la distribution des neurones dans chaque région anatomique. Grâce aux images histologiques des neurones acquises à haute résolution par microscopie optique ces études, qui nécessitent une très haute résolution spatiale, peuvent être mises en œuvre. La Figure 27 montre les images de trois différentes régions anatomiques (thalamus, cortex cérébral, gyrus denté dans l'hippocampe) illustrant la variabilité allant d'une basse densité à une très haute densité de neurones. La segmentation des neurones sous la forme d'une classe de tissu est suffisante pour analyser les neurones dans les régions moins denses où les neurones sont déjà individualisés. Cependant, dans les régions de haute densité, les neurones peuvent être accolés voire se superposer les uns sur les autres. Cela augmente fortement la difficulté d'analyse individuelle des neurones. Si une étape de segmentation par classe n'est pas suffisante pour cela, une étape supplémentaire de post-traitement devient alors nécessaire pour individualiser les neurones. Une fois ce problème complexe réglé, un très grand nombre d'études relatives au comptage, à l'étude de la morphologie et à la distribution des neurones seront alors rendues possibles.

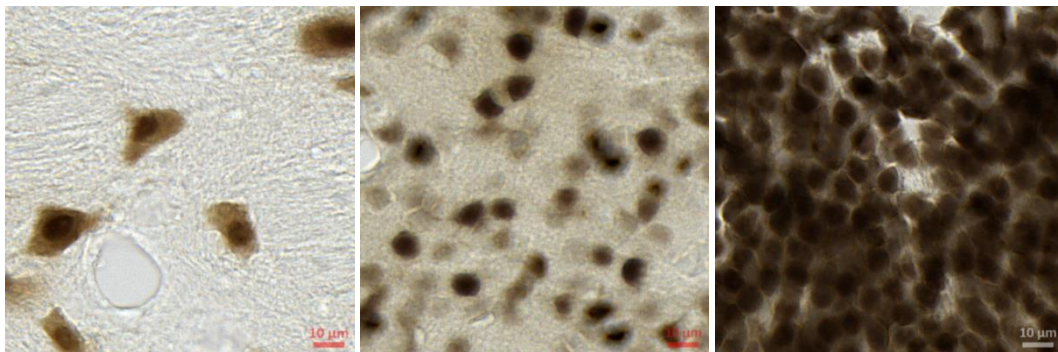


Figure 27 : Images de basse densité à haute densité de neurones dans un cerveau de macaque (thalamus à gauche, cortex cérébral au milieu, gyrus denté dans l'hippocampe à droite).

1.3.3. Complémentarité des techniques d'imagerie du cerveau

Beaucoup de techniques d'imagerie nous permettent d'acquérir des connaissances sur le cerveau humain cependant, les différentes techniques présentent naturellement des limites. L'IRM et la TEP sont deux techniques d'imagerie *in vivo* dont l'acquisition est relativement facile et rapide, elles présentent un intérêt pour les études précliniques car elles permettent un suivi longitudinal qui n'est pas possible par les analyses histologiques réalisées *post mortem*. Elles permettent de voir la structure directement en 3D ainsi que le fonctionnement du cerveau vivant. En contrepartie, la résolution spatiale reste faible, de l'ordre d'un à plusieurs millimètres pour la TEP. En 2015, Stucht et

al. ont acquis une image de cerveau humain par IRM avec une résolution de 0,44 mm. Cependant, la haute résolution de l'imagerie *in vivo* nécessite plus de temps pour l'acquisition et augmente les risques liés à l'anesthésie des animaux. Dans ces conditions, une étude massive des neurones n'est pas possible. L'histologie est réalisée en phase terminale dans la plupart des études précliniques du fait de l'information cruciale accessible par cette technique mais elle ne permet pas de présenter le cerveau en trois dimensions car le plus souvent, seules quelques coupes ou des séries de coupes sont produites (temps de traitement, volume d'anticorps à utiliser, etc.). L'acquisition des coupes de cerveau implique la réalisation d'une série d'étapes qui peuvent provoquer des déformations du tissu. L'extraction de cerveau entraîne d'abord la perte de liquide céphalo-rachidien qui remplit les ventricules cérébraux. Le cerveau est également fixé et déshydraté à ce stade. Toutes ces manipulations produisent un rétrécissement global du cerveau. L'étape de coupe de cerveau conduit à la perte de la structure 3D. Pendant la coloration et le montage, les coupes peuvent aussi être abîmées. Malgré ces contraintes, les très nombreux marquages histologiques disponibles aujourd'hui offrent la possibilité unique d'étudier finement la structure du cerveau et de mieux comprendre les mécanismes du cerveau à l'échelle cellulaire (jusqu'à quelques dixièmes de micromètres).

En résumé, l'histologie et la microscopie optique offrent la possibilité d'analyser les modèles expérimentaux avec une très grande variété de marquages histologiques. La technique de l'histologie permet d'accéder à une résolution spatiale très élevée jusqu'à l'échelle du micromètre. Cela permet d'étudier les organes à l'échelle cellulaire.

1.4. Problématique liée à l'étude des neurones en microscopie optique

Les maladies neurodégénératives présentées dans la partie 1.1.4.2 conduisent à la mort de nombreuses personnes dans le monde et restent pour la plupart incurables aujourd'hui encore. Des phénomènes communs à ces maladies sont l'atrophie du cerveau à l'échelle macroscopique et la perte de neurones au niveau microscopique. Pour mieux comprendre le fonctionnement des organes sains, ainsi que les mécanismes impliqués dans les maladies neurodégénératives, l'étude précise du nombre, de la morphologie et de la distribution des neurones dans chaque région anatomique du cerveau constitue un défi majeur à relever.

MIRCen possède une grande expertise dans le domaine des maladies neurodégénératives. Un certain nombre de laboratoires incluant histologie, biochimie, virologie, imagerie et une équipe traitement de l'image se consacrent à évaluer des thérapies innovantes à l'aide d'un ensemble de plateformes dédiées au développement de modèles animaux pertinents de pathologies humaines.

Dans ce contexte, mon projet de thèse consiste à étudier les neurones marqués par immunohistochimie à l'aide de l'anticorps NeuN sur des coupes histologiques de cerveau de macaque sain. Ces travaux sont basés sur des méthodes de traitement de l'image développées dans notre laboratoire dans le domaine de la segmentation de l'image. Toutefois, dans ce contexte, il subsiste un nombre important de problèmes à résoudre :

1. Un cerveau entier de macaque est très grand (52360 mm^3 en moyenne pour les macaques) comparé à celui de rongeurs (994 mm^3 en moyenne pour les rats) (Chareyron et al., 2011). De plus, les régions anatomiques peuvent être nombreuses sur une seule coupe. Les différents types de neurones ainsi que leur distribution sont très variables. Cela se traduit par une **grande variabilité de l'intensité de coloration des neurones** (Figure 28) rendant difficile la séparation des neurones marqués, du tissu non marqué.

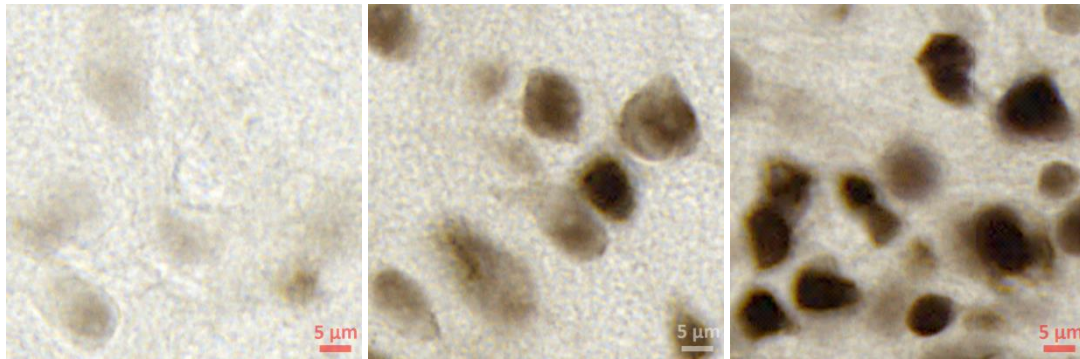


Figure 28 : De gauche à droite, neurones dans le thalamus, le noyau caudé et le cortex. Les intensités de marquage de neurones (en marron) sont variables.

2. Les coupes de cerveau de macaque sont épaisses, dans notre cas 40 µm par coupe. Cette épaisseur est choisie car elle assure une résistance mécanique suffisante pour traiter ces larges coupes flottantes. De plus, cette valeur correspond à l'ordre de grandeur de la taille des plus grands neurones ce qui a été défini par les biologistes comme étant une valeur adaptée pour réaliser des comptages, notamment en stéréologie. Cependant, la production d'une image de projection de cette coupe lors de la numérisation entraîne un certain nombre d'effets négatifs qui peuvent aller d'un **accolement de plusieurs cellules neuronales jusqu'à une superposition totale** avec tout un ensemble de cas intermédiaires (Figure 29). Ces effets peuvent compromettre le comptage et l'analyse des neurones.

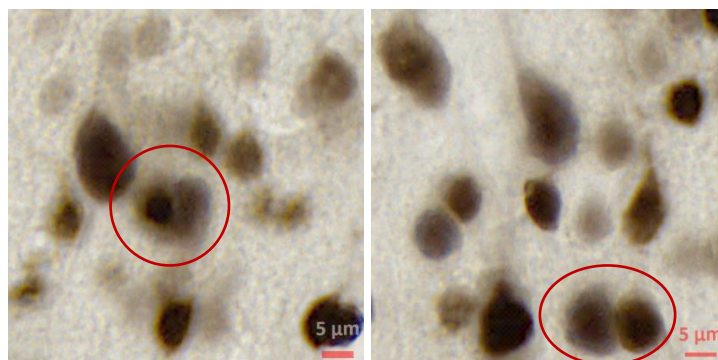


Figure 29 : Cellules superposées à gauche et cellules accolées à droite.

3. **La taille des neurones est très variable** et couvre un domaine allant de 5 à 30 µm de diamètre. Pour les individualiser, une méthode adaptée à la taille des neurones est requise. La Figure 30 montre des neurones juxtaposés de différentes tailles.

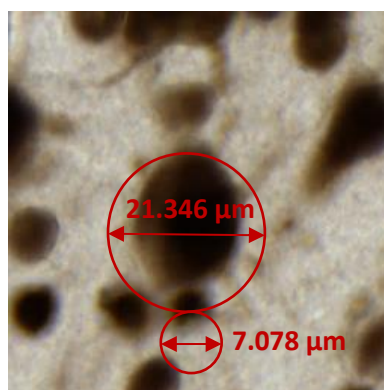


Figure 30 : Image montrant des neurones de différentes tailles accolés, le diamètre du petit neurone est à peu près de 7,078 µm, celui du grand neurone est d'environ 21,346 µm.

4. Classiquement, la technique de la stéréologie qui nécessite de travailler en *Z-stack* au microscope est utilisée par les biologistes pour estimer de manière non biaisée le nombre de neurones dans la région anatomique étudiée. Pour cela, les biologistes se basent sur des stratégies de comptage manuel dans un nombre adapté de sous-régions échantillonnées pour permettre des estimations robustes (une description de cette méthode de référence sera présentée en détail dans la partie 2.1). Cependant, **cette méthode d'analyse stéréologique manuelle est longue et très fastidieuse** ce qui limite la quantité de tissus qui peut être étudiée par cette méthode. L'information obtenue est une estimation à l'échelle de la région et ne donne aucune indication locale sur le niveau d'hétérogénéité de la structure étudiée. De plus, certaines régions anatomiques sont composées d'amas de neurones très denses et de ce fait extrêmement difficiles à individualiser (ex. gyrus denté dans l'hippocampe dans la Figure 31). Dans ce cas précis, un comptage par les biologistes devient quasiment impossible à réaliser. Enfin, un traitement exhaustif de coupes entières voire de cerveaux entiers serait également quasiment impossible à réaliser car il faudrait traiter des téraoctets de données correspondant à des dizaines de millions de neurones et prendrait un temps considérable incompatible avec la durée des études précliniques.

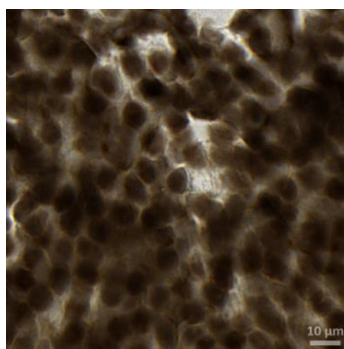


Figure 31 : Neurones dans le gyrus denté de l'hippocampe de cerveau de macaque.

5. Dans le cadre de mon projet de thèse, nous avons eu accès à un jeu de données de 134 coupes de cerveau d'un macaque, la taille d'une coupe numérisée de meilleure qualité (résolution 0,22 $\mu\text{m}/\text{pixel}$) pouvant atteindre jusqu'à 150 Go. Le temps de calcul, la possibilité d'accéder à ces données massives et la capacité de traitement d'une si grande image représentent de réels **défis en terme de capacité matérielle (mémoire vive, réseau, stockage, puissance de calcul, etc.) mais également en terme de développement logiciel (accès aux données, algorithmes, etc.)**.
6. **La plupart des études réalisées dans le domaine de l'individualisation des cellules se sont concentrées sur des régions d'intérêt de petites dimensions** ce qui limite fortement la complexité du problème et n'assure pas la généralité des solutions proposées si l'on vient à s'intéresser à des coupes entières ou à des cerveaux entiers. En effet, dans ces derniers cas, la quantité d'information à traiter explose littéralement et peut atteindre plusieurs centaines de gigaoctets jusqu'à quelques téraoctets de données. L'analyse de ces données massives devient rapidement problématique. De plus, réaliser un développement algorithmique sur des régions spécifiques limitées ne permet pas d'évaluer la capacité de la solution mise en œuvre à être déployée à très grande échelle ni à prendre en compte la diversité des configurations cellulaires qui peuvent être rencontrées dans des données très massives.

Compte tenu de l'ensemble des problèmes et des limitations que nous avons identifiés et qui sont décrits ci-dessus, les objectifs de mon projet de thèse sont les suivants :

- Segmenter la classe de tissu correspondant aux neurones marqués par l'anticorps NeuN sur des coupes de cerveau de macaque de manière aussi précise et rapide que possible.
- Proposer une méthode automatique afin de distinguer les neurones individuels et de séparer les

neurones qui se touchent.

- Proposer une méthode automatique pouvant tenir compte de la taille variable des neurones.
- Compter les neurones dans les principales régions anatomiques du cerveau pour tester la robustesse de la méthode proposée.
- Calculer une série de descripteurs des neurones (colorimétriques, morphologiques, etc.).
- Créer des cartographies paramétriques des descripteurs allant possiblement de l'échelle cellulaire à l'échelle mésoscopique afin de synthétiser le nombre, la morphologie et la distribution des neurones individualisés.
- Créer des bases de validation utilisant des informations produites par des experts (méthodologistes, biologistes) pour valider les différentes étapes du protocole d'analyse des neurones développées à l'occasion de cette thèse.
- Adapter notre environnement logiciel de travail, BrainVISA (<http://brainvisa.info>), afin de pouvoir traiter de très grands volumes de données (gestion des accès partiels à intégrer dans les codes).
- Réduire les temps de calcul. L'optimisation et la parallélisation de l'algorithme est indispensable pour rendre possible l'analyse de cerveaux entiers qui se composent de dizaines voire de centaines de coupes histologiques de cerveau.

2. Méthode de segmentation des neurones

2.1. Etat de l'art

Les informations relatives aux neurones (nombre, morphologie et distribution) sont cruciales dans la recherche biomédicale en particulier dans les neurosciences. Le nombre de neurones est un élément essentiel pour l'étude de la neurogenèse ou de la neurodégénérescence car les neurones perdent leur capacité à proliférer en vieillissant. De plus, les différentes régions anatomiques du cerveau peuvent être caractérisées par la morphologie et la distribution particulière des neurones. Utiliser les informations sur les neurones aide non seulement à étudier le développement du cerveau et la cyto-architecture du système nerveux central dans les cerveaux sains, mais aussi à étudier les mécanismes impliqués dans les maladies neurodégénératives. Par conséquent, une analyse quantitative fine pourrait aider les chercheurs à mieux comprendre la structure et le fonctionnement du cerveau et aider les pathologistes pour le diagnostic de la maladie. Afin de pouvoir analyser les neurones dans toutes les régions anatomiques, voire dans le cerveau entier, une méthode automatique, précise et robuste est nécessaire. Les étapes cruciales de la mise au point d'une telle méthode sont la segmentation et l'individualisation des neurones. Ces étapes sont assez bien décrites dans la littérature. Nous allons présenter des travaux existants relatifs à trois aspects importants qui sont à considérer : la **segmentation** et l'**individualisation** de cellules ainsi que les **méthodes de validation** qui leur ont été associées.

2.1.1. Méthodes de segmentation

- Seuillage

Le seuillage est la méthode de segmentation la plus simple qui convertit une image en niveaux de gris en une image binaire en fonction d'un seuil. Les pixels dont les intensités sont plus grandes qu'un seuil sont fixés à la valeur 1, les autres sont fixés à la valeur 0. Le seuil est généralement évalué par des approches paramétriques ou non paramétriques qui s'appuient sur l'histogramme de l'image. Pour les méthodes paramétriques, les classes d'objet et du fond sont supposées suivre une distribution gaussienne. Ces méthodes cherchent à estimer les paramètres de la distribution gaussienne (Fan et Lin, 2007; Bazi et al., 2007; Wang et al., 2008). Parmi les méthodes non paramétriques, la méthode d'Otsu (Otsu, 1979) est un standard. Le seuil est calculé automatiquement en minimisant la variance intra-classe. Les méthodes comme celles d'Otsu estiment des seuils globaux sur une image entière et la segmentent selon deux classes: l'objet et le fond. Souvent, la luminance d'une image n'est pas uniforme et présente des valeurs variables, c'est ce qui résulte fréquemment de la présence d'objets foncés, d'objets clairs et du fond. Par conséquent, le seuillage adaptatif a été proposé pour s'adapter aux régions de différentes luminances. Le seuillage adaptatif est généralement calculé dans une fenêtre glissante de l'image ou dans les sous-classes segmentées de façon itérative (Shu et al., 2013; Cai et al., 2014).

La méthode de seuillage permet de segmenter des objets par rapport au fond mais ne peut pas séparer les objets s'ils appartiennent à une même composante connexe. Lorsque les objets ne se touchent pas, l'étape de seuillage seule est suffisante. Cependant, si les objets se touchent, il faut une étape supplémentaire pour les séparer.

- K-moyennes

La méthode des K-moyennes (Hartigan et Wong, 1979) est une méthode de classification mise au point pour analyser les données dans le domaine du *data mining*. Etant donné les points initiaux d'un nombre k de classes, chaque point peut être attribué à l'une des k classes de façon à minimiser la

distance de chaque point au centroïde de sa classe. Il est ensuite nécessaire de mettre à jour les centroïdes des k classes et de partitionner à nouveau les points du plan selon k classes en utilisant la même méthode. La méthode itère sur cette étape jusqu'à ce que plus aucun point ne soit changé de classe. En traitement des images, ce sont les pixels de l'image qui sont répartis selon k classes et la distance utilisée est la distance calculée entre l'intensité de chaque pixel et la moyenne des intensités des pixels de chaque classe. Les k classes correspondent aux objets et au fond.

Comme la méthode du seuillage, cette méthode ne peut extraire que les objets du fond. Au sein d'une même classe, pour séparer les objets qui se touchent, une étape supplémentaire est nécessaire.

- Apprentissage / classification

Au cours des dernières années, l'émergence de techniques d'apprentissage a permis à plusieurs méthodes d'analyse cellulaire de voir le jour. La méthode bioVision (Chubb et al., 2006) est une des méthodes d'apprentissage supervisées qui utilise des caractéristiques concernant la colorimétrie et les informations sur l'intensité locale (la moyenne et la variance des valeurs d'intensité dans le voisinage de chaque pixel). Pendant l'étape d'apprentissage, bioVision apprend ces caractéristiques en fonction d'une vérité terrain produite au préalable par un expert. Ensuite, pour chaque classe, la méthode modélise la densité de probabilité des caractéristiques par un mélange de gaussiennes. Pendant l'étape de classification, chaque pixel est classifié, en se basant sur le théorème de Bayes, par estimation des probabilités *a posteriori* d'appartenir à chacune des classes. Cette méthode a été implémentée dans notre laboratoire et a été testée pour la segmentation de la classe neurone dans ce travail. La méthode des forêts aléatoires (*Random Forest*, RF) est une autre technique d'apprentissage qui a montré une amélioration, par rapport à bioVision, de la qualité des segmentations automatiques obtenues sur des plaques A β dans le cas pathologique de la MA (Vandenberghe et al., 2015). Cette méthode sera présentée en détails dans la section 2.3.

Les techniques d'apprentissage/classification permettent généralement de fournir de meilleures performances que les autres méthodes de segmentation. Cependant elles nécessitent la constitution d'une base d'apprentissage segmentée manuellement par des experts ce qui nécessite un travail long et fastidieux.

Les méthodes décrites ci-dessus ont pour but de segmenter les classes objets par rapport au fond. Lorsque les objets se touchent, il est cependant impossible d'analyser les objets individuels avec ces méthodes par un simple calcul de composantes connexes. Dans le cas d'images histologiques, il est fréquent que de nombreux objets segmentés ne soient pas individualisés. Par conséquent, des méthodes complémentaires, permettant d'individualiser les objets présenteraient un grand intérêt. Dans le chapitre suivant, nous allons présenter des méthodes qui traitent du problème de l'individualisation d'objets.

- Morphologie mathématique

La méthode de la morphologie mathématique travaille généralement sur des images binaires. Elle est fondée sur les relations logiques entre les pixels. L'image considérée est traitée comme un ensemble de pixels à l'aide d'un élément structurant. Un élément structurant a une forme définie, classiquement des croix, des carrés et des disques en 2D, et une amplitude qui autorise l'exploration d'un espace donné autour du pixel traité. Les opérations fondamentales de base en morphologie mathématique sont l'érosion et la dilatation. Avec les opérations appropriées, d'ouverture (érosion puis dilatation) et de fermeture (dilatation puis érosion), il est possible de supprimer des petits objets et de remplir les trous dans une image binaire.

Cette méthode (Nedzved et al., 2000; Shu et al., 2013) peut être mise en œuvre pour séparer les objets qui se touchent à l'aide de l'opération d'ouverture ou en utilisant les résidus ultimes produits pendant les étapes d'érosion. Cette méthode ne fonctionne que pour de faibles densités d'objets. Lorsque les objets se trouvent au milieu d'un amas d'objets très compact, la méthode n'est pas adaptée.

- Détection de concavité

Cette approche permet de détecter des points concaves sur le contour des objets accolés. Les points concaves peuvent être déterminés en fonction de différents critères, par exemple, la courbure, le nombre de pixels des objets en fonction du nombre total de pixels dans une fenêtre ayant généralement la forme d'un carré, etc. Ensuite, à l'aide d'un recalage d'ellipse ou d'un algorithme de transformation de distance, les objets qui se touchent peuvent être séparés selon une séparation optimale (Bai et al., 2008; Kothari et al., 2009; Zhang et al., 2013; Qi, 2014).

Cette méthode peut séparer des objets connexes. Cependant, il existe souvent de faux points concaves dus au bruit qui existent dans l'image binaire. De plus, il n'est pas possible d'utiliser cette méthode dans notre cadre d'application car les points concaves sont présents sur le contour externe d'un amas d'objets. Lorsque la densité d'objets augmente, il n'y a plus de points concaves pour identifier les objets qui se présentent au centre de cet amas et il devient impossible d'individualiser les objets (Figure 31).

- Croissance de région

La croissance de région est une méthode de segmentation d'image qui permet de partitionner une image sous forme de régions (Zucker, 1976; Adams et Bischof, 1994). Elle nécessite un ensemble de points (graines) pour initialiser les régions explorées lors du processus de croissance. Ces graines sont déterminées selon un critère défini par l'utilisateur, par exemple, les pixels qui ont les intensités extrêmes (minima/maxima). Les régions croissent alors itérativement par association autour des graines de tous les pixels voisins qui possèdent un ensemble d'attributs de valeurs semblables (couleur, texture, forme, taille, homogénéité, etc.). A chaque itération, les nouveaux points associés servent à leur tour de graines. Les itérations se poursuivent jusqu'à ce que l'ensemble des points aient été affectés à une région.

La clé de la réussite de cette méthode repose au départ sur une sélection appropriée des graines. Le manque de graines conduit à la sous-segmentation, tandis que de la présence de graines supplémentaires produit une sur-segmentation.

- *Watershed*

La méthode du *Watershed* (ligne de partage des eaux) est une transformation définie sur une image en niveaux de gris. L'image est considérée comme une surface topographique. Les pixels d'intensité basse deviennent des bassins tandis que les pixels d'intensité élevée deviennent des élévations. L'objectif est de séparer les bassins. L'idée principale est d'inonder les bassins avec de l'eau. Au cours de la montée de l'eau, les bassins adjacents vont se rejoindre. A ce moment-là, une ligne de partage des eaux est marquée comme frontière entre les deux bassins. Cousty et al. (2009) ont étudié le *Watershed* par l'utilisation de gouttes d'eau s'écoulant sur une surface topographique. Les lignes de séparation sont un ensemble de points à partir desquels une goutte d'eau peut s'écouler vers au moins deux minima régionaux.

Cette technique est très efficace pour certaines applications. Cependant, dans les régions denses où il existe de grandes quantités d'objets connexes, cette technique reste applicable mais génère

souvent des sur- et sous-segmentations. Afin de limiter la sur-segmentation, Vincent et Soille (1991) ont proposé d'utiliser les minima régionaux comme des graines, pour initialiser le *Watershed*. Yang et al. (2006) ont utilisé un *Watershed* contrôlé par des résidus ultimes extraits en fonction de deux étapes d'érosion dont les tailles d'élément structurant sont différentes. L'élément structurant de plus grande taille est utilisé d'abord pour réduire les objets de taille supérieure à un seuillage T_1 ensuite l'élément structurant de petite taille est utilisé pour garder l'information concernant la forme des objets. Shu et al. (2013) ont calculé les graines en utilisant un filtrage morphologique et une croissance de région basée sur l'intensité. Toutefois, à cause de la présence de bruit et du faible gradient d'intensité dans les images, il est fréquent que plus d'une graine soit détectée par objet ou qu'il manque la graine nécessaire à la segmentation d'un objet. Par conséquent, la précision de la détection des graines devient le point le plus important pour obtenir une bonne qualité de segmentation. Une fois que les graines sont bien détectées, les cellules pourront être séparées en utilisant le *Watershed*.

- **Modèle de contour actif déformable**

Cette approche est une technique de segmentation basée sur un modèle déformable qui permet d'extraire la bordure d'un objet (Kass et al., 1988). Un contour initial est dessiné (ou généré automatiquement) autour d'un objet d'intérêt puis, il va être déformé automatiquement jusqu'à ce qu'il converge vers la bordure de l'objet. Cette déformation est dirigée par deux énergies: l'énergie interne et l'énergie externe. L'énergie interne contrôle la morphologie du contour selon un critère de forme, le degré de lissage, la continuité et la longueur du contour. L'énergie externe est une combinaison de forces qui résultent de l'image elle-même (éléments caractéristiques de l'image comme les points de forts gradients, de fortes concavités, etc.) et de forces de contraintes introduites par l'utilisateur. Plus ce contour est proche de la bordure de l'objet, plus la valeur de l'énergie totale de ce modèle est faible. Le but est de minimiser cette énergie totale.

Le modèle présenté ci-dessus est le modèle par défaut, des solutions alternatives existent. Xu et Prince (1997) ont proposé un modèle GVF (*Gradient Vector Flow*) afin de régler les problèmes de modèle par défaut associés à l'initialisation et à la convergence faible vers les frontières concaves. Li et al., (2008) ont proposé de segmenter les noyaux cellulaires en se basant sur le suivi du flux de gradient (*gradient flow tracking*). En générant un champ de vecteur à gradient diffus, ils effectuent une procédure de suivi du flux de gradient pour attirer des points dans le bassin. Les points dans les mêmes bassins sont regroupés dans les mêmes groupes correspondant aux mêmes noyaux. A la fin, un seuillage adaptatif local est appliqué afin d'isoler les noyaux du fond.

- **Graph-cut**

La méthode de *graph-cut* se réfère à une famille d'algorithmes dans lesquels la segmentation de l'image est considérée comme un problème de partition de graphe. Une image est considérée sous la forme d'un graphe pondéré non orienté $G(V, E)$. V représente les pixels de l'image et E représente les arêtes pondérées entre les pixels. Chaque arête pondérée représente la similarité entre deux pixels voisins, elle est généralement mesurée en fonction de l'intensité, de la localisation spatiale, des points concaves mais peut être également calculée en fonction d'autres critères. L'objectif est de trouver des séparations qui partitionnent le graphe en sous-graphes afin de maximiser les connexions au sein de chaque sous-graphe et de minimiser les connexions entre les différents sous-graphes. Les atouts de cette méthode sont sa robustesse au bruit, l'intégration d'informations visuelles et de contraintes topologiques. De ce fait, cette méthode est très populaire actuellement.

Daněk et al. (2009) et Lou et al. (2012) ont individualisé des cellules connexes en utilisant la méthode de *graph-cut* en combinant des informations *a priori* de gradient et/ou de forme de cellules. Al-Kofahi et al. (2010) ont, en premier lieu, séparé les cellules du fond en utilisant la méthode de *graph-cut*. Puis, ils ont détecté les centroïdes des cellules à l'aide de filtre LoG (*Laplacian-of-Gaussian*) et de cartes de distance. Cependant, cette méthode ne résout pas les problèmes de sur- et sous-segmentations même dans les cas simples où seulement quelques cellules sont accolées. En 2015, He et al. ont proposé l'algorithme *iCut* pour individualiser des cellules accolées. Une matrice de pondération est calculée en fonction de la taille de cellules, du contour de l'objet, de la position des pixels de l'objet et de ses points concaves. L'algorithme de *graph-cut* est ensuite appliqué aux valeurs de connexions normalisées par la matrice de pondération. Cette approche n'est cependant pas suffisante pour individualiser des objets superposés dans des régions très denses. La segmentation finale dans la région de forte densité de cellules est homogène en raison de la taille de cellule fixée *a priori* et de l'absence de points concaves dans ces régions. De plus, cette méthode n'est pas adaptée pour gérer des cellules dont la taille varie de façon importante. Dans la littérature, la méthode de *graph-cut* a été uniquement utilisée pour des cas simples où il n'y a pas d'amas des cellules accolées. Pour les cas très complexes, le problème reste entier.

Les méthodes présentées ci-dessus permettent d'individualiser des objets accolés dans différents domaines d'application. Dans la partie suivante, nous allons présenter les méthodes de validation qui ont été proposées et qui permettent d'évaluer la qualité de ces stratégies d'individualisation.

2.1.2. Validation des méthodes de segmentation d'objets

- Comptage d'objets

Pour estimer la performance d'une méthode automatique, He et al. (2015) ont calculé l'exactitude de la méthode en terme de nombre de cellules correctement segmentées. Dans leur évaluation, une cellule segmentée automatiquement correspond à une cellule marquée manuellement par un expert si la distance entre les centroïdes de ces deux cellules est inférieure à un rayon de cellule R prédéfini. Alors, les cellules segmentées automatiquement et marquées manuellement forment une paire unique. Par la suite, le nombre de cellules marquées manuellement sera noté N_g , le nombre de cellules segmentées automatiquement sera noté N_d et le nombre de cellules segmentées qui font partie d'une paire unique sera noté N_m . La performance de la méthode est calculée selon trois critères : le rappel (R), la précision (P), et le *F-score* (F) donnés par les équations (1).

$$R = \frac{N_m}{N_d}; \quad P = \frac{N_m}{N_g}; \quad F = 2 \times \frac{R \times P}{R + P} \quad (1)$$

De nombreux chercheurs (Al-Kofahi et al., 2010; Shu et al., 2013; Zhang et al., 2013; Poulain et al., 2015; Kainz et al., 2015; Molnar et al., 2016) ont évalué la performance de leur méthode en utilisant les mêmes critères. En plus de ces trois critères, Kothari et al. (2009) ont également évalué le pourcentage d'erreur, donné par l'équation (2).

$$\mathcal{E} = \frac{N_d - N_g}{N_g} \times 100\% \quad (2)$$

- Stéréologie

Dans le domaine des neurosciences, la stéréologie est un outil très largement utilisé pour estimer des informations quantitatives telles que le nombre d'objets, leur longueur ou leur volume. Ces mesures sont généralement réalisées pour une région anatomique tridimensionnelle selon une procédure d'échantillonnage non biaisée à partir de sections transversales de tissus espacées régulièrement. La procédure d'échantillonnage systématique doit être soigneusement définie à l'avance afin d'éviter les biais. A un échantillonnage donné correspond une précision de la mesure qui peut être estimée. De très nombreuses études ont été réalisées en utilisant la stéréologie pour caractériser le nombre de neurones dans les cerveaux de plusieurs espèces animales (Gundersen, 1986; West et al., 1991; Mayhew, 1992; Jelsing et al., 2006). Etonnamment, nous n'avons pas trouvé, dans la littérature, de validation des méthodes de segmentation ou d'individualisation d'objets basée sur la stéréologie. Cette technique étant une méthode de référence pour la caractérisation et le comptage d'objets biologiques, il nous a paru très pertinent de la considérer pour valider notre méthode d'individualisation des neurones.

La Figure 32 illustre un schéma d'échantillonnage systématique d'une coupe de l'hippocampe de rat (West et al., 1991).

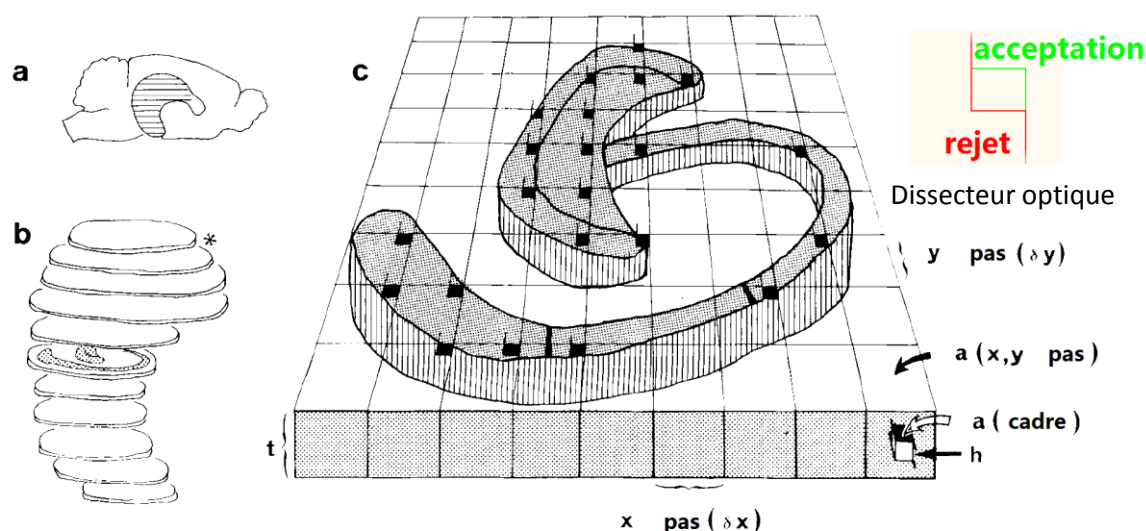


Figure 32 : Schéma d'échantillonnage systématique. (a) L'hippocampe de rat est coupé dans le plan horizontal. (b) La première section de la série à analyser (astérisque *) est choisie aléatoirement parmi les 12 premières sections. Cette section et chaque 12ème section par la suite est colorée pour l'analyse suivante. La fraction de section d'échantillonnage, ssf , est donc $1/12$. (c) Les coupes histologiques sont étudiées grâce à un microscope couplé à une caméra et à un système de déplacement pas à pas en x , y et t . A l'image de la coupe histologique est superposée une trame de comptage définie par des dissecteurs optiques de taille fixe et de distribution régulière. La fraction de surface d'échantillonnage, $asf = a(cadre)/a(x,y,pas)$, est la surface de dissecteur optique par rapport à la surface de l'élément de la trame où est placé le dissecteur. La fraction d'épaisseur d'échantillonnage, hsf , est la hauteur h de l'échantillon du dissecteur par rapport à l'épaisseur de la section, h/t . Le nombre de neurones dans une section peut être calculé selon l'équation 3.

Dans ce cas, l'hippocampe de rat est coupé dans le plan horizontal. Les coupes obtenues sont divisées sous la forme de séries entrelacées (s) qui peuvent être colorées par différents marquages pour les différentes analyses biologiques à mener. Ensuite, dans les coupes contenant la région hippocampique, une coupe de la première série est choisie aléatoirement. Cette coupe et chaque $s^{ème}$ coupe sont ensuite sélectionnées pour l'analyse suivante. La fraction de section d'échantillonnage, ssf , est donc de $1/s$. Dans chacune des coupes, la région à étudier est segmentée

manuellement et à l'intérieur de celle-ci, des dissecteurs optiques (sondes 3D invariantes par la taille, la forme ou l'orientation) sont générés avec une taille et des pas x, y prédéfinis. Ensuite, les neurones qui se trouvent à l'intérieur de dissecteur optique (cadre de comptage) sont comptés avec une règle d'exclusion de tous les neurones touchant les lignes rouges (Figure 32). La fraction de surface d'échantillonnage, asf , est donc la surface du cadre de comptage, $a(cadre)$, par rapport à la surface de l'élément de la trame où est placé le dissecteur, $a(x,y,pas)$. La fraction d'épaisseur d'échantillonnage, hsf , correspond à l'épaisseur du dissecteur h , divisée par l'épaisseur de la coupe t . Le nombre de neurones dans la région d'intérêt de la série correspondante est estimé par l'équation 3.

$$N_{total} = \frac{1}{ssf} \times \frac{1}{asf} \times \frac{1}{hsf} \times \sum Q^- \quad (3)$$

Où $\sum Q^-$ est le nombre total de neurones effectivement comptés dans les dissecteurs.

Dans la section 2.4.6.3, nous allons utiliser la stéréologie afin de valider notre méthode d'individualisation et de comptage des neurones. Grâce à la stéréologie, il est possible d'estimer de manière non biaisée le nombre de neurones dans une structure anatomique ce qui donnera un critère objectif pour évaluer notre méthode de comptage.

Pour conclure cette partie, force est de constater que les critères de validation des méthodes d'individualisation sont relativement limités voire même pourraient être mis en défaut par exemple si l'on venait à prendre en compte les contours ou la forme réelle des objets définis après individualisation. Ce point sera abordé plus précisément dans la suite de notre travail.

2.2. Données biologiques de l'étude

2.2.1. Production des coupes histologiques de cerveau de macaque

Le projet de ma thèse est basé sur un macaque sain âgé de 9 ans. Après l'euthanasie de ce macaque, son cerveau a été extrait puis congelé. Ensuite il a été découpé selon une incidence coronale sur un microtome à congélation. L'épaisseur d'une coupe de cerveau de macaque est typiquement de 40 μm comme dans la plupart des protocoles incluant des primates à MIRCen. De plus, l'ensemble des coupes produites ont été conservées dans une solution tampon jusqu'à l'étape de coloration.

Comme les coupes adjacentes sont proches spatialement et très similaires morphologiquement, nous pouvons diviser l'ensemble des coupes prélevées en plusieurs séries entrelacées ce qui permet, pour un même sujet, de réaliser plusieurs marquages histologiques simultanément. Dans le cadre de ma thèse, nous nous intéressons au nombre, à la morphologie et à la distribution des neurones dans le cerveau de macaque, nous avons donc choisi de marquer l'une des séries des coupes obtenues (134 coupes) en utilisant le marqueur neuronal NeuN. Ce marqueur s'exprime dans la plupart des neurones dans toutes les régions du cerveau. L'expression du marqueur neuronal NeuN est présente non seulement dans le noyau des neurones, mais également dans le cytoplasme de ces neurones qui est coloré plus légèrement. Ensuite, les coupes marquées ont été montées sur des lames de verre (76x52 mm).

2.2.2. Bases de données de l'étude

L'étude des neurones est basée sur des méthodes de traitement de l'image développées au sein du laboratoire. Nous présentons dans cette partie les différentes bases de données créées au cours de mon projet de thèse ainsi que les objectifs méthodologiques associés : la base des coupes sur

lesquelles nous avons travaillé, la base de données d'apprentissage / validation utilisée pour segmenter le marquage neuronal, la base de données utilisée pour individualiser les neurones et la base de données créée pour valider notre méthode grâce à la stéréologie. Toutes ces données ont été extraites sur quelques coupes de cerveau de macaque mais présentent des différences importantes en termes de taille, de nombre et de localisation des neurones.

2.2.2.1. Base d'images de coupes histologiques marquées par NeuN

Notre base de travail regroupe un nombre total de 134 coupes de cerveau de macaque. Ces coupes ont été numérisées avec un scanner de lames virtuelles Axio Scan Z.1 de Zeiss, avec une résolution de 0,22 μm (grossissement x20) dans le plan d'acquisition. Les images obtenues sont des images multi-résolutions au format CZI (format constructeur Carl Zeiss Imaging) qui intègrent des niveaux de résolution plus faibles surtout utilisés pour visualiser les données. La taille d'une coupe de cerveau de macaque scannée est de ~150 Go. Les temps de calcul et la possibilité de traiter ces données massives (accès lecture/écriture, algorithmes de traitement de l'image) sont des questions incontournables qu'il est nécessaire de résoudre dans le cadre de ce travail. Dans un premier temps, nous avons choisi de commencer par travailler sur trois coupes choisies de façon à être assez représentatives de la distribution des neurones dans l'ensemble du cerveau (Figure 33). Ces images contiennent une partie corticale importante ainsi que les principales régions cérébrales internes (striatum, thalamus, hippocampe, etc.). Elles présentent la plupart des différents cas de distribution et de densité de cellules que l'on peut trouver dans le cerveau entier : la coupe 81 (Figure 34), la coupe 91 (Figure 37) et la coupe 101 (Figure 35).

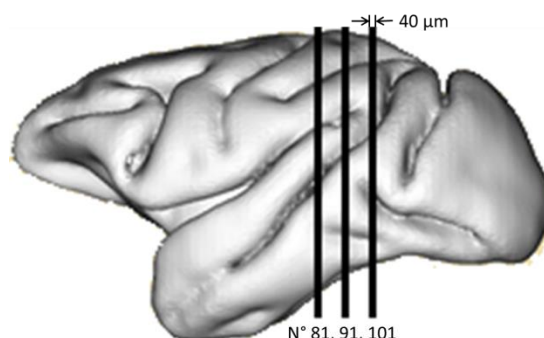


Figure 33 : Illustration de la position rostro-caudale des trois coupes coronales choisies (vue sagittale).

Les trois coupes sélectionnées ont été numérisées avec le logiciel Zen au moyen de quatre méthodes d'acquisition : acquisition en plan unique, acquisition en plan unique avec l'algorithme *unsharp masking*, acquisition en pile Z et acquisition en pile Z avec utilisation de l'algorithme EDF. Ces méthodes de numérisation sont présentées dans la section 1.3.2.3.

2.2.2.2. Base d'images pour la segmentation du marquage NeuN

Une base spécifique d'images a été produite pour entraîner les modèles de segmentation du marquage neuronal puis, pour valider la qualité des segmentations obtenues. Pour réaliser cette base, deux experts (une biologiste b1 et un traiteur d'image i1) ont segmenté manuellement des images en 3 classes : marquage histologique NeuN des neurones, tissu non marqué, absence de tissu. Un ensemble de 100 imageries représentatives (512x512 pixels) ont été extraites. Dans ce travail, nous avons choisi de travailler sur les coupes 81 et 101 afin d'éviter les biais de l'intensité des images histologiques produits par le marquage des neurones le but étant de tester la méthode sur la coupe 91. Au final, 42 imageries ont été extraites dans la coupe 81 (Figure 34) et 58 imageries dans la

coupe 101 (Figure 35). Le déséquilibre du nombre d'images entre les deux coupes est dû aux structures anatomiques présentes ou non dans les différentes coupes histologiques.

La base a été divisée en deux parties : une base d'apprentissage et une base de validation. La répartition des images entre la base d'apprentissage et la base de validation a été réalisée aléatoirement en tenant compte de la latéralisation du cerveau (nombre identique d'images issues des hémisphères gauche et droit) et d'un équilibre de représentation des différentes régions anatomiques. L'annexe 1 présente la répartition des images entre la base d'apprentissage et la base de validation. Au final, nous avons 54 images segmentées dans la base d'apprentissage et 46 images segmentées dans la base de validation.

La coupe 91 a été utilisée pour valider notre méthode d'individualisation et sera présentée plus tard.

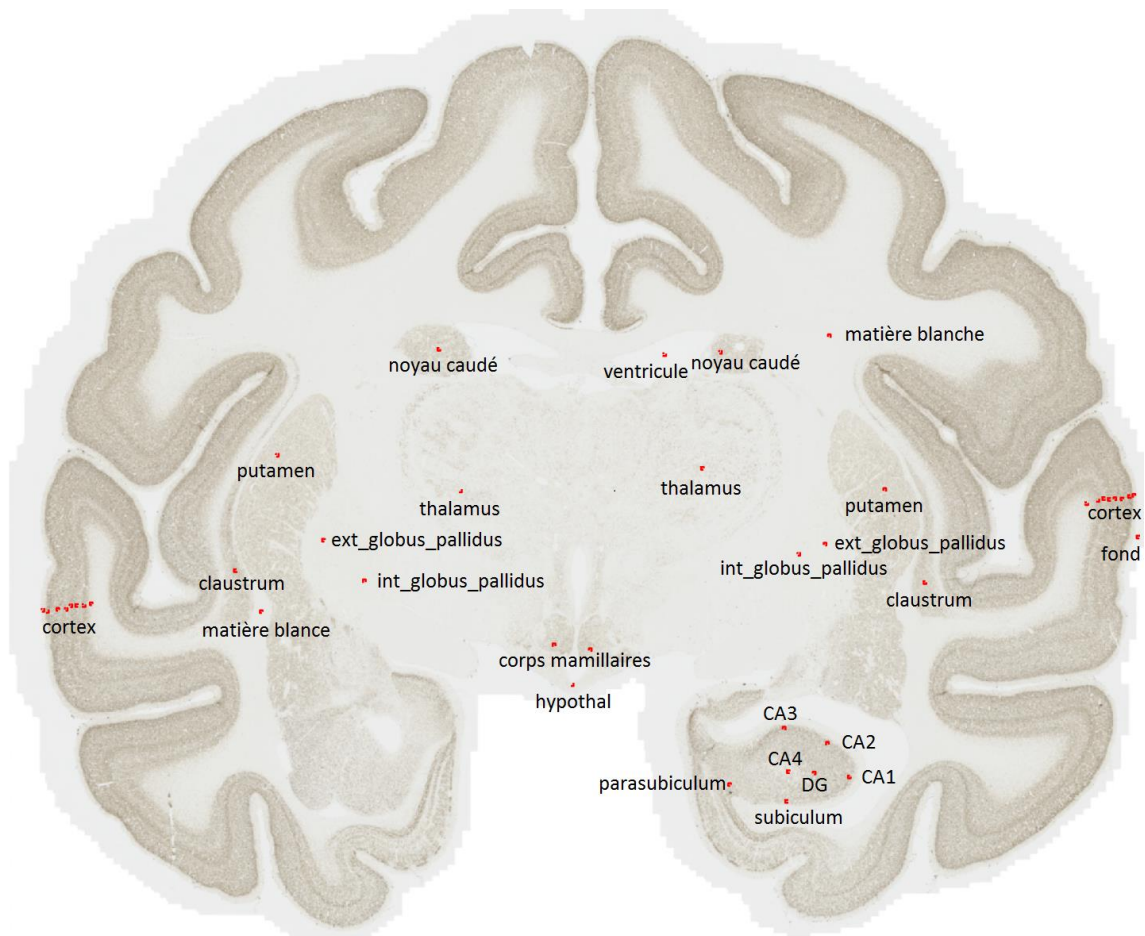


Figure 34: Image de la coupe 81 de dimensions 216000x178000 pixels (47,57 mm x 39,16 mm), ~130 Go. Les carrés rouges représentent la localisation spatiale des images de taille 512x512 pixels constituant la base d'images pour la segmentation des classes. Les légendes correspondent aux structures anatomiques auxquelles elles appartiennent.

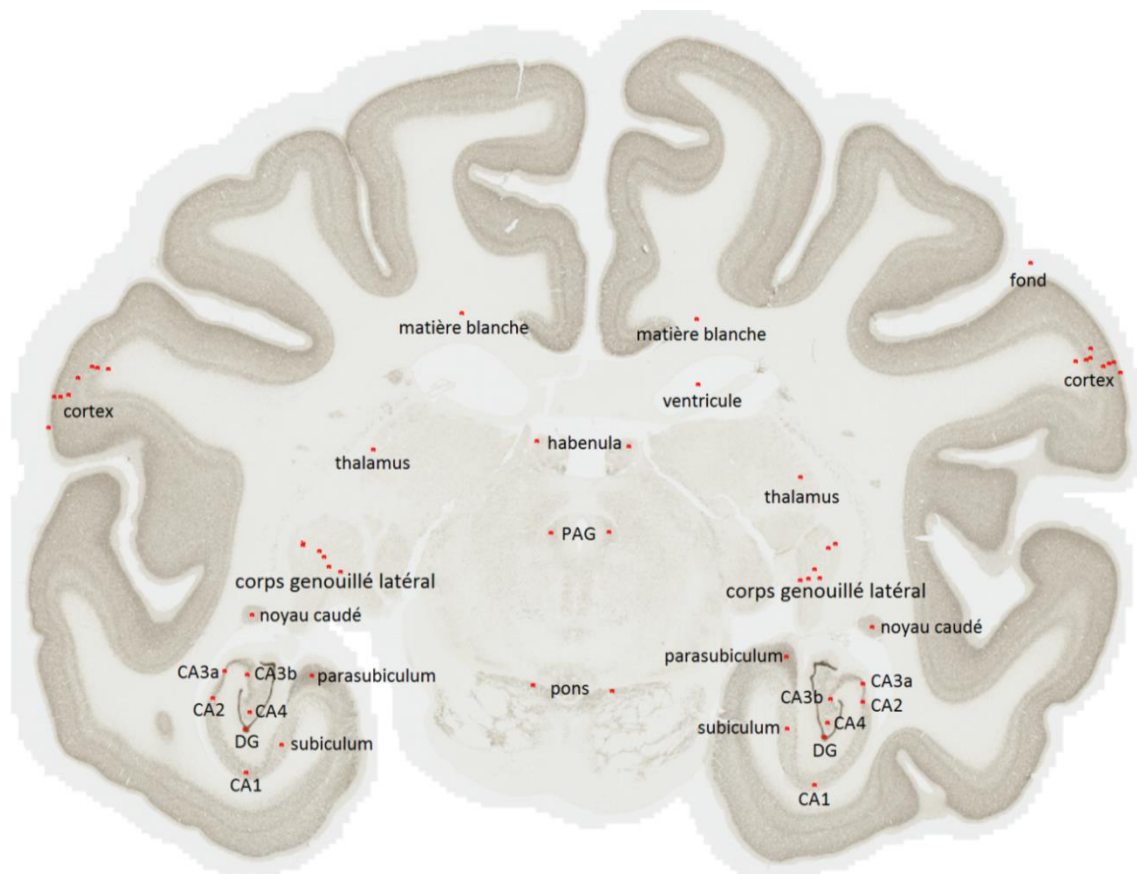


Figure 35: Image de la coupe 101 de dimensions 236000x181000 pixels (52 mm x 39,9 mm), ~148 Go. Les carrés rouges représentent la localisation spatiale des imagerie de taille 512x512 pixels constituant la base d'images pour la segmentation des classes. Les légendes correspondent aux structures anatomiques auxquelles elles appartiennent.

Les 100 imagerie ont été choisies avec l'aide de la biologiste b1 afin de couvrir les principales régions anatomiques du cerveau présentes dans les coupes traitées. Elles comprennent différentes tailles de cellules, représentent différentes densités des cellules et différents niveaux de coloration des cellules. Ces imagerie ont servi à créer le modèle de référence pour la segmentation automatique du marquage neuronal dans les coupes traitées. Cette segmentation pourrait être étendue au cerveau entier à condition de produire une coloration reproductible des coupes histologiques ce qui est le cas dans notre étude. La Figure 36 illustre cette diversité de densité des neurones à travers quelques imagerie issues du thalamus, du noyau caudé, de l'hippocampe et du cortex. Toutes les imagerie sont listées et illustrées dans l'annexe 1 – « Base d'images pour la segmentation des classes tissulaires ».

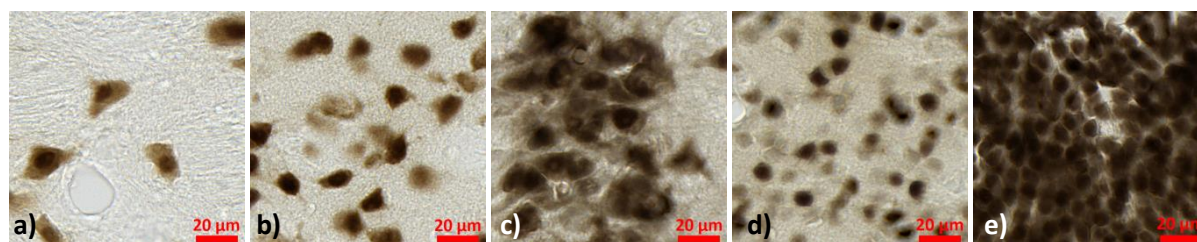


Figure 36: Illustration des imagerie de taille 512x512 pixels intégrées dans la base de données pour la segmentation du marquage neuronal : a) thalamus à gauche de la coupe 81; b) noyau caudé à droite de la coupe 81; c) CA3a à droite de la coupe 101; d) cortex cérébral à gauche de la coupe 101 et e) gyrus denté à droite de la coupe 101.

2.2.2.3. Base d'images pour l'individualisation des neurones

Pour appliquer et valider notre méthode d'individualisation des cellules, la base d'images constituée pour la segmentation du marquage neuronal est trop petite et ne contient pas un nombre suffisant de cellules. De ce fait, une nouvelle base de données a été créée avec des images de plus grandes dimensions qui sont plus représentatives des régions anatomiques du cerveau. En tenant compte de la capacité mémoire des lames de notre cluster de calcul dans le laboratoire (4 Go par cœur de calcul) et de la mémoire vive nécessaire à nos algorithmes, nous sommes arrivés à une estimation des dimensions traitables de l'ordre de 5000x5000 pixels pour les images de cette nouvelle base. Un ensemble de 58 images issues de la coupe 91 et représentant les principales régions anatomiques du cerveau ont été extraites (Figure 37).

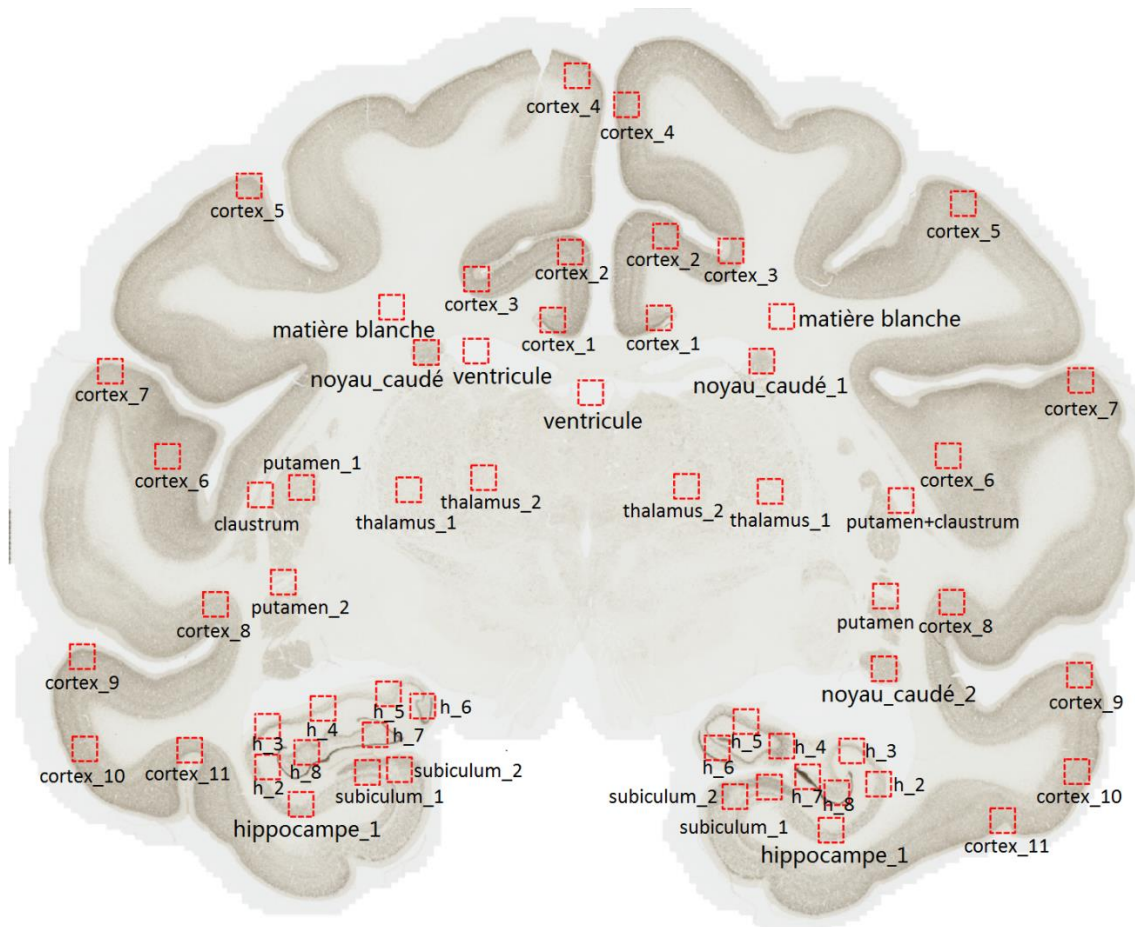


Figure 37: Image de la coupe 91 de taille 230000x188000 pixels (50,74 mm x 41,33 mm), ~145 Go. Les carrés rouges sont les images de taille 5000x5000 pixels qui composent la base d'images utilisée pour réaliser l'individualisation des cellules. « h » est une abréviation pour le nom de la région de l'hippocampe. Les numéros à la fin de chaque région correspondent aux numéros des images dans ces régions.

La Figure 38 illustre quatre images issues du noyau caudé, de l'hippocampe, du cortex et du thalamus. Comme pour la base précédente, les principales régions anatomiques ont été ciblées afin de tenir compte des différentes tailles, densités et organisations des neurones ainsi que des différents niveaux de coloration des neurones. Toutes les images sont listées dans l'annexe 2 – « Base d'images pour l'individualisation des neurones ».

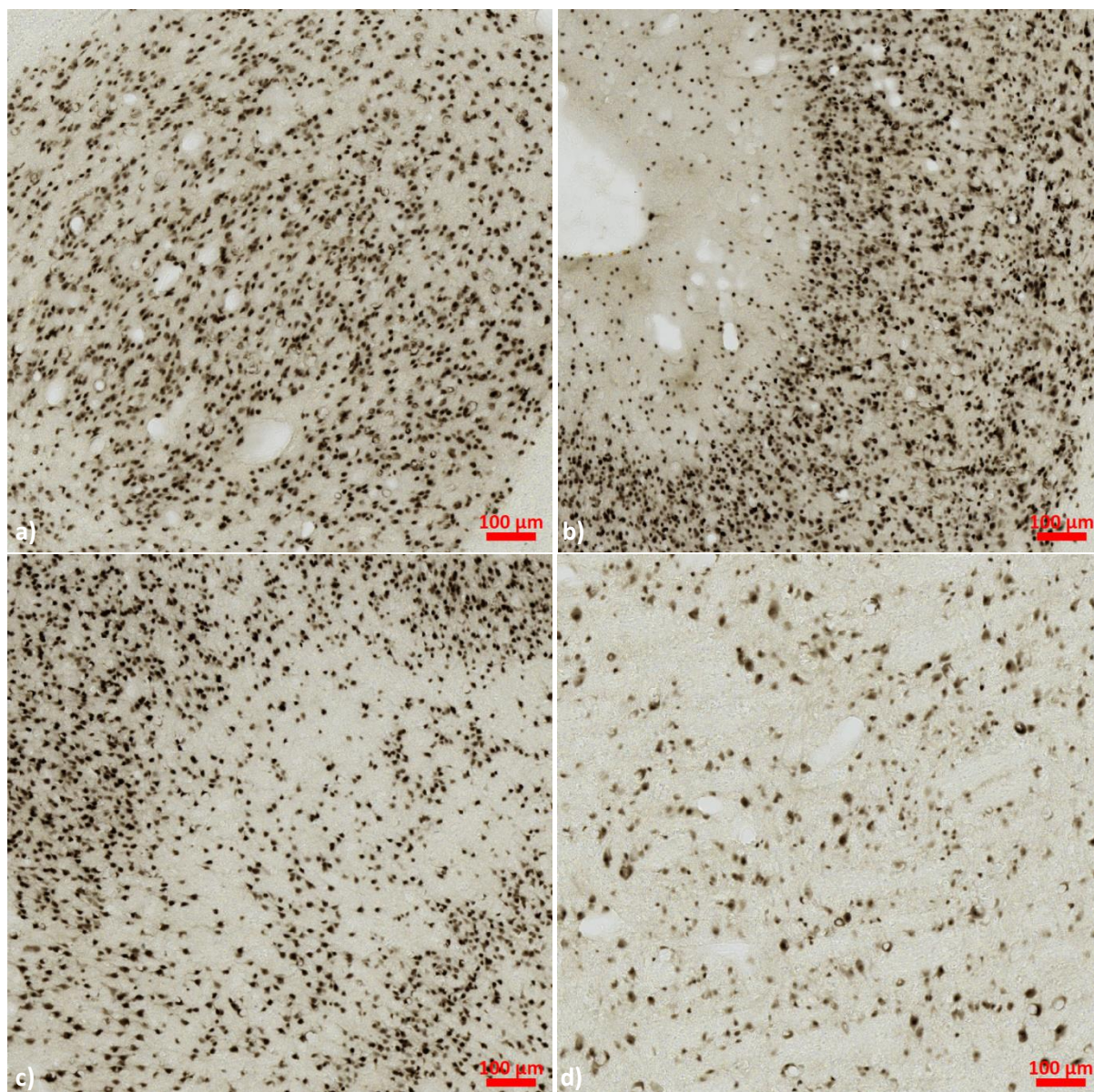


Figure 38: Illustration des images de taille 5000x5000 pixels de la base de données pour l'individualisation des neurones venant de la coupe 91 : a) noyau caudé; b) cortex cérébral; c) subiculum et d) thalamus.

2.2.2.4. *Base d'images pour le comptage stéréologique de référence des neurones*

Une dernière base de données a été créée afin de valider notre méthode d'individualisation des neurones par comparaison avec les résultats obtenus par comptage stéréologique. Cette technique de comptage ne peut pas être mise en œuvre ni sur les imagerie de 512x512 pixels, ni sur les images de 5000x5000 pixels des précédentes bases car elle travaille sur des structures anatomiques entières au sein de la coupe étudiée. De ce fait, une nouvelle base a été créée sous la supervision de biologistes experts en stéréologie. Dans la coupe 91, nous avons sélectionné 8 grandes régions anatomiques (Figure 39). Des régions présentant des densités de neurones variables ont été choisies. Deux biologistes expertes (b2 et b3) ont réalisé le comptage de référence par stéréologie directement sur un microscope optique.

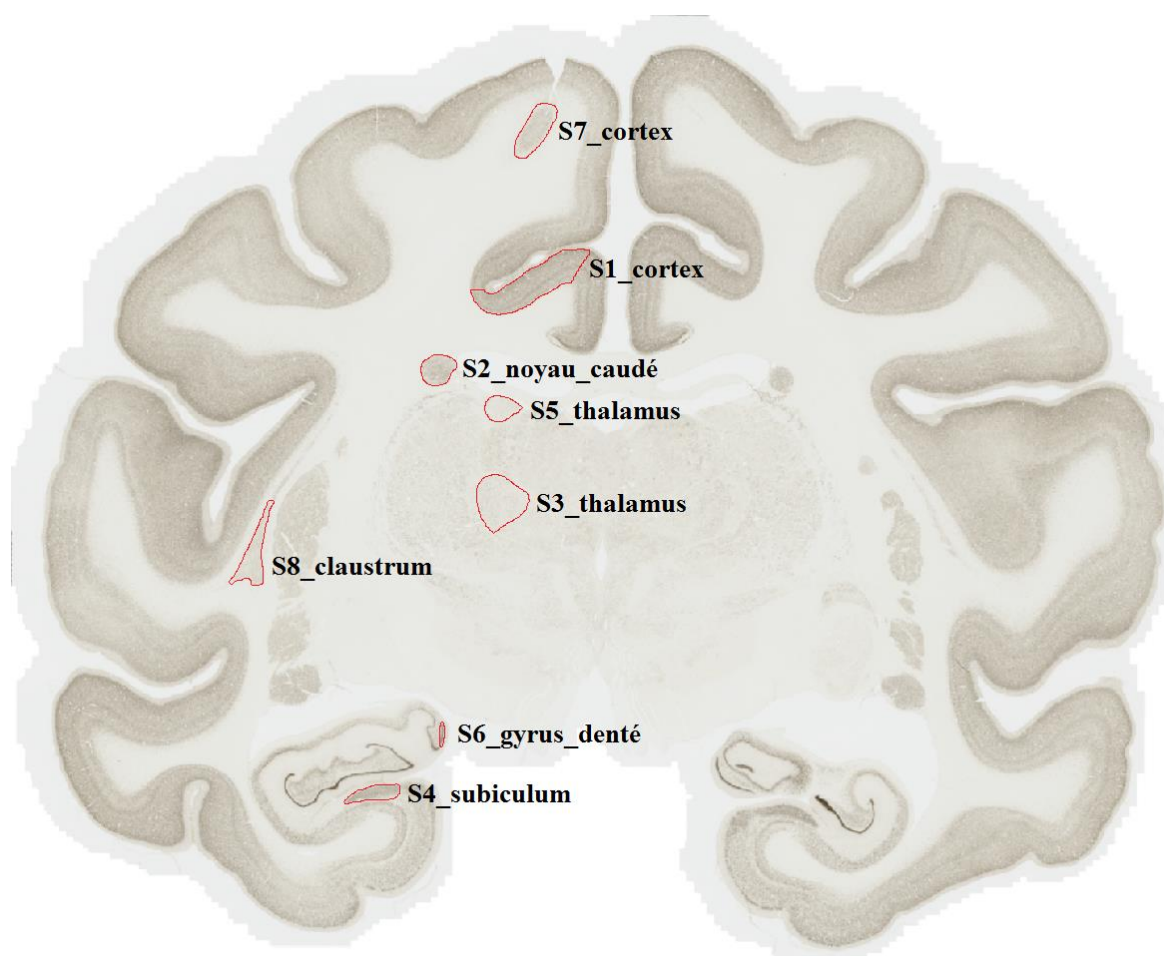


Figure 39 : Image de la coupe 91. Les régions entourées en rouge sont les zones ciblées pour réaliser le comptage stéréologique utilisé comme notre vérité terrain pour valider notre méthode d'individualisation. S1-8 correspond au numéro des RDIs.

2.2.2.5. Résumé des bases d'images créées

Notre travail a porté sur 3 coupes de cerveau de macaque sain à partir desquelles plusieurs bases de données ont été créées pour les différents objectifs méthodologiques. Nous résumons ici toutes les bases de données utilisées dans notre étude.

Tableau 2 : Résumé des bases d'images créées ainsi que leur nombre pour l'étude des neurones selon les objectifs méthodologiques.

| Bases d'images | | Taille des images (pixels) | Coupe histologique | | |
|-------------------------------------|---------------|-----------------------------------|--------------------|-------|--------|
| | | | N° 81 | N° 91 | N° 101 |
| Segmentation de marquage NeuN | Apprentissage | 512x512 | 25 | / | 29 |
| | Validation | | 17 | / | 29 |
| Individualisation des neurones | | 5000x5000 | / | 58 | / |
| Comptage stéréologique de référence | | $3 \times 10^6 - 1,2 \times 10^8$ | / | 8 | / |

2.3. Segmentation du marquage neuronal par apprentissage/classification

Afin d'étudier les neurones dans les coupes histologiques de cerveau de macaque, la segmentation de la classe neurone (marquage histologique NeuN des neurones) à partir du tissu est une étape

primordiale. Différentes méthodes pouvant être utilisées, nous nous sommes basés sur une étude précédente réalisée dans le laboratoire et avons choisi d'utiliser la méthode des forêts aléatoires (*Random Forest*, RF) qui avaient produit de très bons résultats pour des segmentations couleur de plaques amyloïdes (Vandenberghe et al., 2015).

2.3.1. Principe de la méthode des forêts aléatoires (RF)

Les forêts aléatoires sont populaires parmi les méthodes d'apprentissage/classification du fait de leur bonne précision, robustesse et facilité d'utilisation.

La méthode RF est une méthode d'apprentissage/classification qui fonctionne par construction d'une série d'arbres de décision lors de la phase d'entraînement. Pendant la construction d'un arbre de décision, une caractéristique (celle qui divise le mieux les données d'apprentissage sous forme de classes) est choisie à chaque nœud par maximisation du gain d'information (IG). Le gain d'information est un critère d'évaluation qui est basé sur le concept d'entropie (H) (équation 4).

$$\left\{ \begin{array}{l} H = -\sum_{i=1}^m p_i \log_2(p_i) \\ IG(T, f) = H(T) - H(T | f) \end{array} \right. \quad (4)$$

Où m représente le nombre de classe. p_i est la fraction dont la somme est égale à 1 et représente le pourcentage de chaque classe i présente dans le nœud qui résulte d'une scission du nœud parent. T représente une base d'apprentissage. f est une caractéristique. Information Gain = l'Entropie (parent)-Somme conditionnelle de l'Entropie (enfant). La maximisation d' IG est équivalente à la minimisation de la somme pondérée de l'entropie enfant.

Les nœuds de l'arbre sont scindés itérativement selon les caractéristiques (colorimétrie et texture) choisies jusqu'à ce que chaque branche de l'arbre ne contienne plus que les observations venant uniquement d'une classe.

Un seul arbre de décision n'est pas assez précis pour prédire la classe pour les nouvelles données car il sur-apprend les données d'apprentissage. Afin de fournir une meilleure performance de segmentation, pour la méthode RF, un nombre de B d'arbres sont construits en utilisant un tirage avec remise d'un nombre N d'observations dans notre base d'apprentissage (la technique de *bootstrap*) et des caractéristiques. Nous trouvons régulièrement dans la littérature (Amit et Geman, 1997; Geurts et al., 2006) des études utilisant une centaine d'arbres de décision. Le partitionnement des caractéristiques pour chaque arbre de décision a été établi par minimisation de la somme pondérée de l'entropie enfant. La segmentation sous forme de classe en utilisant la méthode RF est obtenue lors de la phase de classification par simple vote majoritaire de la prédiction dans les B arbres de décision.

2.3.2. Segmentation de référence des neurones

Nous avons demandé à deux experts dans le domaine de la biologie (noté b1) et du traitement de l'image (noté i1) de segmenter manuellement les 100 imagerie représentatives venant des coupes N°81 et N°101 en 3 classes (marquage histologique NeuN des neurones, tissu non marqué et absence de tissus). Ces segmentations manuelles correspondent à la vérité terrain (Figure 45. b1-b5) et seront utilisées pour la phase du choix des caractéristiques nécessaires à la génération du modèle d'apprentissage et enfin pour la validation de la méthode de segmentation automatique.

2.3.3. Création du modèle d'apprentissage optimal par RF

Pour créer le modèle d'apprentissage, il a fallu définir les caractéristiques optimales. Nous avons au total étudié 17 caractéristiques de colorimétrie (Cheng et al., 2001) et de texture : $R, G, B, H, S, V, X, Y, Z, L^*, a^*, b^*, M, V, LBP\ 11, LBP\ 40, LBP\ 63$. Elles sont décrites dans le Tableau 3.

Tableau 3 : Caractéristiques à étudier pour créer le modèle d'apprentissage.

| Caractéristiques | | Commentaires |
|------------------|-------------------------------|--|
| Colorimétrie | RGB | <i>Red Green Blue</i> , trois couleurs primaires – rouge, vert, bleu. |
| | HSV | <i>Hue Saturation Value</i> , correspondant à la teinte, la saturation et la valeur, c'est un autre espace de couleur qui est plus intuitif pour la vision humaine. |
| | XYZ | Un espace de couleur défini par CIE (Commission internationale de l'éclairage) qui a été développé pour représenter l'uniformité perceptuelle et répond ainsi au besoin psychophysique d'un observateur humain. Il a trois primaires notés X, Y et Z. Toute couleur peut être spécifiée par la combinaison de X, Y et Z. |
| | $L^*a^*b^*$ | Un autre espace de couleur défini par CIE qui peut être calculé par la transformation non linéaire des composantes de l'espace XYZ. |
| Intensité locale | MV | Il s'agit de la moyenne et de la variance des valeurs d'intensité converties en niveau de gris sur un élément structurant (une croix d'amplitude 10 pixels qui correspond au rayon minimal de neurone attendu – 2,5 μ m) |
| Texture | LBP | Local Binary Patterns. 11, 40, 63 correspondent respectivement au rayon de petit, de moyen et de grand neurone en nombre de pixels |

Pour les caractéristiques de textures, nous précisons en détail le calcul de *LBP* qui est une caractéristique puissante pour reconnaître des textures dans des images (Wang et He, 1990; Ojala et al., 1996; Ojala et al., 2002). Elle est calculée facilement et considérée comme robuste aux changements de niveau de gris dûs à la variation de luminance. La texture *LBP* est calculée dans un cercle de rayon prédéfini dans une image en niveaux de gris. Soit g_c la valeur d'intensité du centroïde c du cercle, g_p la valeur d'intensité des pixels p ($p=0, 1, \dots, P-1$) dans ce cercle. Nous comparons alors chaque pixel p avec le centroïde c , si la valeur d'intensité de p est plus grande ou égale à celle de c , la valeur d'intensité de p est alors remplacée par 1, sinon, elle est remplacée par 0. Les P pixels génèrent un vecteur de caractéristiques, ex. (1,1,1,1,0,0,0,1) dans la Figure 40. La caractéristique de texture *LBP* est la valeur décimale convertie par ce vecteur et est calculée selon l'équation 5.

$$LBP_c = \sum_{p=0}^{P-1} 2^p \delta(g_p - g_c) \quad (5)$$

Où δ est la fonction de Heaviside.

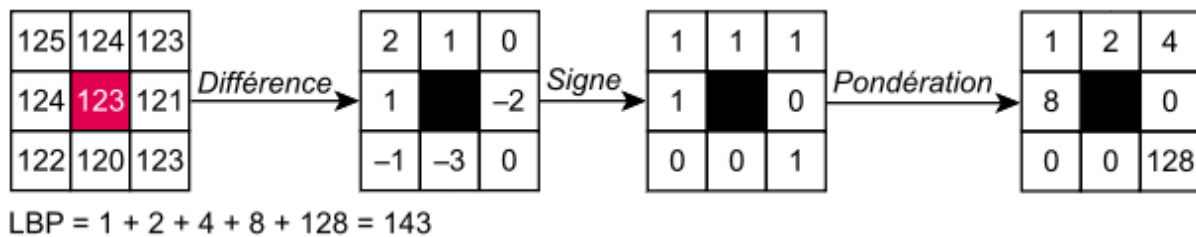


Figure 40: Illustration de l'algorithme pour calculer la caractéristique de texture *LBP*, le rayon est de 1 pixel.

Comme la taille des neurones varie entre 5 et 30 μm de diamètre (soit de 11 à 68 pixels de rayon), dans notre cas, nous avons testé des éléments structurants sous forme de cercles de rayon 11, 40 et 68 pixels.

La Figure 41 montre les résultats de notre étude obtenus avec les différentes combinaisons de caractéristiques lorsque l'on réduit progressivement le nombre de caractéristiques en sélectionnant les plus informatives. Les pourcentages correspondent à l'importance des caractéristiques dans les RF (Breiman et al., 1984), qui est définie comme la diminution totale de l'impureté du nœud en moyenne sur tous les arbres de l'ensemble. Si la diminution est faible, la caractéristique n'est pas importante et inversement.

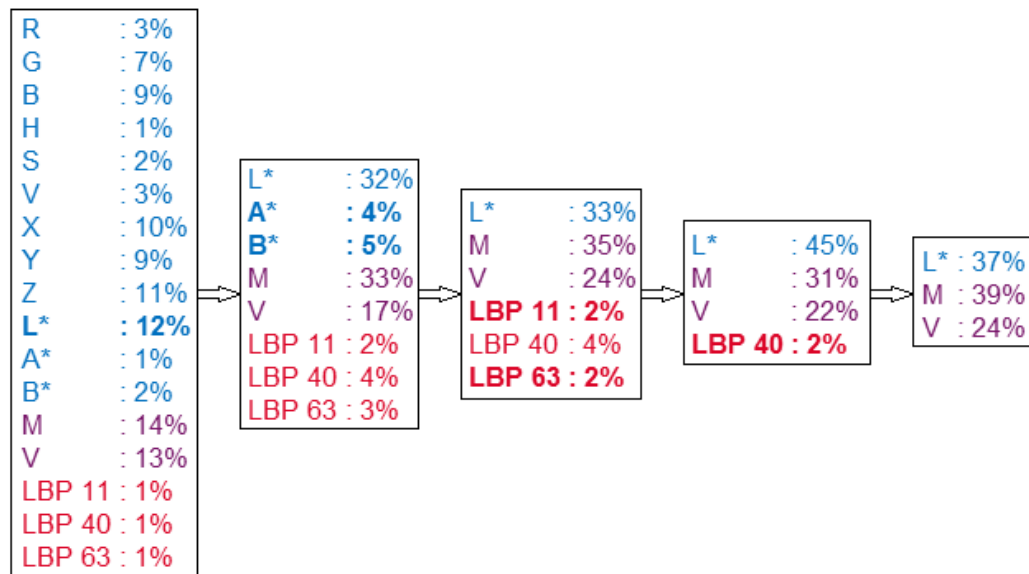


Figure 41 : Etude de différentes combinaisons de caractéristiques. Bleu : caractéristiques de l'espace colorimétrique. Violet : caractéristiques de l'intensité locale. Rouge : caractéristiques de textures. D'après l'étude de Clément Bouvier (thèse du CEA-MIRCen en cours).

La Figure 42 montre la similarité (*F-score*) entre la segmentation manuelle et les segmentations automatiques des 42 imagerie issues de la base de validation selon les différentes combinaisons de caractéristiques (42 imagerie ici au lieu de 46 car il existe 4 imagerie sans tissu marqué issues de zones de matière blanche ou du fond sans tissu et leurs segmentations sont totalement correctes à savoir qu'il n'y a pas de classe neurone détectée). Le *F-score* correspond également à l'indice Dice (Sørensen, 1948) très largement utilisé en traitement de l'image pour évaluer la qualité de segmentation. De façon empirique, des valeurs de *F-score* supérieures à 0,7 sont généralement considérées comme satisfaisantes. Le Tableau 4 liste les valeurs moyennes de *F-score* de chaque combinaison de caractéristiques testée pour 3 groupes de neurones : imagerie de tous les neurones dans les 42 imagerie, imagerie contenant des neurones plus clairs (qui existent dans l'imagerie du globus pallidus interne et l'imagerie du corps genouillé latéral) et imagerie contenant des neurones généralement foncés (qui existent dans les imagerie restantes). La Figure 28 (image de gauche) illustre des exemples de neurones plus clairs qui sont caractérisés par une faible intensité de coloration. Si le *F-score* est très important donc proche de 1, la segmentation automatique est de très bonne qualité et inversement (score proche de zéro) – la superposition des segmentations comparées est faible.

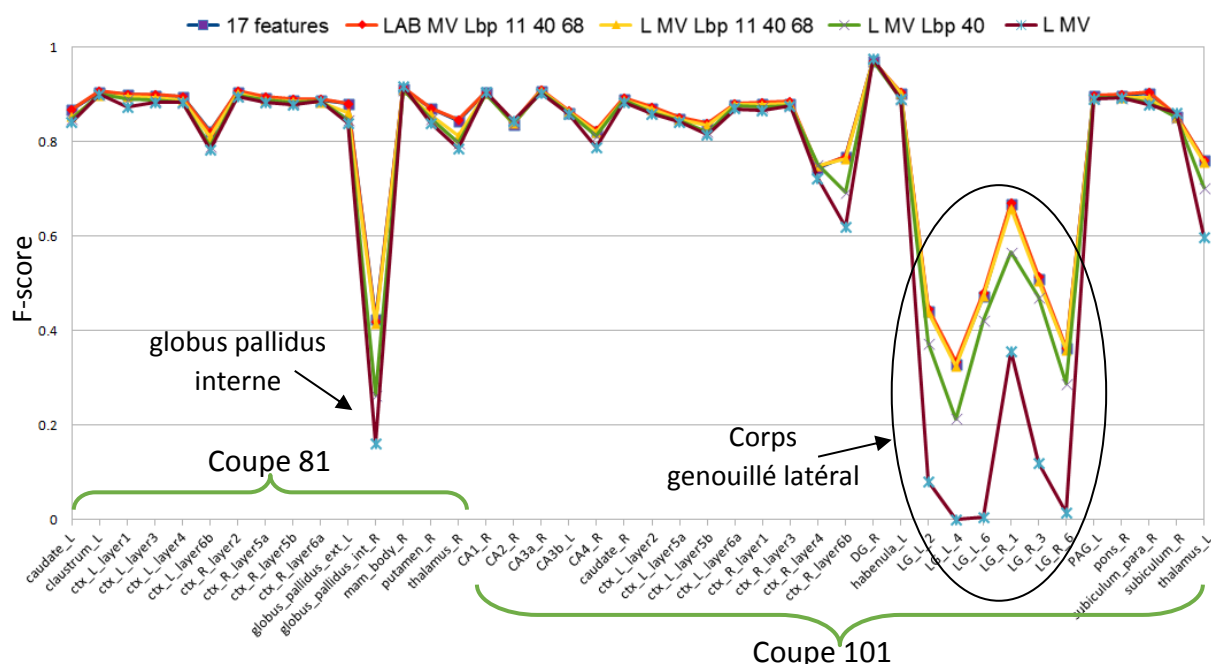


Figure 42 : Comparaison (*F-score*) entre la segmentation manuelle et les segmentations automatiques obtenues en fonction de différentes combinaisons de caractéristiques testées. « ctx » est une abréviation pour le nom de la région du cortex. « LG » est une abréviation anglaise pour le nom de la région du corps genouillé latéral. D'après l'étude de Clément Bouvier (thèse du CEA-MIRCen en cours).

Tableau 4 : Valeurs moyennes de *F-score* pour tous les neurones dans les 42 imagerie, pour des jeux d'images contenant des neurones clairs (qui existent dans une imagerie du globus pallidus interne et six imagerie du corps genouillé latéral) et pour des d'images contenant des neurones généralement foncés (qui existent dans les imagerie restantes).

| <i>F-score</i> | <i>17 features</i> | <i>LAB MV Lbp 11 40 68</i> | <i>LMV Lbp 11 40 68</i> | <i>LMV Lbp 40</i> | <i>LMV</i> |
|--------------------------|--------------------|--------------------------------|-----------------------------|-----------------------|------------|
| Tous les neurones | 0,80 | 0,80 | 0,80 | 0,78 | 0,73 |
| Neurones foncés | 0,87 | 0,87 | 0,87 | 0,86 | 0,85 |
| Neurones clairs | 0,46 | 0,46 | 0,45 | 0,37 | 0,11 |

Nous observons que le *F-score* est identique pour les trois premières combinaisons des caractéristiques (17 features, LAB MV Lbp 11 40 68 et L MV Lbp 11 40 68). Il s'agit des caractéristiques colorimétriques, des caractéristiques d'intensité locale et des caractéristiques de texture. Cela démontre que nous pouvons diminuer le nombre de caractéristiques colorimétriques pour réduire le volume de données à traiter, économiser du temps de calcul et garder en même temps la qualité de la segmentation produite. Lorsque nous diminuons le nombre des caractéristiques de textures jusqu'à la quatrième combinaison (LMV Lbp 40), le *F-score* baisse légèrement de 2,5% pour tous les neurones, de 1,1% pour les neurones foncés et de 17,8% pour les neurones clairs par rapport à la combinaison L MV Lbp 11 40 68. Le *F-score* baisse de 8,8% pour tous les neurones, de 2,3% pour les neurones foncés et de 75,6% pour les neurones clairs dans le cinquième cas de combinaison (LMV) par rapport à la combinaison L MV Lbp 11 40 68. Ceci s'explique du fait qu'il n'y a plus de caractéristiques de texture ce qui démontre l'importance de ce type de caractéristique pour la segmentation du marquage NeuN. En prenant en considération un compromis entre performance et temps de calcul, nous avons choisis d'utiliser la combinaison L MV Lbp 40 pour réaliser les segmentations automatiques du marquage NeuN.

Nous avons extrait les quatre caractéristiques retenues sur les images de la base d'apprentissage et de la base de validation. La Figure 43 illustre les images correspondant à ces quatre caractéristiques sur un exemple.

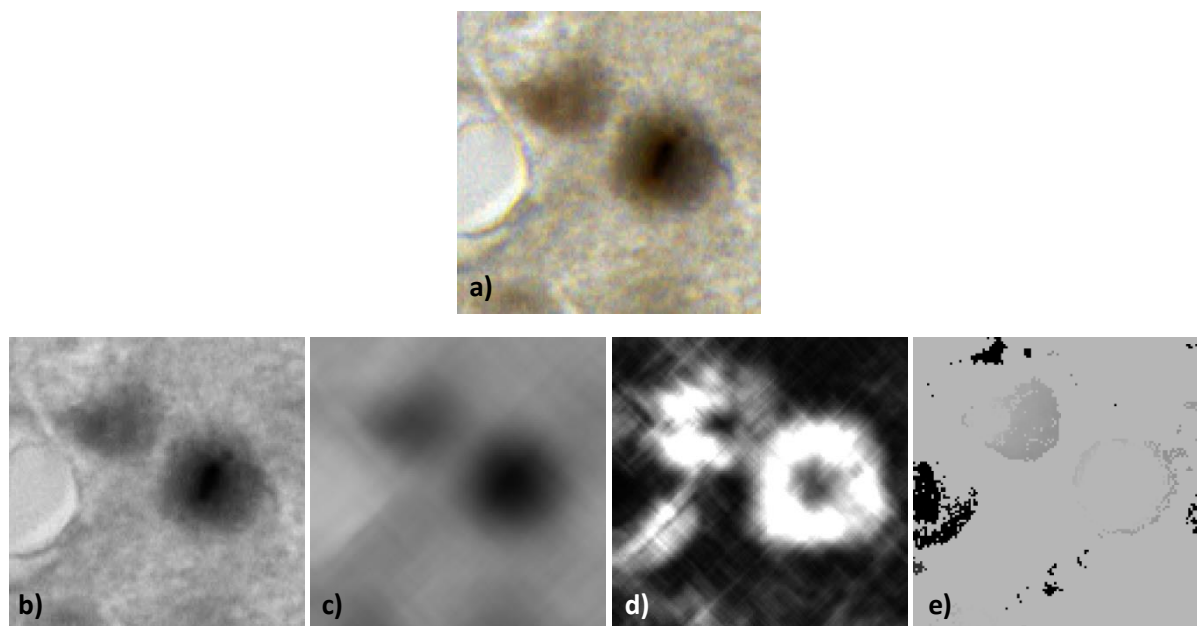


Figure 43 : Extraction des caractéristiques retenues. a) Imagerie dans laquelle les neurones sont marqués par NeuN. Les neurones apparaissent en marron foncé. b-e) Respectivement, les quatre caractéristiques utilisées pour la segmentation du marquage NeuN : b) L^* , c) M , la moyenne et d) V , la variance de nuances de gris convertie à partir de l'image RGB, et e) LBP avec un élément structurant en forme de disque de rayon 40 pixels.

Lors de la phase d'apprentissage par la méthode RF, nous avons choisi de générer le modèle RF en construisant $B=100$ arbres de décision de profondeur 4 (Figure 44). Ces arbres sont construits en utilisant un tirage avec remise d'un nombre N d'observations dans notre base d'apprentissage. Le partitionnement de ces quatre caractéristiques pour chaque arbre de décision est calculé en fonction de la maximisation d' IG ou de la minimisation de la somme pondérée de l'entropie enfant.

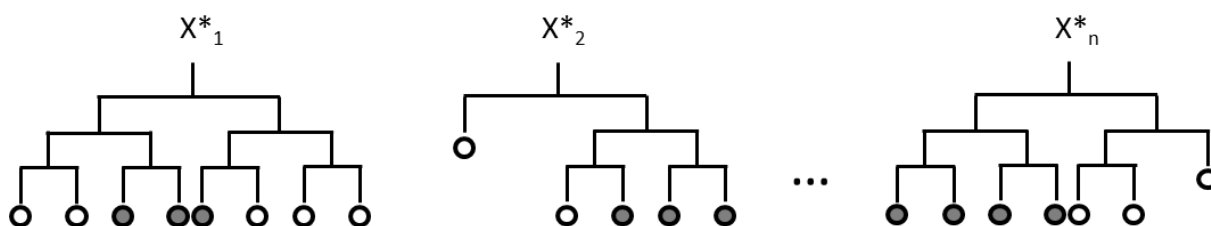


Figure 44 : Illustration du modèle de RF utilisé.

2.3.4. Classification des pixels de la base de segmentation

En fonction des caractéristiques extraites à partir de la base de validation et du modèle RF généré à partir de la base d'apprentissage, le résultat de la segmentation par la méthode RF pour la base de validation est alors obtenu par un vote majoritaire de la prédiction des 100 arbres de décision pour chaque pixel étudié (Figure 45. d1-d5). Ce modèle RF sera utilisé comme le modèle de référence pour la méthode de RF pour la segmentation de la base d'images pour l'individualisation des neurones et de base d'images pour le comptage stéréologique. Afin de vérifier la robustesse de ce modèle de référence, nous avons réalisé une permutation des imagerie entre les bases d'apprentissage et de validation. Les valeurs moyennes de F -score calculées sont respectivement de 0,858 (modèle RF de

référence) et 0,864 (modèle RF croisé) pour la classe neurone. Cela signifie que les deux modèles RF sont équivalents et non biaisés. La segmentation par la méthode RF n'est pas dépendante de la construction de nos deux bases et n'est pas biaisée non plus.

Vandenbergh et al. (2015) ont présenté la meilleure qualité de segmentation par la méthode RF par rapport à la méthode bioVision (Chubb et al., 2006) pour les plaques amyloïdes. Nous avons décidé de mener une étude similaire pour les neurones. Pour la comparaison de ces deux méthodes, nous avons utilisé pour bioVision le vecteur de caractéristiques typique *RGBM*. Le modèle bioVision a été créé avec la base d'apprentissage et a ensuite été utilisé pour la segmentation de la classe neurone dans la base de validation.

La Figure 45 montre la comparaison des résultats de segmentation entre la segmentation manuelle, celle obtenue par bioVision et par RF. Visuellement, la méthode RF détecte plus détails que bioVision, surtout dans les cas de présence de vaisseaux sanguins (Figure 45. d1, le grand d'objet noir représente le vaisseau sanguin, segmenté manuellement par l'expert comme la classe absence de tissus) et pour les neurones clairs (Figure 45. d2, d4). Dans le cas complexe (Figure 45. d5), la méthode RF détecte plus de détails qui n'ont pas été segmentés par les experts à l'intérieur des amas de neurones extrêmement denses et sur les tissus non marqués.

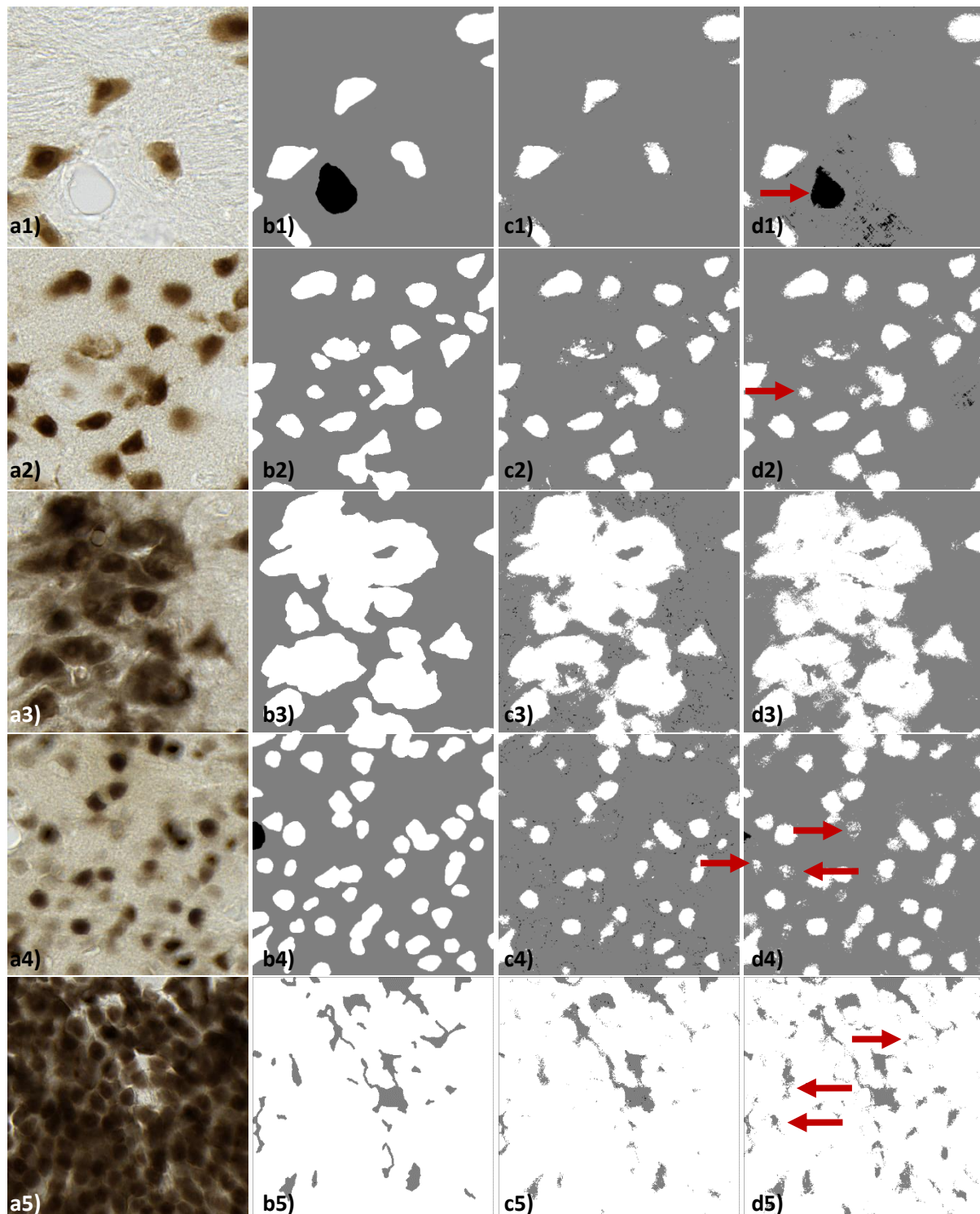


Figure 45 : Comparaison des résultats de segmentation entre la vérité terrain, bioVision et RF. a1-a5) Imagerie originales. b1-b5) Vérité terrain. c1-c5) BioVision utilisé avec 4 caractéristiques (R , G , B , M). d1-d5) RF utilisé avec 4 caractéristiques (L , M , V , Lbp 40). Les flèches rouges pointent les meilleures détections réalisées par la méthode RF.

Afin d'évaluer quantitativement les méthodes de segmentation, les valeurs moyennes de F -score ont été calculées. La Figure 46 montre le F -score obtenu par la méthode RF ($LMVLbp40$) et bioVision ($RGBM$). La performance de la méthode RF (0,78) est globalement supérieure à celle de bioVision (0,68). Les basses valeurs du F -score (0-0,6) correspondent aux imagerie extraites à partir de régions

claires (globus pallidus interne et corps genouillé latéral) dans lesquelles la coloration des neurones est moins forte. La méthode RF, dont la valeur moyenne de *F-score* sur ces régions claires est de 0,37, améliore significativement la segmentation des neurones clairs sans pour autant donner une segmentation satisfaisante (*F-score*<0,7). La méthode bioVision n'obtient qu'une valeur moyenne de *F-score* de 0,07 pour ces régions claires ce qui est très faible. Les imagerie restantes correspondent aux régions généralement foncées, la valeur moyenne de *F-score* est de 0,86, supérieure à celle de bioVision qui est de 0,81. Nous observons que dans l'imagerie ctx_R_layer6b (la couche 6b du cortex dans la coupe 81) et thalamus_L (thalamus dans la coupe 101), les valeurs moyennes de *F-score* pour les deux méthodes sont moins importantes que celle des autres régions plutôt foncées, car ce sont des régions qui contiennent en même temps des neurones foncés et clairs.

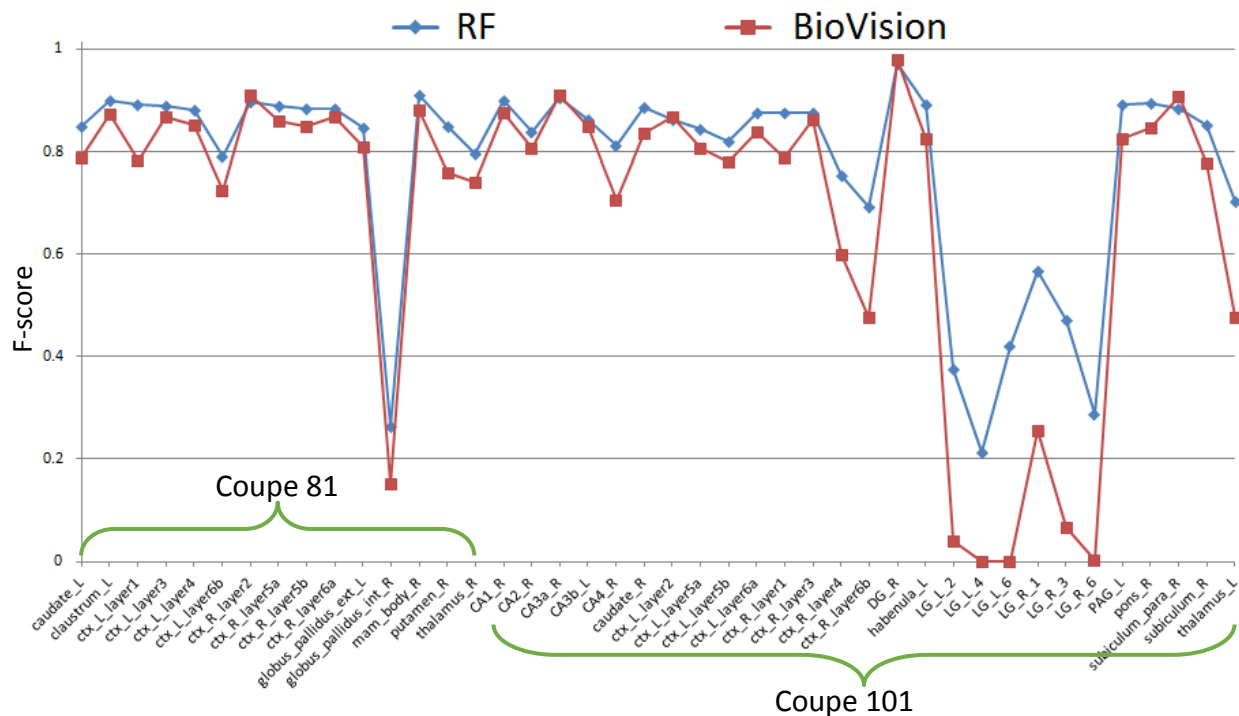


Figure 46 : Comparaison (*F-score*) entre la segmentation manuelle et le résultat de la segmentation automatique de la base de validation réalisée par RF (*LMVLbp40*) et bioVision (*RGBM*). D'après l'étude de Clément Bouvier (thèse CEA-MIRCen en cours).

La Figure 47 illustre les segmentations dans la région du globus pallidus interne où le marquage des neurones est clair par les méthodes RF et bioVision. La méthode RF n'arrive pas à segmenter de façon satisfaisante les neurones clairs, mais visuellement, la méthode RF permet de détecter plus de neurones clairs que la méthode bioVision et donc de limiter la sous-segmentation.

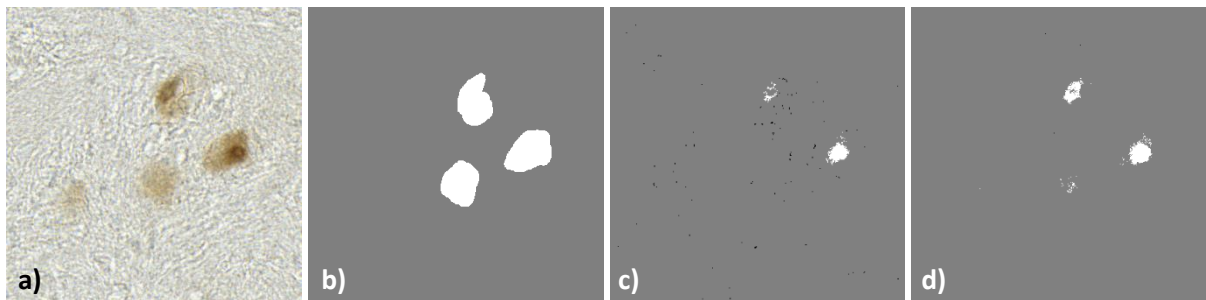


Figure 47 : Illustration des résultats de segmentation dans une région présentant un marquage clair. a) Imagerie originale. b) Vérité terrain. c) bioVision avec 4 caractéristiques (*R*, *G*, *B*, *M*). d) RF avec 4 caractéristiques (*L*, *M*, *V*, *Lbp 40*).

2.3.5. Débruitage des résultats de segmentation du marquage NeuN

Dans les segmentations automatiques obtenues, il existe un bruit important qui correspond à de petites composantes, des nuages de points (un neurone qui n'est pas entièrement segmenté mais se présente sous la forme de nuages de points non compact) et également des irrégularités au niveau des contours externes. Ceci est dû principalement au fait que la segmentation de la classe neurone est réalisée pixel par pixel. Afin de régulariser les segmentations automatiques, nous avons étudié l'impact de deux filtres classiquement utilisés dans ce genre de problème : le filtre médian et l'opérateur de morphologie mathématique (ouverture = érosion + dilations, (Serra, 1982)). Nous considérons comme acquis que le rayon des plus petits neurones est de l'ordre de 2,5 μm ce qui correspond à 11 pixels pour le rayon et 380 pixels pour la surface correspondante si l'on considère la forme d'un neurone comme étant un disque. Ces informations nous ont guidés pour régler les paramètres de filtrage testés. Pour le filtrage médian, nous avons testé différentes tailles de fenêtre sous forme de carré, entre 3 et 11 pixels par pas de 2 pixels. Ensuite, pour supprimer les composantes trop petites pour être des neurones, nous avons décidé de supprimer les objets individuels qui présentaient une surface inférieure à 127 pixels ce qui correspond à un tiers de la surface du plus petit neurone précédemment estimée. Pour ce qui est de l'opérateur de morphologie mathématique, nous avons fixé la taille de l'élément structurant à 1,44 μm ($2,5/\sqrt{3}$), ce qui correspond à un tiers de la taille du neurone le plus petit). Nous avons comparé la classe neurone dans les images filtrées (46 imageries dans la base de validation et les segmentations manuelles, puis calculé la performance selon les critères de Rappel, Précision et *F-score* (Tableau 5).

Tableau 5 : Comparaison de la performance entre les segmentations manuelles et les images de segmentation automatique filtrées pour la classe neurone.

| Filtre | Rappel | Précision | <i>F-score</i> |
|--|-----------|-----------|----------------|
| Segmentation originale | 0,7502202 | 0,9212629 | 0,7967102 |
| Morpho math 1,44 μm | 0,6290369 | 0,9353995 | 0,7151888 |
| médian 3 | 0,7524399 | 0,9332249 | 0,8008133 |
| médian 5 | 0,7512265 | 0,9372958 | 0,8002784 |
| médian 7 | 0,7489645 | 0,9404244 | 0,7987485 |
| médian 9 | 0,7457071 | 0,9427628 | 0,7962224 |
| médian 11 | 0,7418549 | 0,9450035 | 0,7931521 |

A partir de ce tableau, nous observons que la performance diminue lorsque la taille de la fenêtre du filtre médian augmente. Cependant, nous trouvons que la différence est très faible, inférieure à 1%. La Figure 48 montre un exemple de résultats de filtrage ainsi obtenus. Visuellement, il existe encore des aspérités sur les contours des neurones dans l'image filtrée par l'opérateur ouverture. L'opérateur enlève également de manière erronée de petits neurones parce que ceux-ci apparaissent comme des nuages de points qui sont éliminés lors de l'étape d'érosion. Pour le filtrage médian, quand la taille de la fenêtre est supérieure à 5 pixels, les contours sont lissés, et quand la taille est supérieure à 7 pixels, plusieurs petits neurones sont supprimés. Par conséquent, pour le filtre médian, la taille de fenêtre que nous avons choisie est égale à 7 pixels. Avec ce paramètre, l'effet de filtrage est visuellement satisfaisant. De plus, le *F-score* du filtre médian 7 est très proche de celui de la segmentation automatique du marquage neuronal et donne de plus une bonne précision. Il est intéressant de noter que les scores quantitatifs calculés dans cette étude sont tous très proches et ne permettent pas vraiment de décider de la méthode optimale. L'évaluation visuelle, malgré son caractère qualitatif, donne des informations complémentaires pour effectuer ce choix.

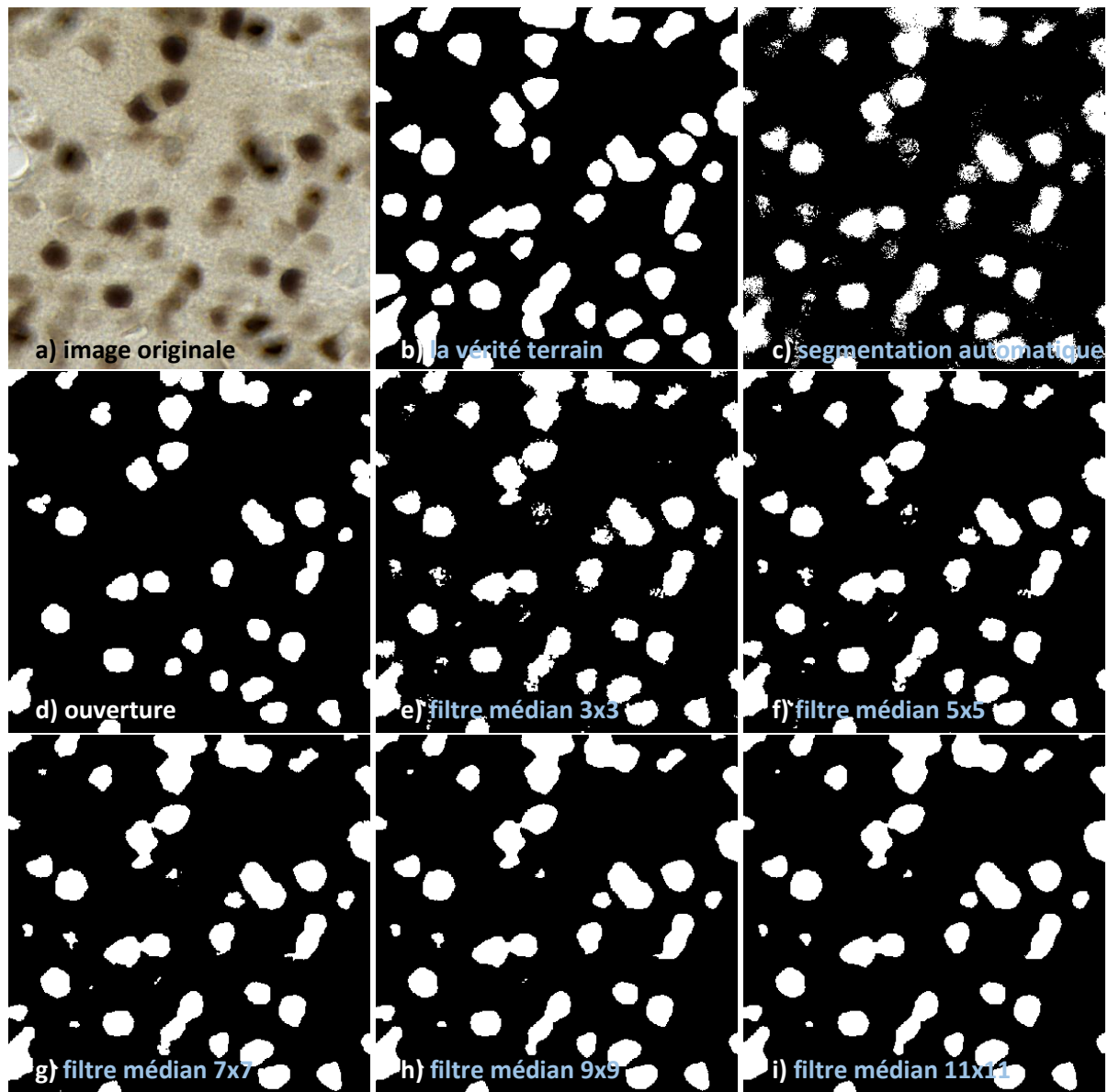


Figure 48: Résultats de deux méthodes de filtrage testées pour régulariser les segmentations automatiques produites par RF. a) Image originale. b) Segmentation manuelle. c) Image de masque obtenue par la segmentation automatique RF du marquage neuronal. d) Image filtrée par la technique de morphologie mathématique (ouverture), la taille de l'élément structurant est de $1,44\ \mu\text{m}$. e-i) Images filtrées par filtrage médian, la taille de la fenêtre en forme de carré est respectivement 3, 5, 7, 9 et 11 pixels. Après le filtrage, les petits objets dont la surface est inférieure à 127 pixels sont supprimés.

2.4. Méthode d'individualisation des neurones

La Figure 45 montre que dans des régions peu denses, la plupart des neurones sont déjà individualisés, cependant dans des régions plus denses, une étape de segmentation seule ne permet pas de séparer les neurones qui se touchent. Cette proximité peut s'expliquer par la proximité physique des cellules mais également par le fait que l'image produite est issue d'une projection de l'épaisseur de la coupe ce qui peut produire cet effet par superposition sans que pour autant les cellules ne soient réellement jointives. Dans cette partie, nous allons présenter notre méthode d'individualisation qui permet de séparer les neurones accolés.

2.4.1. Principe de la méthode – Hypothèses de travail

Les images de la base sur lesquelles nous travaillons ont été obtenues à partir des coupes histologiques colorées par NeuN dont l'expression est présente dans le noyau ainsi que dans le cytoplasme des neurones. La coloration des neurones est plus sombre dans le noyau et plus claire dans le cytoplasme. En termes de signal, le marquage apparaît sous la forme d'un hypo-signal (valeurs d'intensité dans les pixels faibles par rapport au tissu non marqué). Nous supposons donc que la distribution de l'intensité de coloration des neurones du centre vers la frontière peut être modélisée sous la forme d'une distribution normale inverse. La Figure 49 montre un exemple de profil transversal de l'intensité des pixels sur une image histologique. Les bassins dans la partie droite de la figure correspondent à des neurones de différentes tailles. L'idée générale de notre approche est de réduire le bruit, d'extraire les pixels correspondant aux bassins (classe neurone) et enfin de séparer les neurones qui se touchent en proposant une stratégie adaptée à nos données.

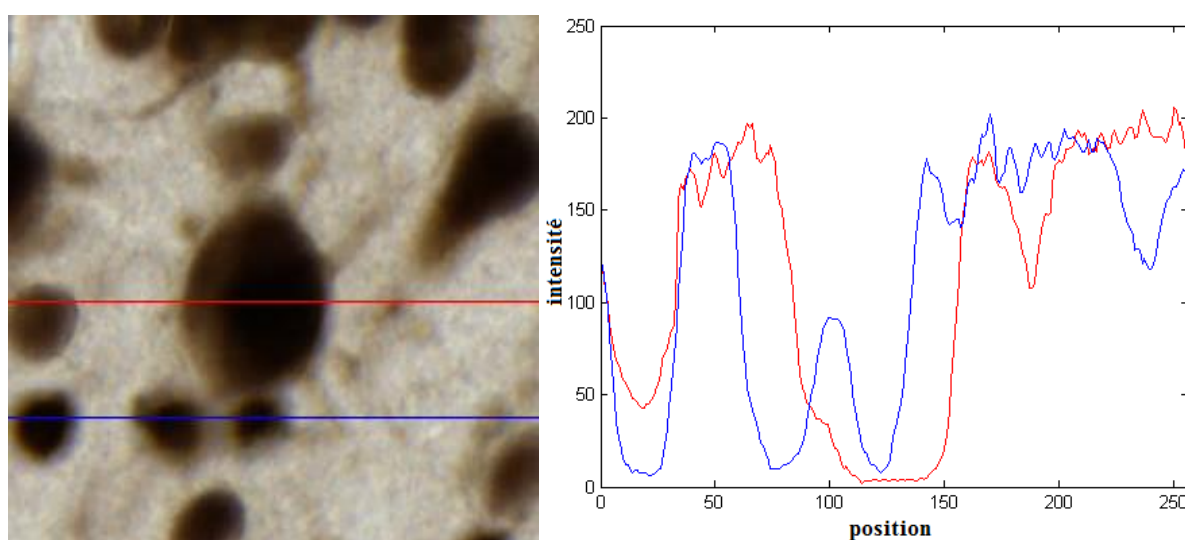


Figure 49: Exemple montrant la distribution de l'intensité des pixels sur des profils dans une image histologique. À gauche: image originale. À droite: profils transversaux de l'intensité des pixels issus de l'image originale.

Le problème important à prendre en compte dans notre travail est la variabilité de la taille des neurones. Il est important de souligner que localement, il ne peut pas être considéré que la taille des neurones est similaire. Dans certains cas, des neurones de différentes tailles se touchent (Figure 49). Il est donc nécessaire de développer des algorithmes capables de gérer cette variabilité. La Figure 50 présente notre stratégie globale de traitement permettant de réaliser l'individualisation des neurones. Par la méthode RF, l'image segmentée correspondant à un masque binaire (noté I_m) du marquage neuronal est calculée. Puis, une étape de réduction du bruit par filtrage médian de l'image I_m est mise en œuvre (filtrage introduit dans la section 2.3.5 et correspondant à la Figure 50b). À partir de l'image filtrée (notée I_{mf}), une image des composantes connexes indépendantes (Figure 50c) est calculée en 8-connexité (voisinage vertical, horizontal et diagonal en 2D). Notre algorithme travaille directement sur ces composantes connexes ce qui présente un intérêt certain pour traiter de très grandes images et possiblement profiter de techniques de parallélisation des calculs. Pour segmenter les neurones, il faut réduire le bruit, dans l'image en niveau de gris calculée comme la moyenne des composantes RGB (Hanbury, 2008) à partir de l'image originale, par un filtre gaussien dont le paramètre σ doit être adapté à la taille des neurones à traiter. La relation entre le paramètre σ et le rayon moyen des neurones sera présentée dans la section 2.4.5. Pour calculer le σ le plus adapté à la taille des neurones, nous proposons de produire tous les résultats de segmentation selon un large intervalle de valeurs possibles de σ . Ensuite, l'idée générale est de sélectionner le σ le plus

adapté en faisant l'hypothèse qu'une gamme de valeurs de σ consécutives produit des segmentations localement stables. Dans un premier temps, nous exécutons deux fois un processus de deux étapes (lissage gaussien (Figure 50d) suivi d'un filtrage min-max (Figure 50e) qui seront présentés en détail dans la section 2.4.3) en fonction de toutes les valeurs possibles de σ (section 2.4.5). Ce processus mis en œuvre pour chaque valeur de σ permet de produire une image originale dans laquelle l'information concernant les centroïdes et les contours des neurones est renforcée. En particulier, il est possible d'extraire à partir de cette image des centroïdes pouvant être considérés comme des centres des noyaux de potentiels neurones. Ensuite, à l'aide d'un modèle discret basé sur des contours déformables (section 2.4.4), les neurones peuvent être séparés et identifiés par un label unique (Figure 50f-h) selon les différentes valeurs de σ . A partir de l'ensemble des résultats obtenus, il faut définir le σ le plus adapté pour chaque neurone selon des analyses dans un premier temps en comparant le nombre de neurones détectés par la méthode avec le nombre de neurones segmentés théoriquement puis, dans un second temps, en étudiant des segmentations localement stables pour des valeurs consécutives de σ (Figure 50i-m, section 2.4.5). Les σ adaptatifs ainsi choisis permettent de générer une image des σ optimaux (Figure 50n) qui seront appliqués à l'image. Après deux itérations du processus de lissage gaussien adaptatif (Figure 50o) et de filtrage min-max (Figure 50p), le résultat final d'individualisation des neurones est obtenu en utilisant le modèle discret basé sur des contours déformables (Figure 50q).

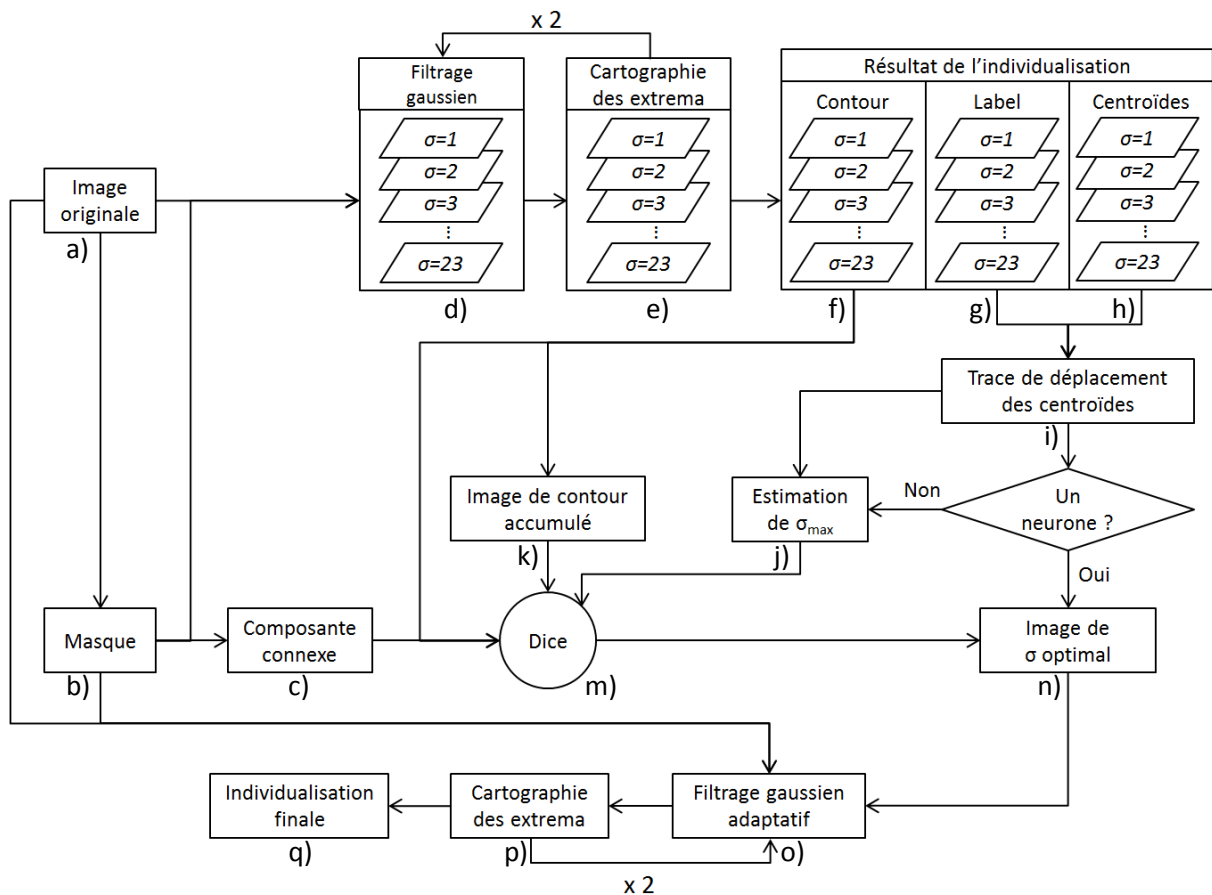


Figure 50: Stratégie globale pour réaliser l'individualisation des neurones.

Nous allons présenter en détail ces étapes une par une.

2.4.2. Segmentation du marquage NeuN par RF sur la base de données d'individualisation

En fonction du modèle de référence généré par la méthode de RF, nous avons traité la base d'images pour l'individualisation des neurones (base présentée dans la partie 2.2.2.3) afin de constituer un ensemble d'images binarisées (5000x5000 pixels) identifiant le marquage neuronal (la classe neurone) et servant de point de départ à l'étape d'individualisation.

Tout d'abord, il faut réduire le bruit dans cet ensemble d'images binarisées – le résultat de segmentation que sont les images I_m . Nous avons étudié deux filtres (filtre médian et l'opérateur de morphologie mathématique) dans la section 2.3.5. Le prétraitement final des données choisi est le filtre médian dont la taille de fenêtre sous forme de carré est 7 pixels. Après ce filtrage, les composantes connexes dont la surface est inférieure à 127 pixels, correspondant à un tiers de la surface des neurones les plus petits, sont supprimés. Nous obtenons alors les images filtrées, notées I_{mf} .

2.4.3. Détection des centroïdes des neurones par calcul d'une cartographie des extrema

Pour individualiser les neurones qui se touchent, nous proposons une nouvelle approche basée sur le calcul d'une image de cartographie des extrema (noté I_e) dans laquelle l'information relative aux centroïdes et aux contours des neurones sont rehaussés et peuvent être visualisés en même temps. Pour cela, une image en niveau de gris (noté I_g) est d'abord calculée à partir de l'image couleur originale (moyenne de R , G et B , (Hanbury, 2008)) et filtrée par un noyau gaussien avec un écart-type σ pour améliorer le rapport signal sur bruit. Cette image filtrée est noté I_{gf} .

La cartographie des extrema I_e est ensuite calculée à partir de cette image filtrée I_{gf} masquée par son image binarisée correspondante (Figure 51). Pour chaque pixel p de l'image I_{gf} , soit $D(r,p)$ le disque ne comprenant pas le pixel p et de rayon r centré sur p , soit $N(D)$ le nombre de pixels dans $D(r,p)$, soit $Max(p)$ le nombre de pixels dont la valeur de l'intensité dans $D(r,p)$ est inférieure ou égale à $I_{gf}(p)$ et $Min(p)$ le nombre de pixels dont la valeur de l'intensité est supérieure à $I_{gf}(p)$. La valeur calculée est alors:

$$I_e(p) = \frac{Max(p) - Min(p)}{N(D)} \quad (6)$$

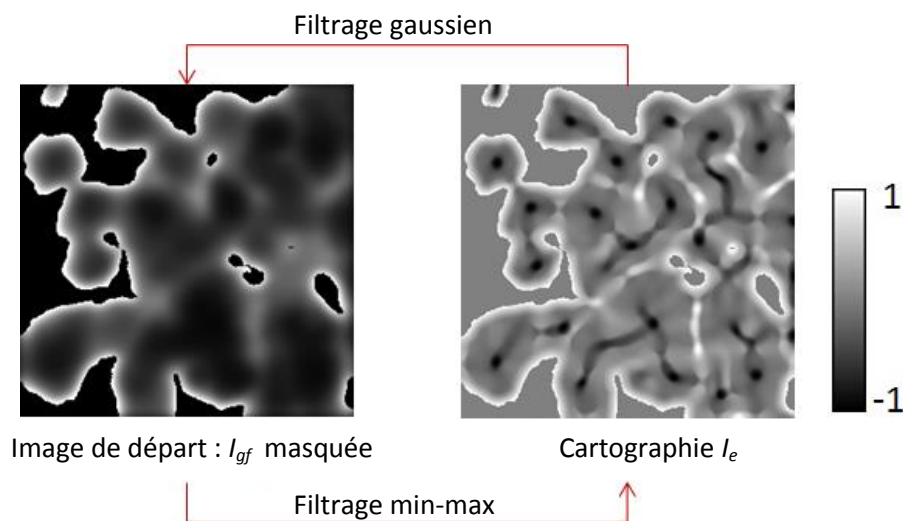


Figure 51 : Illustration de la création de la cartographie I_e .

Par conséquent, les pixels dont la valeur filtrée est -1 sont absolument des minima dans le disque local alors que ceux dont la valeur est 1 sont absolument des maxima. Les pixels de valeur -1 sont sélectionnés comme centroïdes des neurones et à chaque centroïde est attribué un label unique permettant l'identification de chaque neurone. r est fixé à 10 pixels car il correspond au rayon minimal attendu pour les neurones les plus petits ($2,5\ \mu\text{m}$). Ceci introduit un *a priori* biologique pertinent dans le traitement en évitant de détecter des minima trop proches qui ne correspondraient pas à la réalité mais plus probablement à du bruit. Nous appellerons « filtrage min-max » le calcul de la cartographie des extrema. Le processus en deux étapes (lissage gaussien et filtrage min-max) peut être répété itérativement n fois pour affiner la détection des extrema. La Figure 52 illustre les centroïdes détectés en fonction du nombre n d'itérations.

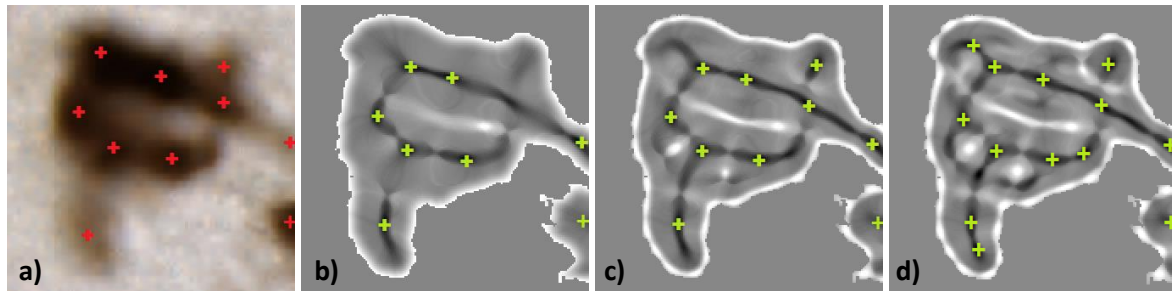


Figure 52 : Création de la cartographie des extrema à partir de l'image filtrée par filtrage gaussien ($\sigma = 4$). a) Image couleur originale. b,c,d) I_e après respectivement une, deux et trois itérations du processus (lissage gaussien et filtrage min-max). Les croix rouges sont les centroïdes correspondant aux centres des neurones marqués par un expert alors que les croix vertes sont les centroïdes détectés automatiquement par la méthode.

Dans l'exemple présenté dans la Figure 52, une estimation cohérente entre l'expert et la méthode automatisée est visuellement obtenue pour deux itérations (10 centroïdes détectés). Ceci s'explique du fait qu'une seule itération conduit à des sous-détections des neurones et trois itérations conduisent à des sur-détections. Ce paramètre permet de renforcer la détection de petits neurones, bruités ou plus clairs qui sont très difficiles à détecter.

Une étude spécifique a été menée pour optimiser le nombre d'itérations. Pour cela, nous avons testé les différentes valeurs de n comprises entre 1 et 3 et mesuré l'erreur relative de comptage (équation 2) et le F -score (équation 1) entre le calcul automatique du nombre de neurones et un comptage manuel. Cette étude sera détaillée dans la partie 2.5.1.

La Figure 53 montre une image de plus grande dimension dans laquelle les centroïdes sont visualisés sous la forme des points blancs. De façon qualitative, nous observons que la plupart des centroïdes détectés automatiquement correspondent bien à des neurones.

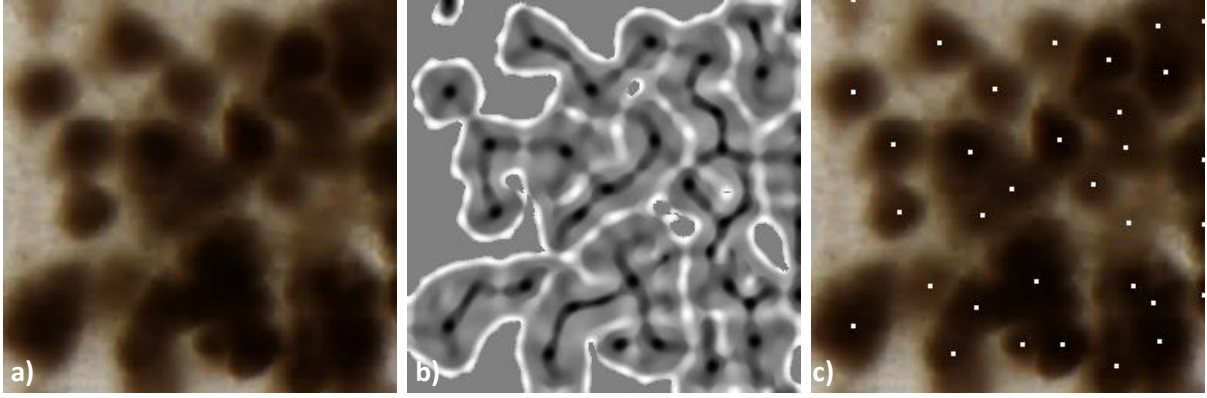


Figure 53: Détection des centroïdes des neurones par calcul de la cartographie des extrema. a) Image originale. b) Cartographie des extrema produite après deux itérations du processus de lissage gaussien ($\sigma = 4$) et de filtrage min-max. c) Centroïdes des neurones détectés automatiquement et superposés sur l'image originale.

2.4.4. Individualisation des neurones par croissance de région

L'individualisation des neurones est obtenue à l'aide d'un modèle discret basé sur l'information de contour. Des contours déformables sont initialisés sous forme de cercles de rayon prédéfini centrés sur chaque centroïde précédemment détecté. Tous les pixels à l'intérieur de ces contours initiaux sont labélisés par un identifiant (id) unique attribué à leur centroïde. Comme les contours des neurones sont proches des maxima dans I_e , nous proposons de donner à chaque point du contour une vitesse d'expansion qui dépend de sa courbure (contrôle la direction et le pas de mouvement du contour) et de la valeur de l'intensité de I_e (contrôle le pas d'avancement du contour en exploitant les valeurs dans I_e). Soit p un point du contour déformable, $\kappa(p)$ le terme dépendant de la courbure, $\mu(p)$ le terme dépendant de l'intensité, o la position du centroïde du neurone traité. La position suivante de p notée p' est obtenue par l'équation (7).

$$\overrightarrow{op'} = \overrightarrow{op} + \frac{\overrightarrow{op}}{\|\overrightarrow{op}\|} \times \kappa(p) \times \mu(p) \quad (7)$$

Soit $k(p)$ la courbure du contour au point p et c une valeur de courbure choisie par l'utilisateur. $\kappa(p)$ est obtenue par l'équation (8) pour laquelle il se produit une expansion du contour pour les courbures inférieures à c et un retour en arrière du contour pour les courbures supérieures à c . Plus la valeur de c est grande, plus le contour avance rapidement.

$$\kappa(p) = c - k(p) \quad (8)$$

Le terme $\mu(p)$ dépend de l'intensité de l'image I_e et s'inspire de (Perona and Malik, 1990). Soit t un coefficient défini empiriquement à $0,7 \times \max(I_e)$. $\mu(p)$ est calculé par l'équation (9). Quand la valeur de l'intensité du point p dans l'image I_e est supérieure à t , cette valeur est très probablement le bord d'un neurone, il faut donc réduire la vitesse d'expansion du contour déformable dans ce point, inversement.

$$\mu(p) = \exp\left(-\left(\frac{I_e(p)+1}{2t}\right)^2\right) \quad (9)$$

Pendant l'expansion du contour déformable, il faut s'assurer que l'intensité de p' dans l'image originale est plus claire que celle de p (hypothèses d'un centre de neurone plus sombre). Après chaque progression du contour déformable, une régularisation du contour déformable est réalisée

par filtrage moyen puis, une interpolation est calculée pour obtenir un contour déformable fermé. Lorsque la distance entre deux points consécutifs p et q est supérieure à une distance maximale prédéfinie d_{max} , de nouveaux points sont créés et interpolés automatiquement selon les équations (10).

$$\left\{ \begin{array}{l} N = \left\lceil \frac{\|\vec{pq}\|}{d_{max}} \right\rceil - 1 \\ \vec{op}_i = \vec{op} + \frac{\vec{pq}}{N+1} \times i \end{array} \right. \quad 1 \leq i \leq N \quad (10)$$

Ensuite, les pixels autour de chaque point p du contour déformable au sein de d_{max} sont examinés. Ceux qui sont plus proches du centroïde par rapport à p et qui ne sont pas encore labélisés sont assignés au même label que le centroïde.

Tous les contours déformables étudiés pour la même composante connexe contenant plusieurs cellules sont traités simultanément (approche compétitive) et l'intersection entre les contours déformables est strictement interdite. Dans notre expérience, nous avons fixé le nombre d'itérations pendant le processus d'expansion à 100 ce qui correspond à un entier qui assure un déplacement supérieur à la taille maximale des neurones. Avec cette valeur, les contours déformables peuvent atteindre les frontières des cellules. Une fois que le processus est terminé, les pixels restants non labélisés dans la composante connexe sont affectés au label de leurs voisins les plus proches (fenêtre en forme de carré de taille 3x3) par un vote majoritaire. Cette étape est itérée jusqu'à ce que tous les pixels de la composante connexe aient été traités.

La Figure 54 illustre un exemple de l'expansion des contours déformables des neurones. La Figure 54a est l'image originale. La Figure 54b est l'initialisation des contours déformables des neurones (les pixels à l'intérieur de chaque contour déformable initialisé ont été labélisés par les mêmes id que leurs centroïdes). La Figure 54c-k montre l'expansion des contours déformables de l'itération 1 à 9. A la fin de ce processus, les pixels non labélisés restants doivent être labélisés selon le label de leurs voisins majoritaires. La Figure 54l montre le résultat final de l'individualisation des neurones. Il faut noter que dans cet exemple, les neurones sont de petites dimensions et pour cette raison, 9 itérations semblent suffisantes pour obtenir une individualisation correcte. Cependant, dans d'autres régions du cerveau, il existe de grands neurones et dans ce cas-là, le nombre d'itérations doit être plus grand pour assurer que les contours déformables puissent atteindre les bords de tous les neurones.

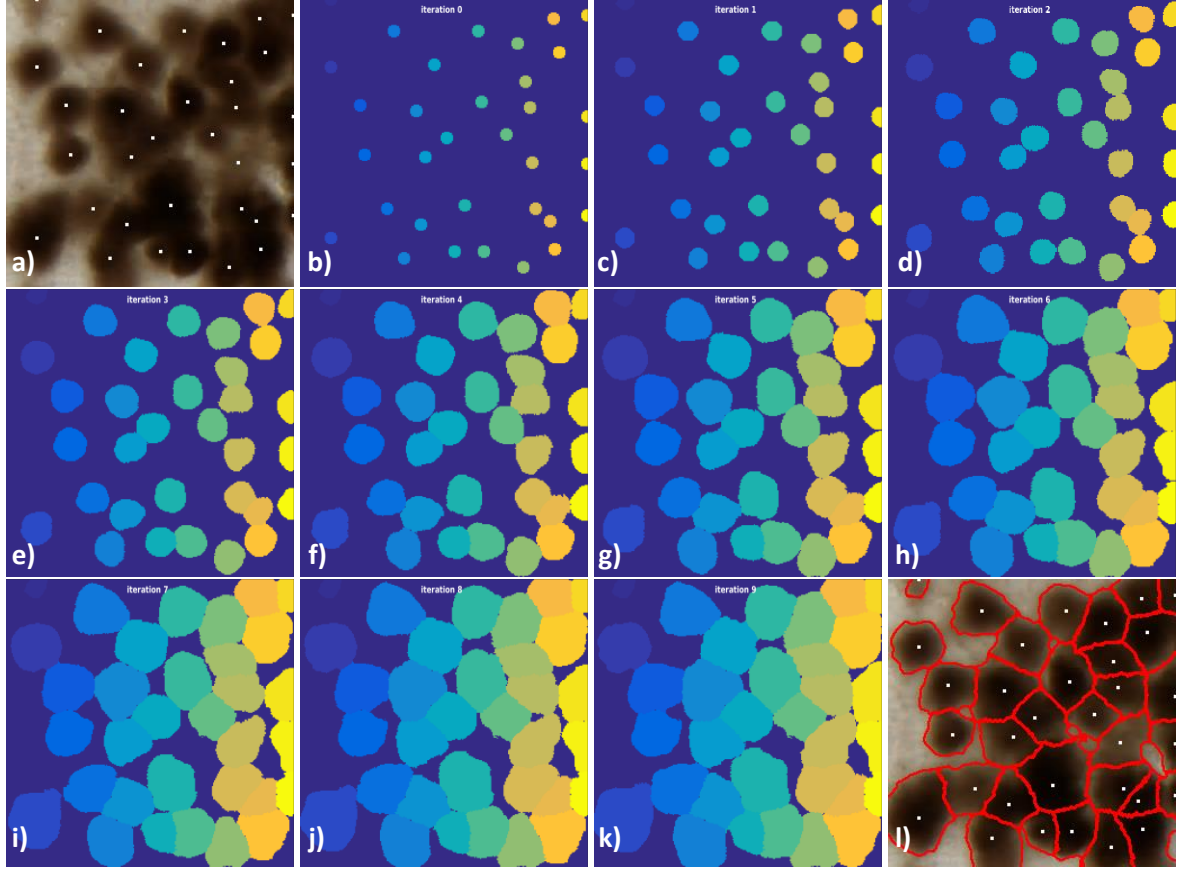


Figure 54: Illustration d'un exemple d'expansion des contours déformables de tous les neurones. a) Image originale. b) Initialisation des contours déformables. c-k) Itérations lors de l'expansion des contours déformables, de l'itération 1 à 9. l) Résultat final de l'individualisation des neurones. Les centroïdes sont illustrés pour des points blancs et les contours des neurones sont illustrés en rouge.

2.4.5. Optimisation multi-échelle de l'individualisation des neurones

Comme nous l'avons introduit précédemment, nous considérons que la taille des neurones dans le cerveau varie entre 5 et 30 μm de diamètre. Par conséquent, il faut adapter le paramètre σ du filtrage gaussien en fonction de la taille des neurones et du niveau de bruit pour améliorer la qualité de l'image en niveau de gris. Les valeurs limites de σ sont calculées en fonction des équations (11) en considérant la règle des trois sigmas par rapport au rayon des neurones.

$$\begin{cases} \Pr(\mu - 3\sigma \leq x \leq \mu + 3\sigma) \approx 0.9973 \\ 2 \times 3\sigma + 1 \leq d \end{cases} \quad (11)$$

Où x est une observation d'une distribution normale des rayons des neurones, μ est la moyenne de la distribution, σ est son écart-type et d est le diamètre de neurone. Cette règle indique que pour une distribution normale, environ 99,73% des valeurs se situent dans un intervalle centré autour de la moyenne μ dont les bornes se situent à 3σ de part et d'autre de la moyenne. En considérant que le diamètre des neurones les plus grands est de 30 μm , la valeur de σ doit varier de 1 à 23 pixels. Au vu de cet intervalle, nous avons travaillé en considérant les valeurs entières de σ ce qui constitue un nombre de tests gérable en temps de calcul avec un pas d'échantillonnage correct.

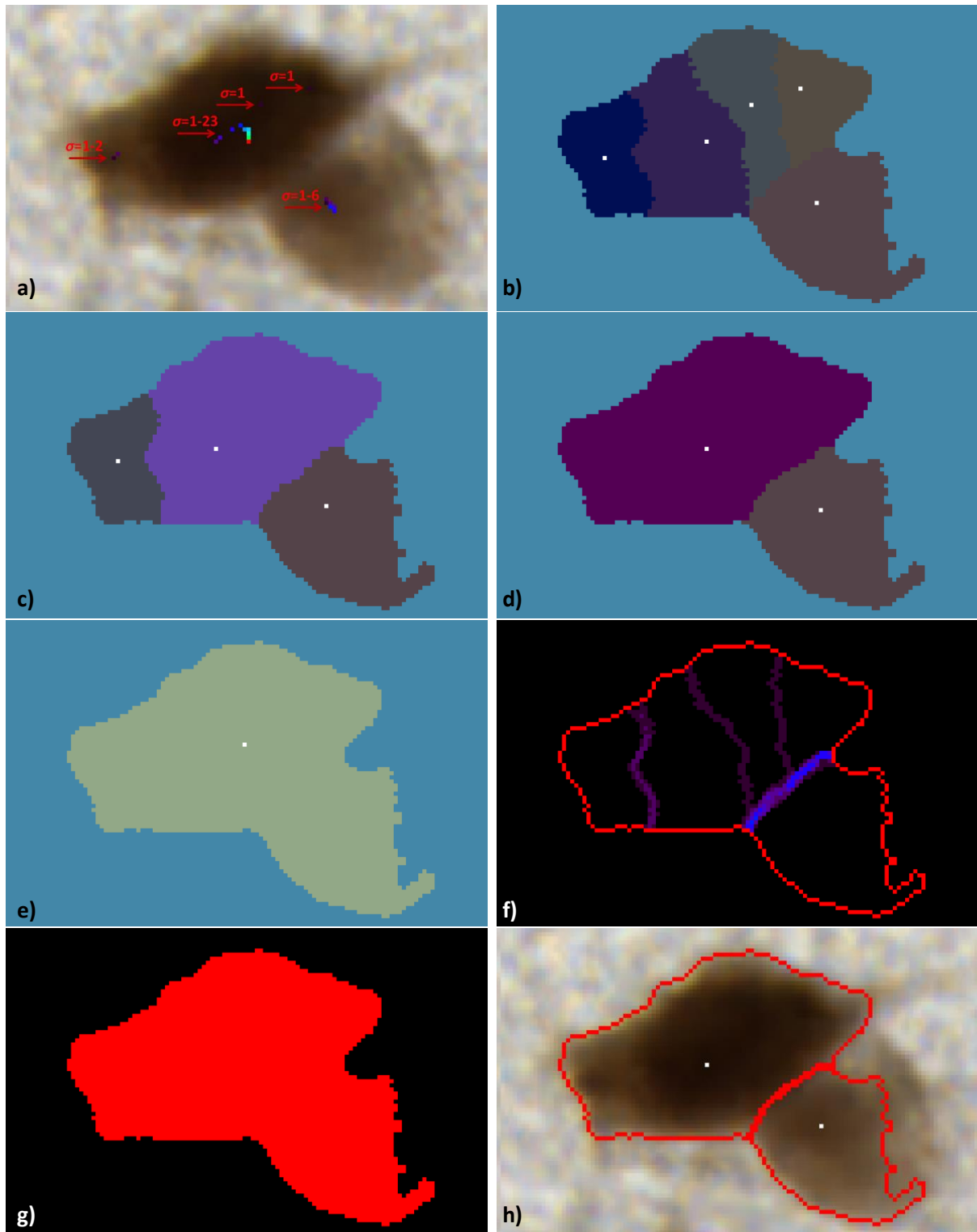
Comme *a priori* nous n'avons pas localement d'information sur la taille des neurones dans une image, il est nécessaire de proposer une méthode pour trouver le σ le plus adapté pour chaque neurone dans les différentes régions étudiées. L'idée générale est de produire tous les résultats de

segmentation selon toutes les valeurs possibles de σ en fonction desquels le σ le plus adapté est sélectionné. Pour cela, nous cherchons des valeurs consécutives de σ qui produisent des segmentations localement stables. Puis, nous regardons la reproductibilité des résultats de contours détectés par les voisins du σ détecté en utilisant un indice de similarité.

Afin de produire tous les résultats, nous exécutons deux fois le processus de lissage gaussien et de filtrage min-max pour chaque valeur de σ pour affiner les extrema dans l'image de cartographie I_e . Avec les centroïdes et le modèle discret basé sur les contours déformables, nous obtenons trois images pour chaque valeur de σ : images des labels, des contours et des centroïdes. Le σ le plus adapté va être sélectionné en fonction de ces images et des informations supplémentaires que nous allons détailler.

Dans notre travail, la valeur de σ varie de 1 à 23 pixels. La plus grande valeur de σ permet de détecter les grands neurones et produit souvent des sous-individualisations pour des amas de petits neurones accolés. De petites valeurs de σ permettent de détecter les petits neurones et produisent quant à elles des sur-individualisations pour de grands neurones. Afin de réduire l'impact du problème de sous-individualisation, nous réduisons d'abord l'espace d'exploration des valeurs de σ en cherchant un σ_{max} aussi petit que possible pour chaque composante connexe pour s'assurer de la détection des « grands neurones », soit de vrais grands neurones, soit d'amas de petits neurones accolés. Ceci est réalisé dans une composante connexe en tenant compte de la relation entre le nombre de neurones individualisés réellement estimés avec toutes les valeurs de σ (Figure 50h) et le nombre de neurones calculé théoriquement en fonction des équations (11). Il faut noter que le nombre des neurones individualisés équivaut au nombre de centroïdes détectés.

Nous étudions alors l'évolution du nombre de centroïdes détectés par composante connexe au cours de l'augmentation de la valeur de σ ce qui correspond au nombre de neurones individualisés réellement. Par exemple, la Figure 55b montre 5 neurones segmentés lorsque σ est égal à 1. Les centroïdes correspondants sont les points blancs. En augmentant σ à la valeur 2 (Figure 55c), 2 neurones disparaissent, 3 centroïdes sont détectés en blanc. Lorsque σ varie de 3 à 7, deux centroïdes sont détectés de façon stable. A partir de σ égal 8, il n'y a plus qu'un centroïde détecté (Figure 55d). Pour synthétiser ces résultats, nous accumulons par superposition tous les centroïdes détectés sur l'image originale (Figure 55a). Les points qui vont de violet à rouge (palette des couleurs de l'arc-en-ciel) correspondent aux centroïdes détectés en fonction des valeurs de σ variant de 1 à 23. Les centroïdes détectés représentent indirectement les neurones potentiellement détectés. Nous générons alors le Tableau 6 qui présente le nombre de nouveaux neurones détectés au cours de la décroissance de σ . Par exemple, pour la composante connexe 1 (cc1) présentée dans la Figure 55, quand σ est égal à 23 il n'y a qu'un neurone détecté par notre méthode. Lorsque σ continue de diminuer, il n'y a pas de nouveau neurone détecté jusqu'à σ égal à 7. A partir de σ égal à 6, un nouveau neurone apparaît. Entre σ variant de 5 à 3, il n'y a pas de changement observé. Quand σ est égal à 2, un nouveau neurone est détecté. Enfin, lorsque σ est égal à 1, deux nouveaux neurones sont détectés. Ce tableau sert à chercher le σ_{max} dans les étapes suivantes. La Figure 56 illustre un autre exemple plus complexe, dans lequel la composante connexe correspond à cc2 dans le Tableau 6.



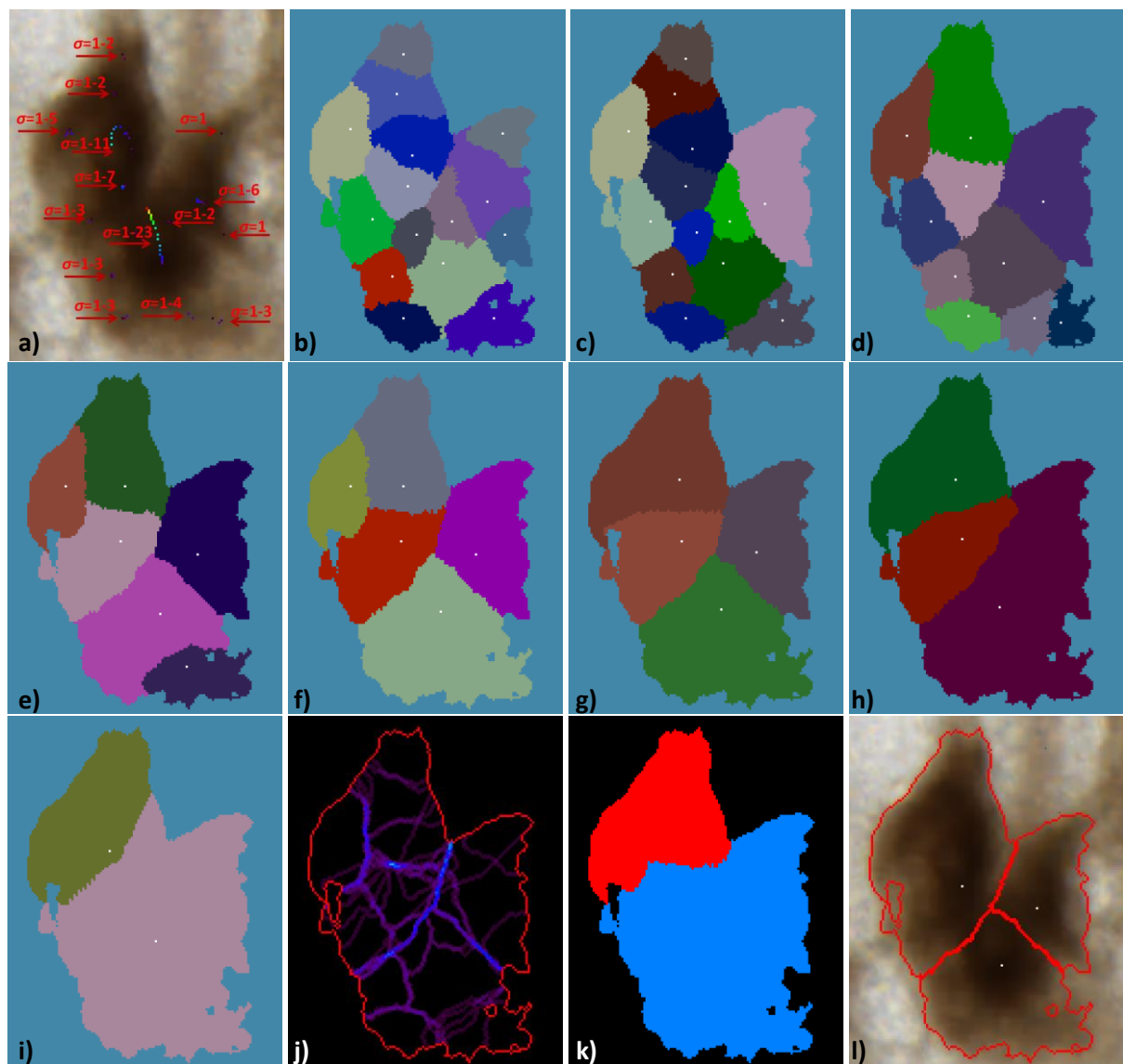


Figure 56: Illustration de l'optimisation du paramètre σ . a) Image originale fusionnée avec les centroïdes détectés en fonction des différentes valeurs de σ . b-h) Images des labels des neurones pour σ égal à 1-7. i) Les images des labels des neurones sont identiques pour σ égal à 8-11. A partir de σ égal à 12, il n'y a plus qu'un label. j) Superposition de tous les contours obtenus en fonction des toutes les valeurs de σ . La couleur violet foncé correspond à la valeur 1 et rouge correspond à la valeur 23. k) Image de σ optimal. Rouge correspond à la valeur 12 et bleu à la valeur 5. Le calcul de ces valeurs sera expliqué dans les étapes suivantes. l) Résultat final de l'individualisation. La surface totale de cette composante connexe est 10189 pixels. Les surfaces des trois neurones sont respectivement de 4686, 3371 et 2132 pixels.

Tableau 6: Evolution du nombre de nouveaux neurones détectés au cours de la décroissance de σ pour deux composantes connexes étudiées. Le tableau se lit de droite à gauche, 0 signifie que le nombre de neurones détecté est stable.

| Composante connexe | σ | | | | | | | | | | | | | | | | | | | | | | |
|-----------------------|----------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
| cc1 | 2 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| cc2 | 2 | 3 | 4 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

A partir du Tableau 6, nous observons des valeurs consécutives de 0 correspondant à des valeurs consécutives de σ . Ceci correspond à des états stables pour lesquels il n'y a pas de nouveaux neurones détectés. Par la suite, cette information sera exploitée pour définir le σ optimal.

Le σ le plus grand dans le Tableau 6 étant capable de détecter des neurones est noté σ_d (≤ 23). Selon ce tableau, nous pouvons calculer le nombre maximal des neurones détectables dans une composante connexe, soit la somme du nombre de centroïdes détectés pour σ variant de 1 à σ_d . Si la somme est égale à 1, cette composante connexe est effectivement un neurone individuel. N'importe quelle valeur de σ ($\leq \sigma_d$) peut alors être appliquée pour détecter ce neurone. Si la somme est différente de 1, cette composante connexe peut être composée de plusieurs neurones, soit de grands neurones, soit de petits neurones, soit un mélange des deux.

Ensuite, nous cherchons σ_{max} en fonction du Tableau 6. Pour chaque composante connexe traitée, nous considérons tout d'abord l'estimation du nombre de neurones (noté nb_e) obtenue avec le σ_d (première valeur de σ en partant de la droite qui détecte des neurones). Ainsi, il est possible de calculer une surface moyenne des neurones avec cette valeur. Grâce aux équations (11), en considérant cette surface moyenne par neurone, nous pouvons estimer le σ théorique pour segmenter ces neurones. Nous vérifions alors dans le Tableau 4, qui présente le cas réel, que le nombre de neurones (noté nb_r) qui seraient effectivement détectés avec ce σ théorique serait bien équivalent avec le nombre de neurones estimé (nb_e) précédemment. Si c'est le cas, l'exploration des valeurs de σ pourront se limiter de σ égal à 1 jusqu'à cette valeur théorique. Dans le cas contraire ($nb_r > nb_e$), le calcul du σ théorique va être renouvelé en considérant le nouveau nombre de neurones qui seraient détectés avec cette valeur théorique. Le processus est ainsi itéré jusqu'à obtention de la valeur de σ_{max} .

Nous pouvons reformuler plus en détail cette stratégie sous la forme suivante. Nous commençons par l'étude de σ_d (première valeur de σ en partant de la droite qui détecte des neurones). La surface de la composante connexe traitée (noté S) est accessible, le nombre (nb_e) de neurones individualisés lorsque σ est égal à σ_d est obtenu à partir du Tableau 6. Comme la même valeur σ_d sert à détecter nb_e neurones, nous faisons l'hypothèse que leurs surfaces sont similaires. Ensuite, nous pouvons calculer une nouvelle valeur σ_t théorique qui permettrait de segmenter cette composante connexe en nb_e neurones selon les équations (11). Par la suite, nous cherchons dans le Tableau 6 lorsque σ est égal à σ_t , le nombre de neurones (nb_r) qui ont été réellement individualisés. Si nb_r est supérieure à nb_e , c'est-à-dire que cette composante connexe n'est pas constituée de nb_e grands neurones mais de nb_r neurones de plus petite taille. Dans ce cas, nous continuons de calculer la nouvelle valeur σ_t de la même façon en fonction de nb_r toujours avec les équations (11). Ainsi de suite le processus est itéré jusqu'à obtenir nb_r égal à nb_e , enfin, la dernière valeur σ_t est attribuée à σ_{max} . Comme σ_{max} est inférieure à 23, cela permet de réduire l'espace d'exploration des σ .

Cette estimation de σ_{max} pour chaque composante connexe est expliquée ci-dessous en s'appuyant sur le Tableau 6 et les équations (11).

| Algorithme 1 : Estimation de σ_{max} |
|--|
| Donnée: T : Tableau 6 pour une composante connexe, S : la surface de cette composante connexe, $\sigma_d, \sigma_t, nb_e, nb_r$ |
| Résultat: σ_{max} (nombre entier) |
| <pre> $nb_r \leftarrow \sum T(\sigma \geq \sigma_d), nb_e \leftarrow 0, \sigma_t \leftarrow \sigma_d, \text{iteration} \leftarrow 0$ while ($nb_r > 0$) and ($nb_e < nb_r$) and ($\text{iteration} < 23$) iteration \leftarrow iteration + 1 $nb_e \leftarrow nb_r$ $\sigma_d \leftarrow \sigma_t$ $\sigma_t \leftarrow \left\lceil \frac{1}{6} \times \left(2 \times \sqrt{\frac{S}{\pi \times nb_e}} - 1 \right) \right\rceil$ if $\sigma_t > \sigma_d$ $\sigma_t \leftarrow \sigma_d$ end $nb_r \leftarrow \sum T(\sigma \geq \sigma_t)$ end $\sigma_{max} \leftarrow \sigma_t$ </pre> |

Ensuite, nous allons affiner localement la valeur optimale de σ au sein de la composante connexe à l'échelle des neurones. Pour cela nous allons utiliser l'information relative aux contours détectés avec les différentes valeurs de σ comprises entre 1 et σ_{max} précédemment calculé. Notre hypothèse est que le σ optimal doit correspondre à une zone pour laquelle plusieurs valeurs consécutives de σ doivent fournir une segmentation similaire. Pour accéder à cette information, nous générons une image dans laquelle les contours produits dans l'intervalle de σ donné sont accumulés (notée I_a , les Figure 55f et Figure 56j illustrent deux images I_a par accumulation de tous les contours obtenus en fonction de toutes les valeurs de σ). Les images individuelles des contours sont notées I_i . Le choix du σ optimal revient à déterminer la valeur de σ qui produit une segmentation similaire à l'image d'accumulation des contours. Pour réaliser cette tâche, nous devons comparer les segmentations individuelles avec l'image d'accumulation en utilisant un critère de similarité, ici le critère de Dice. Afin de comparer des contours binaires entre eux, nous allons produire, à partir de l'image d'accumulation, des images binarisées par seuillage (noté s) en utilisant les valeurs de seuillage allant de 1 à σ_{max} . Ces images binarisées sont noté I_{as} . Cela nous permet de décrire toutes les conformations d'accumulation possibles sans forcément privilégier les superpositions maximales correspondant aux contours extérieurs de la composante connexe traitée.

Le critère de Dice est alors calculé selon l'équation (12).

$$Dice = \frac{2 \times N_c(\sigma, s)}{N_i(\sigma) + N_{as}(s)} \quad (12)$$

Où σ et s procèdent du même intervalle de 1 à σ_{max} . N_i dépendant de la valeur de σ représente le nombre de pixels des contours d'une composante connexe dans chaque image des contours I_i , N_{as} dépendant de s représente le nombre de pixels des contours d'une composante connexe dans l'image I_{as} , et N_c , dépendant de ces deux paramètres σ et s , est le nombre de pixels des contours qui existent dans la même localisation géographique de ces deux images I_i et I_{as} .

Les Tableau 7 et Tableau 8 montrent deux exemples de tableau de coefficients de Dice qui nous aident à définir le σ le plus adapté à l'échelle des neurones. Les valeurs des coefficients de Dice varient de 0 à 1. Plus la valeur est proche de 1, plus les deux échantillons sont considérés similaires,

et inversement. Nous définissons un seuillage des coefficients de Dice, noté s_{dice} , au-delà duquel, les deux échantillons sont considérés comme étant assez similaires. Le calcul de s_{dice} est adaptatif en fonction du Tableau 6 et des tableaux de coefficient de Dice. Dans le Tableau 6, il existe des états stables qui, pour des valeurs consécutives de σ , n'entraînent pas l'apparition de nouveaux neurones. Afin d'éviter de rater de petits neurones qui ne peuvent qu'être détectés par de petites valeurs de σ , nous cherchons alors s'il existe un état stable précédent celui du σ_{max} estimé (l'état stable actuel). Si la réponse est oui, s_{dice} est la valeur moyenne des coefficients de Dice pour toutes les valeurs consécutives de σ dans l'état stable précédent celui de σ_{max} , selon les tableaux de coefficient de Dice (Tableau 7 et Tableau 8). Cela permet de détecter des coefficients de Dice maxima locaux pour de petites valeurs de σ dans cet état stable selon l'étape suivante. Si la réponse est non, s_{dice} est la valeur moyenne des coefficients de Dice pour toutes les valeurs consécutives de σ dans l'état stable actuel de σ_{max} . Si la valeur s_{dice} calculée est supérieure à 0,8, nous la fixons à 0,8, une valeur expérimentale au supérieur à laquelle les deux contours sont considérés comme étant assez similaires.

Ensuite, pour chercher une segmentation dépendant de σ similaire à l'image d'accumulation des contours, nous cherchons les coefficients de Dice qui sont des maxima locaux le long des valeurs de σ pour chaque valeur de s (présentés en gras et en rouge dans les Tableau 7 et Tableau 8) et ont des valeurs supérieures au seuillage s_{dice} . Par la suite, le long des valeurs de s , nous calculons le nombre des coefficients de Dice maxima locaux pour chaque valeur de σ . Le σ qui a plusieurs coefficients maxima locaux (au moins 2 maxima locaux qui représentent une segmentation localement stable) le long des valeurs de s et qui est également inférieur à σ_{max} est sélectionné comme le candidat de σ , le plus adapté à la taille des neurones. Par exemple, les chiffres en gras et en rouge sont des coefficients maxima locaux dans les Tableau 7 et Tableau 8, les candidats de σ sont 3, 4 et 7 pour $cc1$ et sont 5, 8, 9 et 12 pour $cc2$.

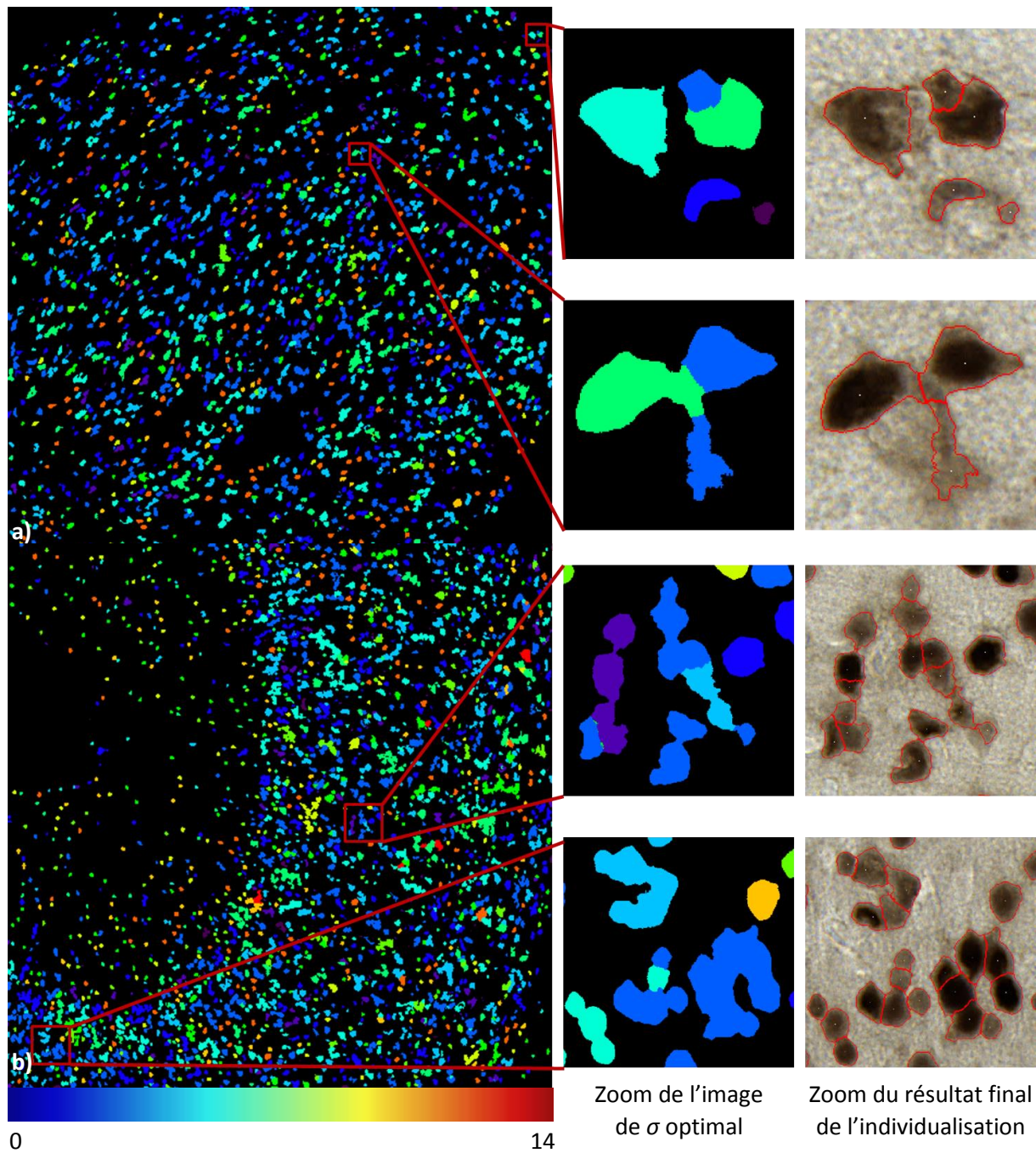
Tableau 7: Tableau de coefficients de Dice pour $cc1$. Pour chaque seuillage s , les coefficients de Dice sont identiques pour σ égal à 7 jusqu'à 23. Pour chaque σ , les coefficients de Dice sont identiques pour une valeur de seuillage comprise entre 7 et 23.

| Coefficient de Dice | | s | | | | | | |
|---------------------|---|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| σ | 1 | 0.9304 | 0,7589 | 0,7063 | 0,6854 | 0,6733 | 0,6641 | 0,6607 |
| | 2 | 0,7688 | 0,9292 | 0,8934 | 0,8746 | 0,858 | 0,8183 | 0,8146 |
| | 3 | 0,6806 | 0,8772 | 0,9348 | 0,9421 | 0,939 | 0,912 | 0,9081 |
| | 4 | 0,6806 | 0,9006 | 0,9762 | 0,9653 | 0,9525 | 0,912 | 0,9081 |
| | 5 | 0,6806 | 0,9006 | 0,9444 | 0,9554 | 0,9424 | 0,912 | 0,9081 |
| | 6 | 0,6792 | 0,8843 | 0,9268 | 0,9437 | 0,9406 | 0,9136 | 0,9097 |
| | 7 | 0,6005 | 0,8133 | 0,8908 | 0,9295 | 0,9554 | 0,9961 | 1 |

Tableau 8: Tableau de coefficients de Dice pour $cc2$. Pour chaque seuillage, les coefficients de Dice sont identiques pour σ égal à 12 jusqu'à 23. Pour chaque σ , les coefficients de Dice sont identiques pour une valeur de seuillage comprise entre 10 et 23.

| Coefficient de Dice | | s | | | | | | | | | |
|---------------------|----|--------|--------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| σ | 1 | 0,6634 | 0,7408 | 0,6359 | 0,5732 | 0,5502 | 0,5392 | 0,5241 | 0,5137 | 0,5109 | 0,5102 |
| | 2 | 0,6296 | 0,7729 | 0,6745 | 0,6123 | 0,5896 | 0,5699 | 0,5554 | 0,5427 | 0,5408 | 0,5401 |
| | 3 | 0,5716 | 0,7509 | 0,8048 | 0,7553 | 0,6984 | 0,6381 | 0,6135 | 0,6001 | 0,597 | 0,5962 |
| | 4 | 0,4932 | 0,6858 | 0,8009 | 0,8254 | 0,7872 | 0,7271 | 0,7008 | 0,6877 | 0,6844 | 0,6836 |
| | 5 | 0,4629 | 0,6999 | 0,8355 | 0,8628 | 0,8246 | 0,766 | 0,7404 | 0,7258 | 0,7224 | 0,7216 |
| | 6 | 0,4352 | 0,6393 | 0,7781 | 0,8235 | 0,8209 | 0,7916 | 0,7737 | 0,763 | 0,7595 | 0,7586 |
| | 7 | 0,4017 | 0,59 | 0,7378 | 0,8136 | 0,8315 | 0,8169 | 0,8191 | 0,8113 | 0,8078 | 0,8069 |
| | 8 | 0,3713 | 0,5031 | 0,6486 | 0,7503 | 0,8113 | 0,8597 | 0,8663 | 0,8587 | 0,8551 | 0,8542 |
| | 9 | 0,357 | 0,4994 | 0,6329 | 0,7262 | 0,7892 | 0,8436 | 0,8658 | 0,8743 | 0,8771 | 0,8779 |
| | 10 | 0,3617 | 0,4962 | 0,6413 | 0,7262 | 0,7844 | 0,838 | 0,8551 | 0,8666 | 0,8694 | 0,8701 |
| | 11 | 0,3527 | 0,4845 | 0,6263 | 0,7284 | 0,795 | 0,8502 | 0,8728 | 0,8847 | 0,886 | 0,8851 |
| | 12 | 0,2906 | 0,4871 | 0,663 | 0,7915 | 0,8793 | 0,9498 | 0,9782 | 0,9954 | 0,9991 | 1 |

Une fois que nous avons trouvé tous les candidats de σ , nous les étudions un par un, de la plus grande valeur à la plus petite pour chercher des zones de segmentation localement stable. Par exemple, pour $cc1$ dont les candidats de σ sont 3, 4 et 7, la première valeur de σ à étudier est 7. Le nombre de neurones segmentés en fonction de ce σ est 1 (obtenu à partir de Tableau 6). Alors les valeurs des pixels correspondant à cette composante connexe $cc1$ dans l'image de σ optimal est mise à jour à la valeur 7 car il n'y a qu'un neurone. Nous étudions alors σ égal à 4. La composante connexe est segmentée en 2 neurones. Nous mesurons la similarité (Dice) entre ces deux neurones segmentés et les neurones segmentés par $\sigma + 1$ égal à 5 à ce moment-là. La mesure de similarité entre deux valeurs consécutives de σ (σ venant des candidats et $\sigma+1$) permet de trouver une zone de segmentation localement stable (σ local) pour de petits neurones à l'intérieur d'une composante connexe. Nous ne tenons pas compte des neurones segmentés par $\sigma-1$ car cela pose souvent des problèmes de sur-segmentation. Si la similarité entre deux neurones produits avec σ et $\sigma+1$ est assez grande (supérieure à 0,95 qui est une valeur expérimentale), nous considérons qu'ils sont effectivement un seul et même neurone et mettons à jour les valeurs des pixels correspondant à ce neurone par la nouvelle valeur de σ dans l'image de σ optimal, et ainsi de suite pour chaque candidat de σ . De plus, en comparant les tableaux de Dice (Tableau 7 et Tableau 8) et le Tableau 6, nous observons que des candidats de σ correspondent bien à des états stables dans le Tableau 6. Par conséquent, le tableau de Dice fournit une voie pour trouver les σ les plus adaptés aux neurones accolés dans une composante connexe. Les Figure 55g et Figure 56k présentent l'image finale de σ optimal correspondant aux composantes $cc1$ et $cc2$. La Figure 57 illustre deux exemples de ce type d'image correspondant à la Figure 38a (noyau caudé) et b (cortex). Nous observons bien l'adaptation des valeurs de σ aux neurones de différentes tailles.



Il faut mentionner que le gyrus denté présente une caractéristique particulière – des milliers de neurones sont accolés dans cette région. Le tableau de coefficients de Dice dans ce cas est particulier car il ne présente généralement pas de coefficients de Dice localement maximaux pour un σ le long des valeurs de s . L'algorithme ne sait donc pas si les amas sont composés de milliers de petits ou de grands neurones. Pour simplifier le problème, nous avons fixé la valeur de σ à 5 pour le gyrus denté (la justification de cette valeur de σ sera présentée dans la section 2.5.1).

Après avoir terminé la mise à jour de l'image de σ optimal, nous effectuons le processus de lissage gaussien adaptatif et de filtrage min-max deux fois avec cette cartographie de σ . À l'aide du modèle basé sur l'approche des contours déformables, le résultat d'individualisation finale des neurones peut alors être obtenu. Les Figure 55h et Figure 56l présentent les exemples de deux résultats finaux d'individualisation.

2.4.6. Validation de la méthode d'individualisation des neurones

Une fois que les neurones ont été individualisés par notre méthode, il faut évaluer la qualité de l'individualisation produite. Cette étape est primordiale. Ici, nous présentons trois méthodes d'évaluation : par comparaison des comptages des neurones automatique/manuel, par l'étude de caractéristiques des centroïdes et des contours des populations des neurones et par stéréologie sur des RDIs anatomiques (méthode de référence en biologie).

2.4.6.1. Validation par comptage des neurones par RDI

Afin de valider la méthode d'individualisation, un score qui prend en compte la position et le nombre de neurones segmentés manuellement par les experts est défini. Pour chaque neurone individualisé automatiquement, le nombre de centroïdes marqués par les experts et contenus dans ce même neurone est calculé (Figure 58).

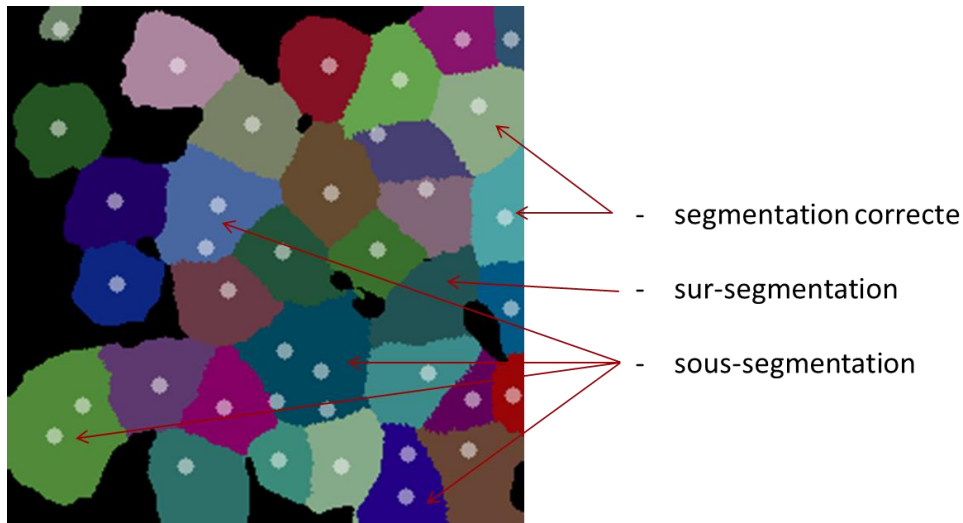


Figure 58: Illustration de la méthode de validation de la méthode d'individualisation. Les différentes couleurs définissent les neurones séparés automatiquement. Les disques en blanc représentent les centroïdes marqués par l'expert.

Si un seul centroïde de l'expert est contenu dans une cellule segmentée automatiquement, ce neurone est considéré comme un vrai positif. Sinon, ce neurone est soit sur-segmenté (aucun centroïde de l'expert présent) soit sous-segmenté (plusieurs centroïdes de l'expert présents). Les scores de Rappel (R), Précision (P) et le F -score (F) sont calculés selon les équations (13).

$$\left\{ \begin{array}{l} R = \frac{N_t}{N_e} \\ P = \frac{N_t}{N_a} \\ F = 2 \times \frac{R \times P}{R + P} \end{array} \right. \quad (13)$$

Où N_a est le nombre de neurones segmentés automatiquement. N_e est le nombre de centroïdes marqués par l'expert. N_t est le nombre de neurones qui sont segmentés automatiquement de manière correcte. Au niveau du comptage, N_t se compose de deux types de neurones segmentés : les neurones correctement segmentés (un seul centroïde de l'expert contenu dans un neurone unique automatiquement segmenté, noté N_{tc}) et les neurones automatiquement sous-segmentés par rapport à la méthode automatique (plusieurs centroïdes de l'expert contenus dans un neurone

unique automatiquement segmenté, noté N_{sous1} de valeur 1). Dans ce dernier cas, au niveau du comptage, un neurone est pris en compte parce que le code détecte bien au moins un neurone, mais les centroïdes supplémentaires détectés par l'expert correspondent quant à eux à des neurones sous-segmentés, notés N_{sous2} . Un exemple de comptage de neurones est montré dans le Tableau 9.

Tableau 9: Exemple correspondant à la Figure 58 qui montre comment nous comptons les neurones et calculons les scores.

| Description | Résultat |
|--|----------|
| N_e : nombre de centroïdes marqués par l'expert | 38 |
| N_a : nombre de neurones segmentés automatiquement | 33 |
| N_t : nombre de neurones segmentés automatiquement de manière correcte | 32 |
| N_{tc} : nombre de neurones correctement segmentés par la méthode, un seul centroïde de l'expert versus un neurone unique segmenté automatiquement | 28 |
| N_{sous1} : nombre de neurones sous-segmentés automatiquement par rapport à la méthode | 4 |
| N_{sous2} : nombre de neurones supplémentaires sous-segmentés par rapport à la segmentation de l'expert | 6 |
| N_{sur} : nombre de neurones sur-segmentés par la méthode | 1 |
| N_m : nombre de neurones non détectés mais marqués par l'expert | 0 |
| R : Rappel | 0,84 |
| P : Précision | 0,97 |
| F : F -score | 0,90 |

La relation entre ces paramètres concernant le nombre de neurones est décrite ci-dessous dans les équations 14.

$$\begin{cases} N_e = N_{tc} + N_{sous1} + N_{sous2} + N_m \\ N_a = N_{tc} + N_{sous1} + N_{sur} \\ N_t = N_{tc} + N_{sous1} \end{cases} \quad (14)$$

Grâce à cette méthode, nous pouvons calculer les scores sur notre base d'images à plus grande échelle. Cependant, cette méthode a une limite car elle requiert d'effectuer une segmentation manuelle ce qui peut s'avérer long et fastidieux lorsque nous manipulons de grandes images contenant de nombreuses cellules. Pour les grandes images, c'est en général la stéréologie qui permet de compter les neurones de la façon la plus efficace et qui constitue aujourd'hui la référence pour les questions de comptage d'objets.

2.4.6.2. Validation basée sur des caractéristiques de populations des neurones

La méthode présentée ci-dessus compte le nombre de neurones correctement segmentés par la méthode automatique par rapport au nombre de neurones qui existent réellement (comptage par l'expert). C'est une estimation globale largement utilisée dans les travaux précédemment publiés mais qui ne tient pas compte de certaines propriétés des cellules détectées par les différentes méthodes à savoir l'emplacement des centroïdes et la position exacte des contours des cellules. Ici, nous nous intéressons à la qualité de l'individualisation des neurones selon ces critères et proposons d'affiner ainsi la validation de notre méthode. Ceci implique une segmentation manuelle supplémentaire tant des centroïdes que des contours des neurones.

Comme il n'est pas possible de segmenter tous les neurones de la base d'images (5000x5000 pixels), une stratégie rendant cette étude possible a été développée. Un sous-jeu de 9 images

représentatives de cette base d'images a été extrait. Pour chacune de ces images, nous avons choisis de segmenter manuellement 100 neurones tirés aléatoirement sous forme d'une boîte englobant (Figure 59a). Nous avons jugé ce nombre suffisamment élevé pour faire des calculs statistiques et tractable en termes de temps de production. Afin de s'assurer qu'un neurone est bien présent dans chacune des boîtes englobantes, une contrainte sur la présence d'une fraction de la classe neurone a été introduite pour les différentes régions (pourcentage variant de 1 à 10%). Une interface graphique a été spécialement développée dans BrainVISA pour réaliser cette tâche et faciliter le travail de l'opérateur (Figure 59b). L'expert i1 en traitement de l'image a effectué ce travail. Sur chacune des boîtes englobantes, l'expert i1 a tracé pour le neurone le plus central son centroïde ainsi que le contour correspond de ce neurone. Ensuite, nous pouvons comparer ce neurone segmenté par l'expert avec le neurone correspondant segmenté par une méthode automatique. En première intention, nous avons comparé avec notre méthode d'individualisation mais du fait de la généricité de cette approche, nous avons pu de plus comparer les segmentations de l'expert avec une méthode de *Watershed* disponible au sein de notre laboratoire. Pour mener ces études, plusieurs paramètres ont été calculés et comparés : la distance entre les centroïdes détectés (manuel / méthodes automatiques), la distance moyenne entre les pixels des contours et le coefficient Dice. Ceci permet d'évaluer la qualité des segmentations produites par différentes méthodes automatiques par rapport aux segmentations manuelles (Figure 59c).

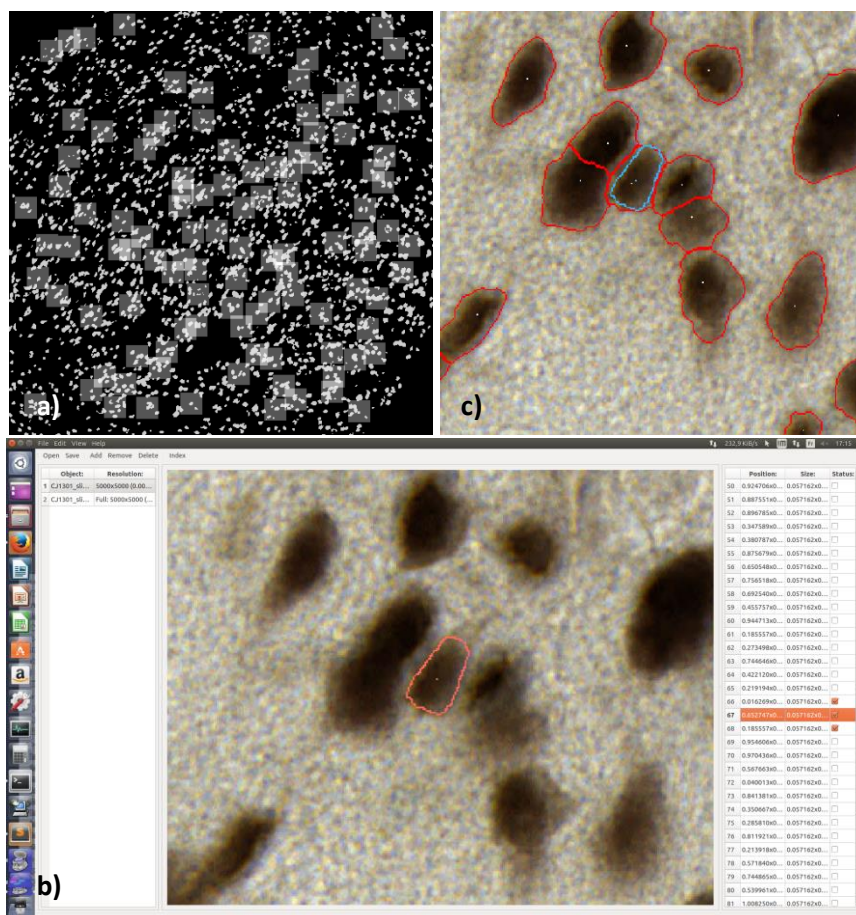


Figure 59 : a) Génération de 100 boîtes englobantes (carrés grisés) sur une image de la base d'images pour l'individualisation des neurones. b) Interface graphique développée pour permettre la segmentation manuelle par l'expert. Les 100 boîtes englobantes sont listées à droite. L'image correspondant à une boîte englobante, dans laquelle l'expert peut marquer le centroïde et dessiner le contour, s'affiche au milieu avec un centroïde et le contour associé en rouge. c) Fusion de la segmentation manuelle et de la segmentation automatique. Rouge : la segmentation automatique. Bleu : la segmentation manuelle.

Après que l'expert ait terminé de segmenter les neurones, les paramètres d'évaluation ont été calculés pour chaque image. A titre d'exemple, les résultats obtenus pour une cellule présentée dans la Figure 59b entre la segmentation d'un expert et notre méthode automatique d'évaluation sont listés dans le Tableau 10 ci-dessous.

Tableau 10 : Exemple des paramètres calculés correspondant au neurone segmenté dans la Figure 59.

| N°_boîte_englobante | Label | Distance_centroïde (pixel) | Distance_moyenne_contour (pixel) | Ecart-type_Distance_contour | Dice |
|---------------------|-------|----------------------------|----------------------------------|-----------------------------|-------|
| 67 | 1414 | 4,47 | 1,85 | 2,19 | 0,896 |

« N°_boîte_englobante » correspond au numéro de la boîte englobante affiché à droite dans l'interface graphique pour réaliser la segmentation manuelle. « Label » est le label attribué au neurone individualisé automatiquement. « Distance_centroïde » est la distance entre le centroïde détecté par l'expert et le centroïde détecté par la méthode automatique. « Distance_moyenne_contour » est la distance moyenne entre le contour dessiné par l'expert et le contour du neurone individualisé par la méthode automatique. « Ecart-type_distance_contour » est l'écart-type de la distance moyenne entre le contour dessiné par l'expert et le contour du neurone individualisé par la méthode automatique. Le « Dice » est l'indice de superposition de la surface du neurone segmenté par l'expert et du neurone individualisé par la méthode automatique.

A partir de ces paramètres, il est possible de comparer de manière quantitative deux méthodes entres elles.

2.4.6.3. Validation par comparaison avec le comptage stéréologique

La stéréologie est actuellement la technique la plus utilisée par les biologistes pour estimer le nombre de neurones dans les régions d'intérêt du cerveau. Dans cette partie, nous avons souhaité valider notre méthode par cette technique. Pour cela, nous avons demandé à deux biologistes maîtrisant les techniques de comptage stéréologique d'estimer le nombre de neurones sur la base de données dédiée au comptage stéréologique (8 régions anatomiques).

Chaque biologiste a compté deux fois les neurones de façon indépendante sur une même région anatomique directement au microscope à l'aide du logiciel Mercator (Explora Nova). Une série de régions anatomiques choisies en accord avec les biologistes experts ont été analysées (section 2.2.2.4). Sur chaque région anatomique, les biologistes ont d'abord tracé les contours de la région à étudier (Figure 60), à l'intérieur de laquelle, ils ont réalisé un échantillonnage systématique aléatoire d'un nombre de dissecteurs optiques (cadre de comptage) avec des paramètres optimaux pour obtenir une estimation non biaisée (la taille de dissecteur optique, la distance entre deux dissecteurs optiques, introduit dans la section 2.1.2). La fraction d'échantillonnage varie de 1/10 à 1/2 (les détails seront présentés dans la partie 2.5.6) dans ce type d'étude. Dans chaque dissecteur optique, il y a une frontière verte (la ligne d'acceptation) et une frontière rouge (la ligne de rejet, Figure 61) qui permettent de donner les règles de prise en compte des neurones qui les intersectent (gestion de l'effet de recouvrement partiel). Ensuite, les experts pointent les centres des neurones détectés, ne sont comptabilisés que les neurones se trouvant à l'intérieur de ces dissecteurs optiques ou touchant les frontières vertes, pas ceux touchant les frontières rouges. En travaillant directement sur le microscope, il est possible pour les experts de modifier la mise au point de la netteté en fonction de la profondeur ce qui facilite l'identification de neurones superposés.

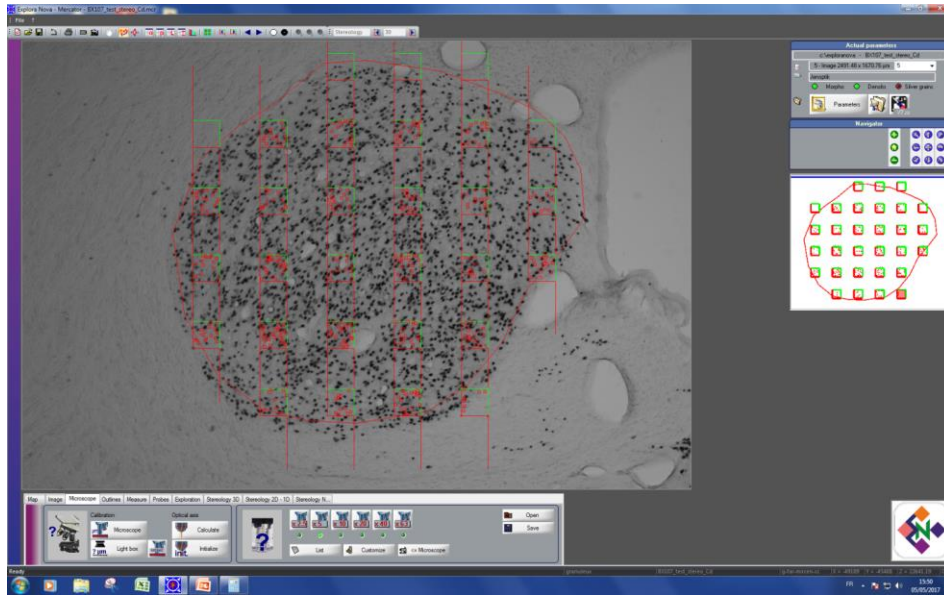


Figure 60 : Logiciel Mercator. L'expert trace d'abord une région – le contour rouge à l'intérieur duquel est la région du noyau caudé. Avec des paramètres optimisés, un échantillonnage optimisé de dissecteurs optiques est réalisé dans lesquels les neurones seront comptés par l'expert.

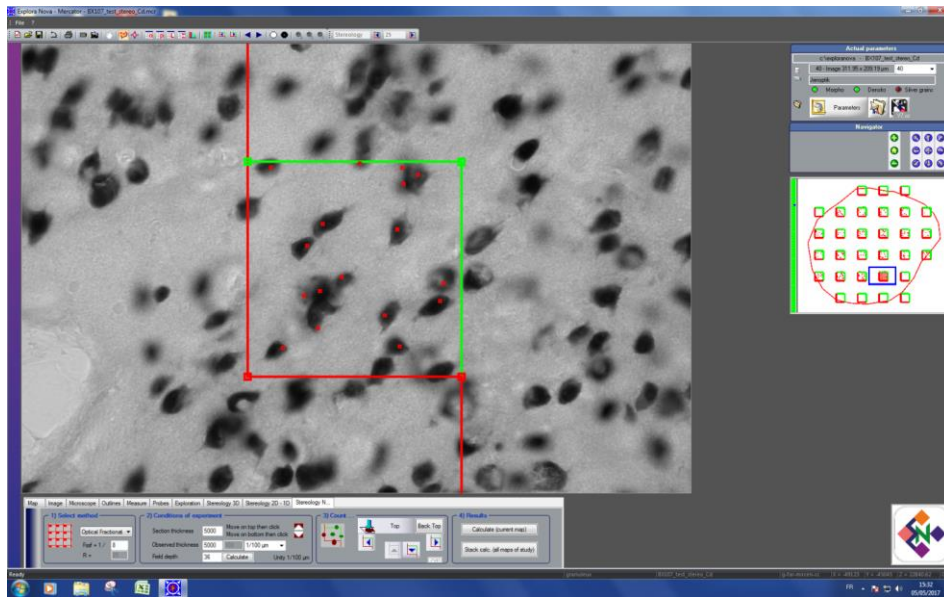


Figure 61 : Illustration du comptage des neurones dans un dissecteur optique. Les neurones sont comptés s'ils se trouvent entièrement dans le cadre de comptage ou s'ils touchent les lignes d'acceptation (verte) sans toucher les lignes de rejet (rouge).

Après avoir réalisé le comptage des neurones sur cet échantillonnage, le logiciel va estimer le nombre total de neurones dans cette région selon l'équation (3) que nous rappelons ci-dessous (l'équation 15).

$$N_{total} = \frac{1}{ssf} \times \frac{1}{asf} \times \frac{1}{hsf} \times \sum Q^- = \frac{1}{asf} \times \sum Q^- \quad (15)$$

Où ssf est la fraction des coupes, asf est la fraction de surface d'échantillonnage, hsf est la fraction d'épaisseur, $\sum Q^-$ est le nombre total de neurones effectivement comptés dans ces dissecteurs

optiques (les paramètres sont présentés dans la section 2.1.2 – Stéréologie). Ici, nous travaillons sur un cas particulier à savoir une seule coupe et non une région entière. Nous comptons donc sur toute l'épaisseur de cette coupe pour nous mettre dans la même situation que la méthode proposée, donc *ssf* et *hsf* sont toutes les deux égales à 1.

Par la suite, nous pourrions comparer les comptages stéréologiques et les comptages automatiques réalisés avec notre méthode, celle-ci ayant été appliquée sur la base des 8 régions. Pour chaque structure anatomique, nous avons tout d'abord numérisé les régions d'intérêt. Comme il existe plusieurs méthodes de numérisation, nous avons testé notre méthode automatique sur les trois principales pour évaluer le mode optimal et ainsi produit trois images par région anatomique (acquisition avec un plan unique de mise au point, acquisition en plan unique avec l'algorithme *unsharp masking* et enfin acquisition en pile Z avec l'algorithme EDF). Comme les méthodes stéréologique et automatique ne travaillent pas directement sur les mêmes données (en ligne sur un microscope pour la stéréologie versus après numérisation pour la méthode automatique), il a fallu s'assurer que les régions comparées étaient bien les mêmes. Pour cela, les structures avaient été spécifiquement choisies afin de permettre une segmentation précise et très comparable pour les deux méthodes et par les mêmes experts afin de limiter les biais (captures écrans réalisées pour superviser le tracé des experts, choix de régions facilement identifiables et segmentables). Le comptage automatique a été lancé sur les 8 régions avec 3 modes de numérisation ce qui correspond à 24 images.

2.5. Résultats et discussion

2.5.1. Optimisation des paramètres de la méthode

- Nombre d'itération n

Dans la section 2.4.3, nous avons parlé de l'optimisation du nombre d'itération n en fonction de l'erreur relative de comptage. L'opérateur *i2* – un traiteur d'image – a marqué tous les centroïdes pour tous les neurones dans la base d'images pour l'individuation (vérité terrain). Il faut noter que cette base de données se compose de 58 images, dans laquelle 8 images ont été écartées car 4 étaient floues et 4 ne comportaient pas de tissu marqué (matière blanche ou fond). Les images ôtées sont listées ci-dessous :

- CJ1301_slide91unsharp_caudate_R_2
- CJ1301_slide91unsharp_ctx_R_8
- CJ1301_slide91unsharp_ctx_R_9
- CJ1301_slide91unsharp_putamen_R
- CJ1301_slide91unsharp_vtcl_L
- CJ1301_slide91unsharp_vtcl_med
- CJ1301_slide91unsharp_WM_L
- CJ1301_slide91unsharp_WM_R

L'erreur relative de comptage a été calculée en fonction de l'équation (2) pour les 50 images. Nous synthétisons les erreurs relatives en regroupant les images correspondant à la même région anatomique (50 images regroupées en 7 régions). Les résultats moyens par région anatomique sont présentés dans le Tableau 11.

Tableau 11 : Erreur relative de comptage basée sur le nombre d'itération.

| Erreur relative (opérateur i2) | Nombre d'itérations n | | |
|-----------------------------------|-------------------------|--------------------|--------------------|
| | 1 | 2 | 3 |
| noyau caudé | 0,163±0,057 | 0,073±0,073 | 0,111±0,067 |
| claustrum | 0,520±0,051 | 0,074±0,003 | 0,028±0,015 |
| cortex | 0,248±0,055 | 0,051±0,033 | 0,048±0,045 |
| hippocampe | 0,231±0,179 | 0,186±0,161 | 0,276±0,221 |
| putamen | 0,427±0,183 | 0,085±0,081 | 0,066±0,086 |
| subiculum | 0,185±0,023 | 0,033±0,026 | 0,047±0,028 |
| thalamus | 0,302±0,050 | 0,167±0,061 | 0,217±0,063 |
| Moyenne | 0,257±0,131 | 0,105±0,114 | 0,137±0,167 |

A partir de ce tableau, nous observons que pour toutes les images, la valeur moyenne de l'erreur relative la plus petite ($0,105 \pm 0,114$) est trouvée pour un nombre d'itération égal à 2. Si nous regardons les résultats région par région, nous trouvons que dans les régions du claustrum et du putamen, l'erreur relative est plus petite pour 3 itérations. Pour le cortex, l'erreur relative est assez similaire pour 2 et 3 itérations. L'examen visuel de l'image montre que dans ces régions, il y a beaucoup d'amas constitués de petits neurones clairs, flous qui sont mieux détectés avec un nombre d'itérations plus grand. La Figure 62. a1-c1 présente les résultats d'individualisation pour une agrégation de neurones lorsque le nombre d'itérations varie de 1 à 3. Le plus grand nombre d'itérations peut correctement séparer les neurones dans certains cas. Cependant, cela induit souvent des problèmes de sur-segmentation (Figure 62. a2-c2).

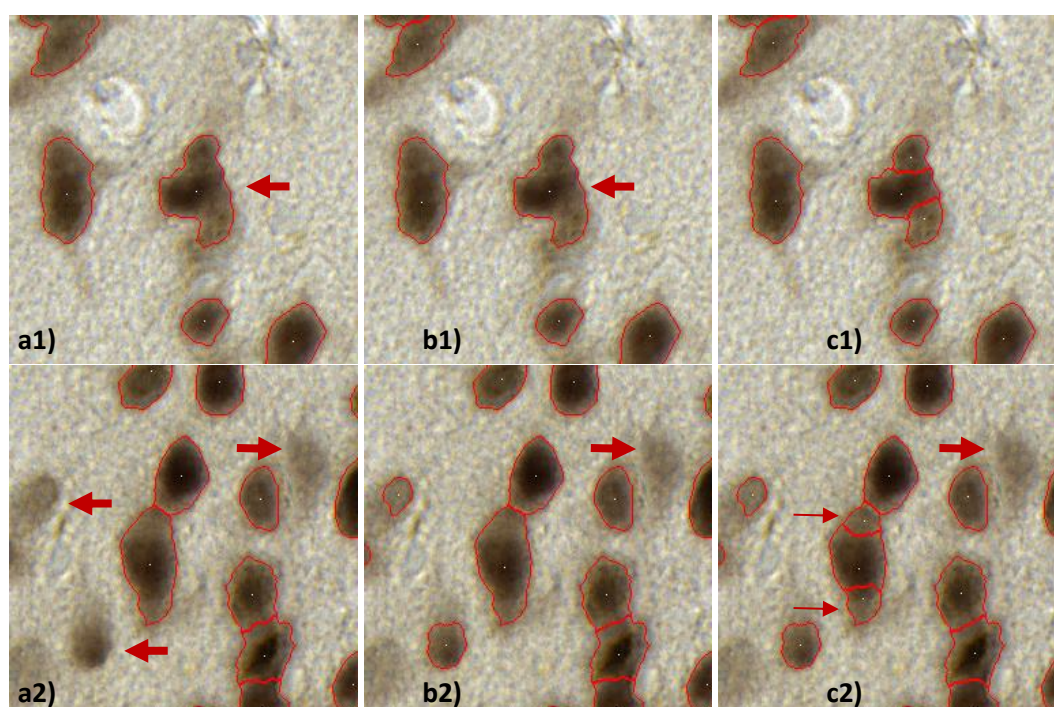


Figure 62: Comparaison des résultats d'individualisation en fonction du nombre d'itérations dans la région du putamen. a1-a2) Résultats d'individualisation pour un nombre d'itération égal à 1. b1-b2) Résultats d'individualisation pour un nombre d'itération égal à 2. c1-c2) Résultats d'individualisation pour un nombre d'itération égal à 3. Les flèches fines pointent des sur-segmentations et les flèches épaisses pointent des sous-segmentations.

De plus, nous avons également étudié l'optimisation de ce paramètre n avec un autre critère : le F -score (Figure 63) selon les équations (1). Pour une itération, on observe une sous-détection des

neurones très nette. Les résultats obtenus pour 2 et 3 itérations sont relativement proches même si la précision calculée pour 2 itérations est meilleure.

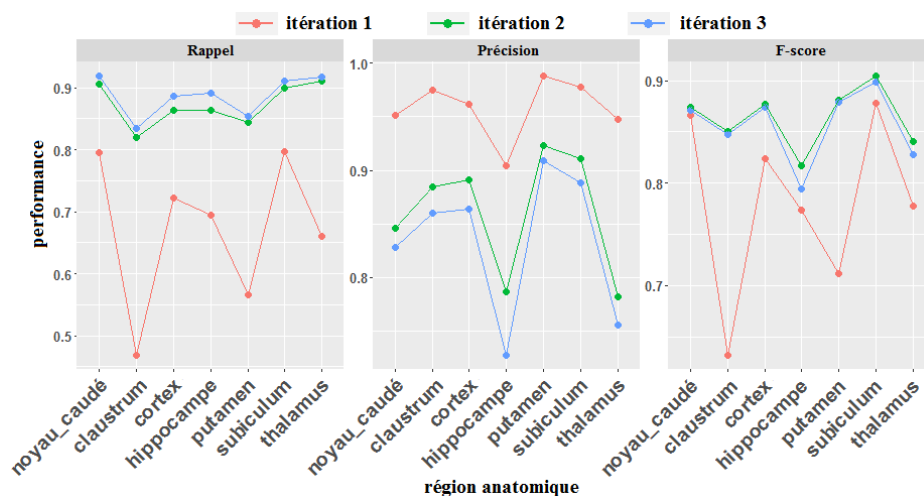


Figure 63 : Comparaison des performances de comptage en fonction de nombre d'itérations.

En ce qui concerne le temps de calcul, ce dernier est directement liée au nombre de neurones présents dans l'image. Nous avons testé notre méthode sur une image concernant la région du noyau caudé issue de la base d'images pour l'individualisation. Dans cette image, deux milliers de neurones ont été détectés par l'opérateur i2. Le temps d'exécution pour obtenir le résultat final d'individualisation pour cette image pour un nombre d'itération variant de 1 à 3 est respectivement 45, 89 et 106 minutes. Le temps d'exécution augmente au fur et à mesure que le nombre d'itération augmente.

En considérant les deux paramètres testés (erreur relative et *F-score*) ainsi que le temps de calcul nécessaire, nous avons décidé de fixer le nombre d'itérations à 2 pour la suite de notre travail. Une étude plus poussée pourrait sans doute permettre d'affiner le choix du nombre d'itération en fonction des régions anatomiques étudiées. Cependant, il serait difficile d'assurer ainsi une qualité optimale, le choix de $n=2$ nous a paru être le plus robuste.

- Optimisation de la valeur de σ pour le gyrus denté

Dans la section 2.4.5, nous avons parlé d'une optimisation spécifique de la valeur de σ pour la région du gyrus denté. En effet, pour cette région, une composante connexe extrêmement grande est détectée car elle agrège plusieurs milliers de neurones ce qui la rend difficile à traiter par notre méthode. Dans la base d'images pour l'individualisation, nous avons sélectionné les images (6 sur 58) qui contiennent les régions du gyrus denté. Ensuite, en fonction des images binarisées identifiant le marquage NeuN calculées par la méthode de RF, nous avons généré de nouvelles images binarisées en ne gardant que les régions du gyrus denté. Avec ces nouvelles images binarisées, nous avons obtenu le résultat final d'individualisation et la vérité terrain, déterminée par l'opérateur i2, uniquement pour les régions du gyrus denté. Nous avons ensuite calculé le *F-score* et l'erreur relative de comptage basés sur la vérité terrain dans le gyrus denté en fonction de différentes valeurs de σ . La Figure 64 illustre les résultats obtenus pour ces deux critères. Les différentes couleurs représentent différentes images qui contiennent des parties du gyrus denté. Le *F-score* maximal moyen est de 0,885 et l'erreur relative minimale moyenne est de 7,7% pour σ égale à 5.

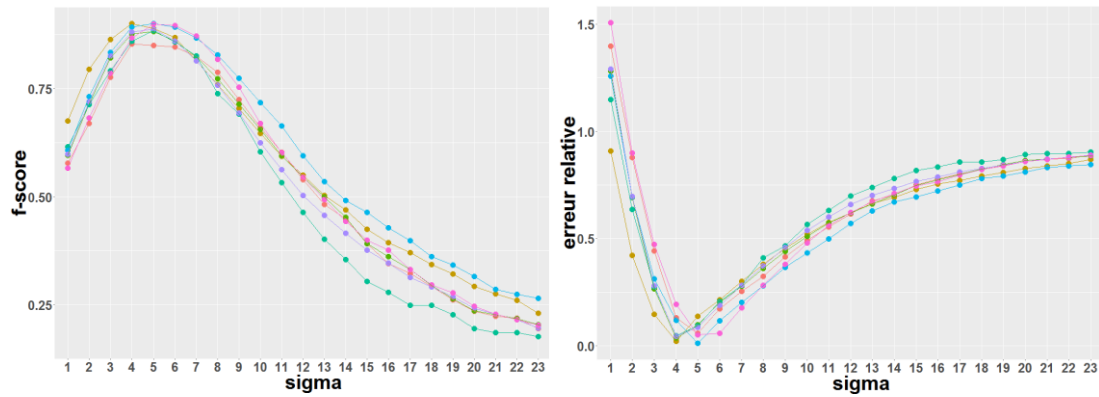


Figure 64 : Critères d’optimisation en fonction des différentes valeurs de σ pour la région du gyrus denté. Les différentes couleurs représentent les images qui contiennent des régions du gyrus denté.

La nécessité de traiter le gyrus denté de façon spécifique se justifie du fait qu’il représente un cas particulier avec une agrégation massive de très nombreux neurones de petite taille. La taille de gyrus denté à gauche de la coupe 91 est de l’ordre de 6×10^6 pixels. En pratique, l’identification par notre méthode des composantes connexes concernées est très simple par rapport aux autres composantes connexes qui sont significativement plus petites.

2.5.2. Données computationnelles

Dans la Figure 50, nous avons présenté notre stratégie globale pour individualiser les neurones. Nous résumons ici dans le Tableau 12 le temps d’exécution et la mémoire nécessaire dans chaque étape pour individualiser deux milliers de neurones (détectés manuellement par l’opérateur i2) dans une image issue de la base d’images pour l’individualisation concernant la région du noyau caudé.

Tableau 12 : Temps d’exécution et mémoire vive nécessaire pour l’individualisation des neurones dans l’image intégrant la région du noyau caudé. Ce test concerne deux milliers de neurones détectés par l’opérateur i2. «*» représente le nombre d’itérations et «#» représente 23 valeurs de σ . La taille de l’image originale est de 75 Mo.

| Etape de traitement dans la Figure 50 | Parallélisable | Temps d’exécution (s : second, min : minutes) | Mémoire nécessaire (Mo) |
|--|----------------|---|-------------------------|
| Calcul des composantes connexes (c) | Oui | 3 s | 125 |
| Filtrage multi-gaussien (d) | Oui | 90 s \times 2 (*) | 200 \times 23 (#) |
| Calcul de cartographie min-max (e) | Oui | 9 min \times 2 (*) | 200 \times 23 (#) |
| Calcul des résultats d’individualisation (f, g, h) | Oui | 3 min | 500 \times 23 (#) |
| Génération d’image de contour accumulé (k) | Non | 22 s | 600 |
| Calcul d’image de σ optimal (i, j, n) | Non | 77 s | 3175 |
| Filtrage gaussien adaptative (o) | Non | 1473 s \times 2 (*) | 325 |
| Calcul de cartographie min-max (p) | Non | 325 s \times 2 (*) | 325 |
| Calcul de résultat final d’individualisation (q) | Non | 60 s | 475 |

Le temps de traitement pour cette image est de 89 minutes pour l'individualisation de deux milliers de neurones. Nous avons lancé notre méthode sur l'ensemble de la base d'images pour l'individualisation sur un ordinateur sous Linux 64 bits (CPU: Intel Xeon E5-2630 v3 à 2,4 GHz, RAM : 128 Go). 24 des 32 noyaux de calcul ont été utilisés pour paralléliser les calculs des 58 images de 5000x5000 pixels (module *Soma-Workflow* de BrainVISA, parallélisation CPU). Le temps moyen d'exécution pour obtenir les résultats finaux d'individualisation pour un nombre d'itération égal à 2 était de 15,6 heures, soit 16,1 minutes par image.

2.5.3. Visualisation des résultats de l'individualisation

Nous avons appliqué notre méthode d'individualisation des neurones à l'ensemble de la base des images pour l'individualisation. En raison de la grande taille des images (5000x5000 pixels), nous illustrons nos résultats sur de petites parties d'images extraites (une imagerie par région). La visualisation de l'individualisation est effectuée en superposant les contours des neurones détectés directement sur l'image couleur. La complexité de la tâche d'individualisation est directement liée à la densité des neurones dans les images. Afin d'illustrer la variabilité des cas traités, nous avons utilisé le critère de la fraction d'occupation de la classe neurone dans les images. La Figure 65 présente différents niveaux de fraction d'occupation (variant de 2,6% à 26,5%) de la surface des neurones par rapport à la surface de l'image. Les images dans la Figure 65. a-f représentent des cas avec quelques neurones individuels et un petit nombre de neurones qui se touchent. La Figure 65g représente un cas de marquage extrêmement dense dans lequel de très nombreux neurones se touchent dans le gyrus denté (cas particulier déjà abordé dans la section 2.4.5). Pour le visualiser plus clairement, nous avons changé la dynamique de couleur de cette même image dans la Figure 65h car le marquage est si dense qu'il n'est pas possible de discerner les neurones individuels. La méthode d'individualisation automatique proposée fournit visuellement de bons résultats.

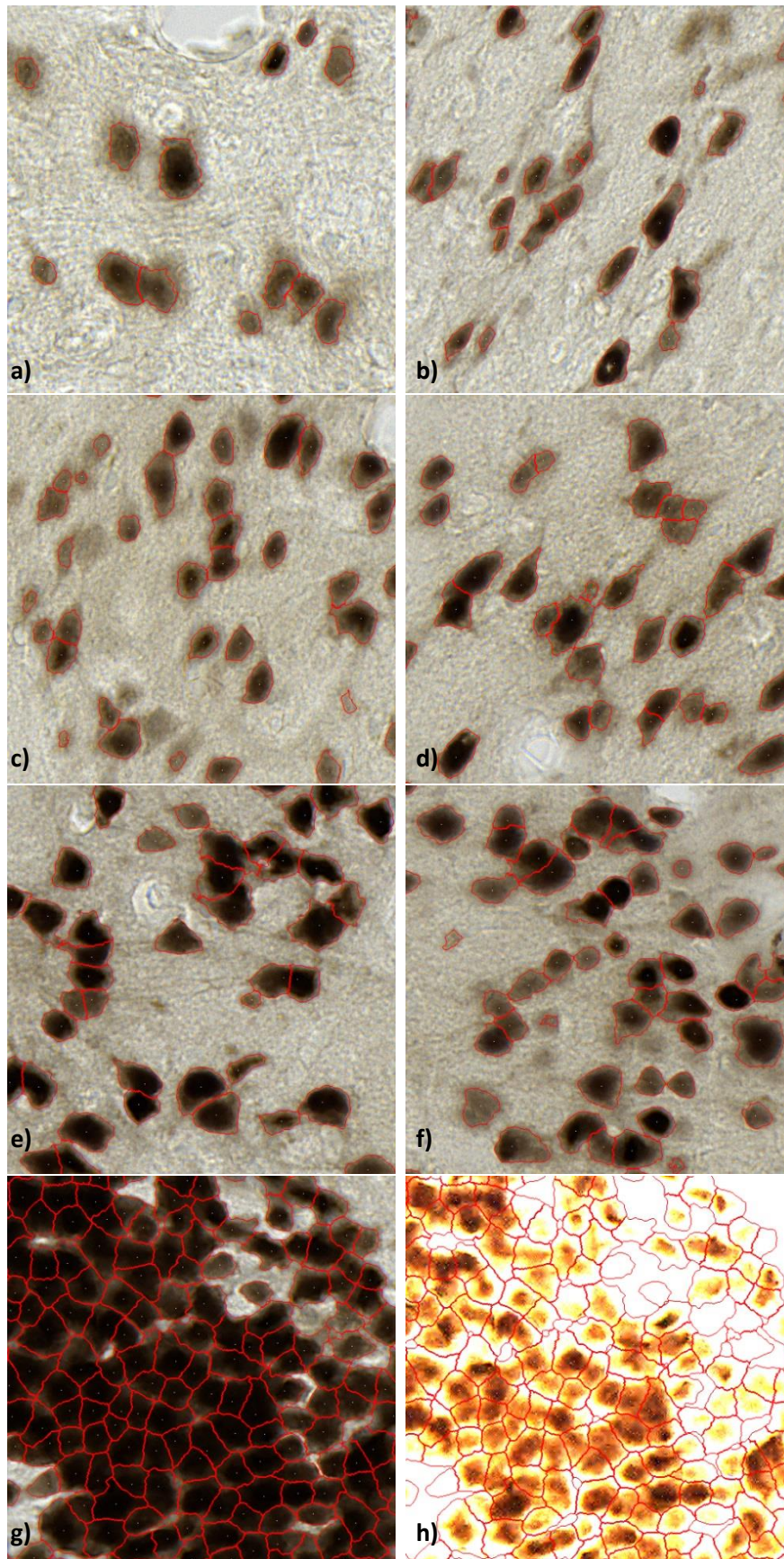


Figure 65 : Visualisation de quelques résultats d'individualisation (imagerie de 512x512 pixels extraites à partir des résultats pour la base de données pour l'individualisation. a) Thalamus. b) Claustrum. c) Putamen. d) Noyau caudé. e) Subiculum. f) Cortex. g) Gyrus denté. h) Gyrus denté après changement de la dynamique de couleur.

Malgré les bons résultats d'individualisation que nous avons obtenus dans la plupart des images, il y a des cas particuliers où notre méthode ne peut pas bien segmenter les neurones. En identifiant ces cas par rapport aux données fournies par les experts, nous avons pu identifier un certain nombre de situations dans lesquelles notre méthode peut être mise en défaut. La Figure 66 montre des échecs de notre méthode concernant des neurones qui sont atypiques.

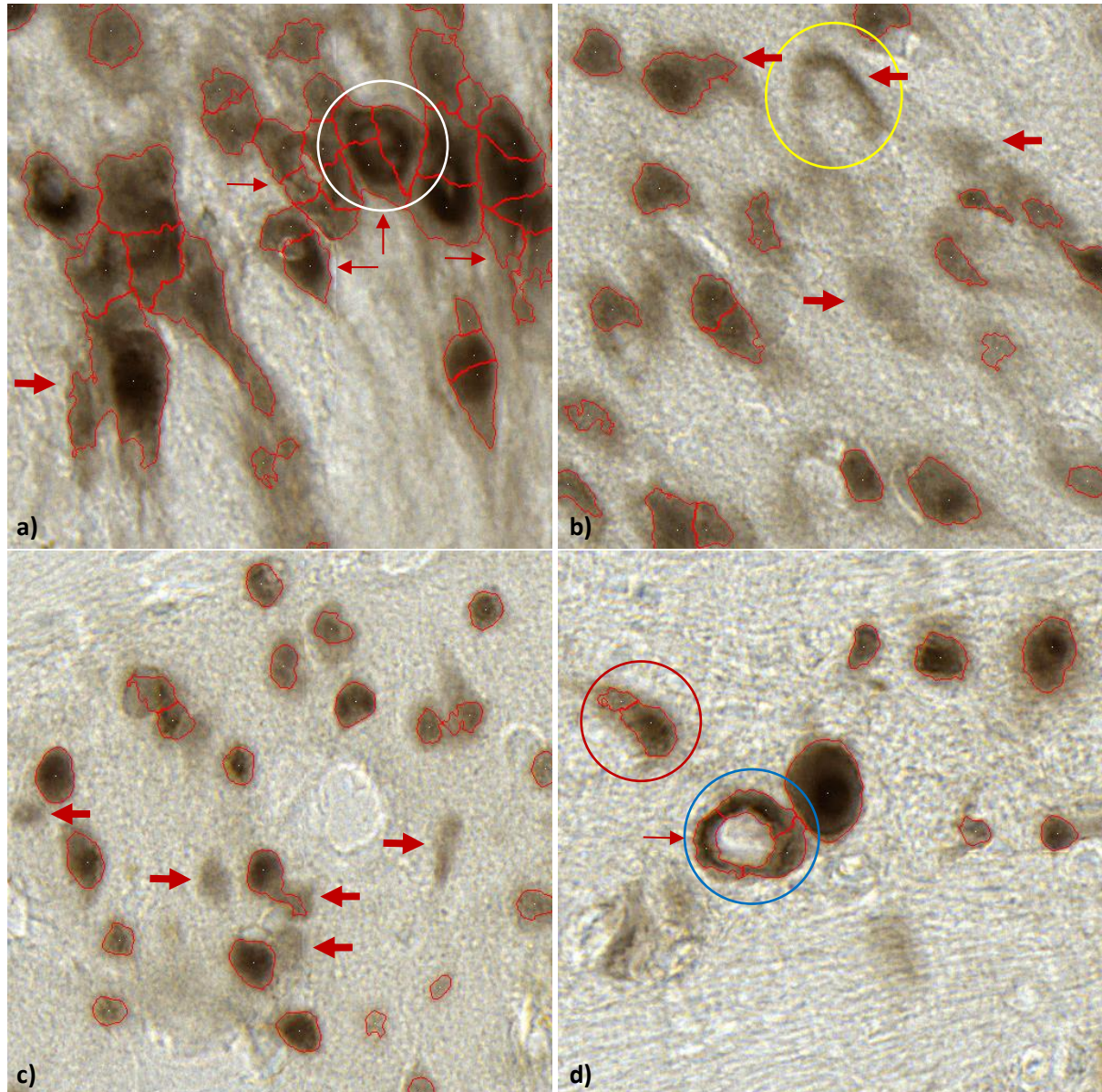


Figure 66 : Illustration de quelques problèmes d'individualisation des neurones – sur-individualisation (flèche fine) et sous-individualisation (flèche épaisse). a) CA2 dans l'hippocampe. b) CA3 dans l'hippocampe. c) Putamen. d) Thalamus.

Le neurone entouré par un cercle blanc (Figure 66a) présente une coloration différente de celle des autres neurones. La couleur au centre de ce neurone est plus claire ce qui induit souvent un problème de sur-segmentation. Le neurone entouré par un cercle jaune (Figure 66b) présente une forme en arc ce qui cause souvent une sous-segmentation lors de la segmentation ou une sur-segmentation lors de l'individualisation. Le neurone entouré par un cercle bleu (Figure 66d) présente une forme d'anneau ce qui conduit également à une sur-segmentation (cas similaire à Figure 66a). La présence de ces types de neurones est difficile à expliquer. Il est possible que dans certains cas le

plan de coupe fasse que nous n'avons qu'une partie de la cellule et que du coup que le noyau ne soit pas présent d'où cet aspect plus clair. Quand ces cas se produisent, notre méthode fonctionne difficilement car les hypothèses sur lesquelles se base notre méthode ne sont pas respectées. Dans la Figure 66c, nous voyons des neurones qui possèdent une forme relativement « classique » mais qui ne sont pas détectés par notre méthode. Ceci se produit du fait que la méthode de segmentation ne détecte pas ces neurones clairs qui présentent une intensité très faible et proche de celle du tissu non marqué. Dans ce cas, il est légitime de se poser la question de la pertinence de segmenter ces neurones. Cela mène à des sous-segmentations qui diminuent la performance de notre méthode selon les critères que nous avons testés car les experts ont pris en compte ces neurones. Une première solution serait d'améliorer les protocoles de production des coupes avec un marqueur plus spécifique ou de modifier la technique d'acquisition pour augmenter le contraste. Un doctorant de notre laboratoire (Clément Bouvier) travaille actuellement sur une optimisation des algorithmes de segmentation en explorant les combinaisons optimales de caractéristiques à utiliser. Un autre problème, est dû au fait que nous travaillons avec des images 2D, ce qui augmente la difficulté de vérification de notre méthode. Par exemple, il est impossible de déterminer si l'objet entouré par le cercle rouge (Figure 66d) représente un ou deux neurones, ce qui impacte l'exactitude de la validation de notre méthode. Ceci résulte de notre choix de travailler sur un plan de numérisation unique guidé tant par le temps d'acquisition que par le volume des données à traiter. En effet, un objectif à moyen terme sera de traiter des coupes entières voire des cerveaux entiers ce qui correspondra à des volumes de données très importants. L'acquisition de plusieurs plans, si elle est envisageable, nécessiterait de modifier notre méthode et générerait de nouveaux problèmes en particulier pour combiner les résultats des différents plans analysés.

Malgré ces différents problèmes qui ne représentent qu'une petite fraction des détections totales de neurones (principalement concentrées dans les régions CA1-3 de l'hippocampe), la segmentation et l'individualisation de la plupart des neurones ont permis de valider notre hypothèse de travail et ont démontré le potentiel de notre stratégie de traitement des coupes histologiques. Nos travaux ont été réalisés sur 3 coupes histologiques de cerveau de macaque ce qui peut sembler peu si on ne tient pas compte du nombre de pixels qu'elles représentent. L'extension de ce travail au reste du jeu de données disponible pour ce macaque donnera l'occasion de tester la robustesse de notre approche.

2.5.4. Comptage des neurones par RDI

Cette partie présente l'évaluation de la performance de notre méthode à l'aide de l'étude sur des paramètres calculés selon la vérité terrain réalisée par l'opérateur i2 (segmentation des centroïdes sur les 50 images de la base d'images pour l'individualisation). Nous rappelons les principaux paramètres étudiés ci-dessous :

- fraction : fraction d'occupation des neurones segmentés par RF par rapport au nombre total de pixels dans les images de 5000x5000 pixels.
- N_{cc} : nombre de composantes connexes.
- N_e : nombre de neurones segmentés par l'opérateur i2 (présenté dans la section 2.4.6.1).
- complexité : la complexité (noté C) de la zone ($C = N_e/N_{cc}$), plus elle est grande, plus le nombre de neurones qui sont accolés est important.
- N_g : nombre de neurones individualisés par notre méthode.
- N_t : nombre de neurones individualisés automatiquement de manière correcte ($N_t = N_{tc} + N_{sous1}$).
- N_{tc} : nombre de neurones automatiquement segmentés correctement, un seul centroïde de l'expert versus un neurone unique segmenté automatiquement

- $N_{s\text{ous}1}$: nombre de neurones sous-individualisés par la méthode.
- $N_{s\text{ous}2}$: nombre de neurones sous-individualisés supplémentaires par rapport à la segmentation de l'opérateur i2.
- N_{sur} : nombre de neurones sur-individualisés par la méthode.
- N_m : nombre de neurones non détectés mais marqués par l'opérateur i2.
- R : Rappel.
- P : Précision.
- F : $F\text{-score}$.

Cette partie de notre travail a produit une grande quantité de résultats chiffrés. Pour synthétiser les résultats, nous avons calculé les valeurs moyennes pour chaque paramètre en regroupant des images issues de la même région anatomique, ces résultats sont listés dans l'annexe 3 – « Valeurs moyennes des différents paramètres des données d'individualisation par région anatomique ». Par la suite, nous présentons les résultats moyens par région anatomique que nous avons jugé les plus intéressants.

La Figure 67 montre les résultats préliminaires obtenus en fonction de différents paramètres d'intérêt et des principales régions anatomiques étudiées. Pour rappel, RF représente l'algorithme de *Random Forest*. b1 représente un opérateur neurobiologiste, i1 représente un opérateur traiteur d'image. Les deux opérateurs ont réalisé la segmentation manuelle des 100 imagerie de taille 512x512 pixels sous la forme de 3 classes. Les classes ainsi obtenues et issues des 54 images ont servi de base d'apprentissage pour générer le modèle de référence RF qui a été appliqué à la base d'images pour l'individualisation pour identifier la classe neurone. Les résultats d'individualisation présentés dans la Figure 67 sont basés sur deux modèles :

- RF(b1) : *Random Forest* utilisant le modèle créé avec la base d'apprentissage produite par b1,
- RF(i1) : *Random Forest* utilisant le modèle créé avec la base d'apprentissage produite par i1.

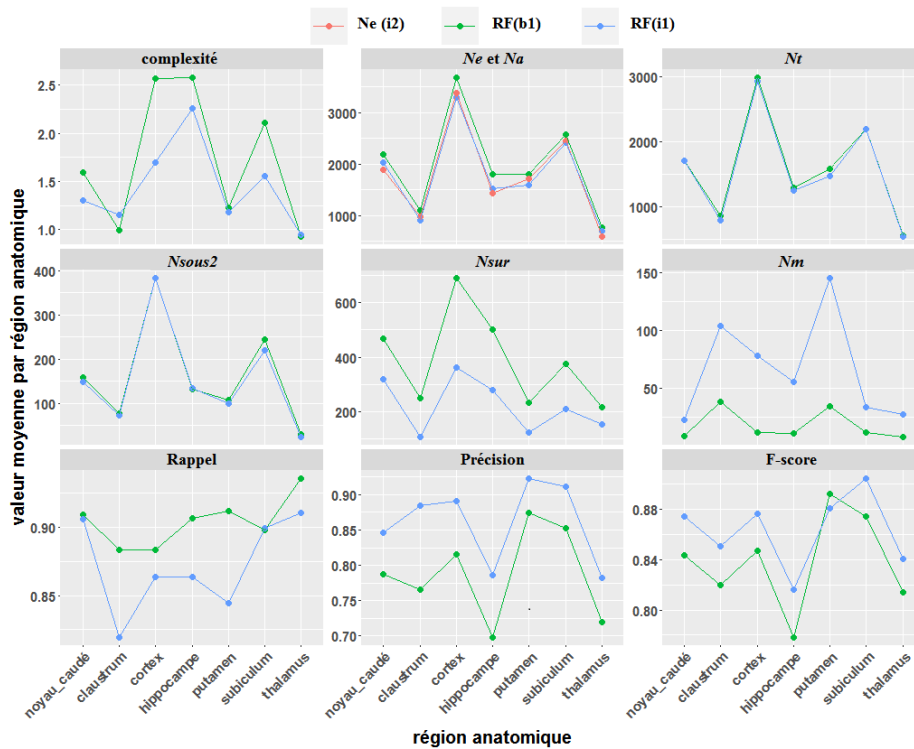


Figure 67 : Illustration des résultats obtenus en fonction de différents paramètres testés (complexité, N_e , N_a , N_t , $N_{s\text{ous}2}$, N_{sur} , N_m , R , P , F) et des différentes régions anatomiques.

Le critère de complexité calculé dans la Figure 67 est irrégulier, nous observons que globalement les régions du cortex, de l'hippocampe et du subiculum sont plus complexes que les autres du fait de la présence de composantes associant plusieurs neurones. Dans ces régions-là, il est donc nécessaire d'individualiser les neurones accolés dans les composantes connexes. A partir de nos résultats, nous pouvons conclure que, sur l'ensemble d'une coupe de cerveau, il n'y a pas de région où tous les neurones sont déjà individualisés. De ce fait, il sera systématiquement nécessaire de mettre en œuvre l'étape d'individualisation même dans des régions où il y a très peu de neurones qui se touchent.

Concernant le comptage manuel des neurones correspondant au paramètre N_e cela représente une grande charge de travail (50 images de taille 5000x5000 pixels, avec un nombre de neurones compris entre 500 et 4000 neurones par image) qui a été réalisée par le traiteur d'image i2. Dans les images de mêmes tailles, le nombre moyen de neurones est décroissant dans les régions suivantes, respectivement : le cortex, le subiculum, le noyau caudé, le putamen, l'hippocampe, le claustrum puis le thalamus. Cependant, ce résultat est à interpréter avec précaution car c'est une valeur moyenne par région et il existe des hétérogénéités aussi bien dans les régions que dans les images extraites. Nous allons introduire le paramètre de densité de neurones pour estimer la densité de neurones dans les régions anatomiques dans la partie 2.5.6 pour les données plus anatomiques. De plus, le nombre de neurones individualisés correspondant au paramètre N_o dans notre méthode est très proche du comptage manuel. Nous observons que le nombre de neurones calculé en fonction du modèle RF (b1) est plus grand que celui calculé en fonction de modèle RF (i1).

Pour les paramètres suivants (N_t , N_{sous2} , N_{sur} , N_m), les résultats obtenus suivent les mêmes tendances. Le nombre de neurones individualisés automatiquement de manière correcte (N_t) et la sous-segmentation (N_{sous2}) par rapport à la segmentation de l'opérateur i2 sont assez reproductibles. Par contre, la sur-segmentation N_{sur} produit des résultats variables entre les deux modèles testés. L'opérateur b1 segmente beaucoup plus largement les pixels de la classe neurone par rapport à l'opérateur i1 ce qui augmente le risque de sur-individualisation sur les modèles générés à partir de ces segmentations. Pour le critère N_m , ceci explique pourquoi RF (b1) individualise plus de neurones (N_o) par rapport à RF (i1). Nous avons enfin évalué les différents modèles grâce à trois caractéristiques – R , P et F . Le modèle RF (b1) par rapport à RF (i1) permet de détecter plus de neurones par rapport au nombre total de neurones qui existent réellement selon R , cependant entre les neurones détectés, il existe plus de mauvaises détections de neurones générées par RF (b1) selon P . En ce qui concerne le F -score, le modèle RF (i1) donne les meilleurs résultats d'individualisation, car il produit moins de sur-segmentation par rapport au modèle RF (b1).

La Figure 68 montre le nombre moyen de neurones détectés (N_e) par l'opérateur i2, des neurones détectés automatiquement (N_o) ou non (N_m) par la méthode d'individualisation ainsi que la variabilité intra-région. Cette figure présente les pourcentages de l'écart-type par rapport à la moyenne du nombre de neurones détectés par région. Plus le chiffre est faible, plus la région est homogène.

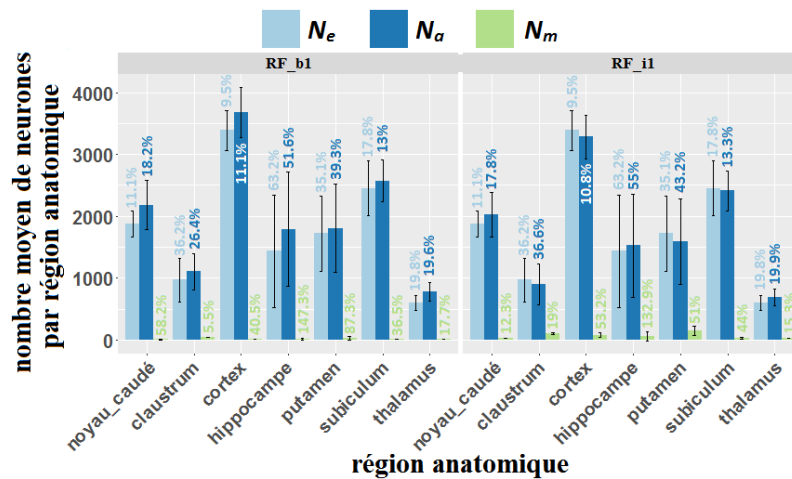


Figure 68 : Nombre moyen des neurones détectés par région anatomique. N_e : nombre de neurones segmentés par l'opérateur i2. N_a : nombre de neurones individualisés par notre méthode. N_m : nombre de neurones non détectés mais marqués par l'opérateur i2.

Nous observons que l'hippocampe est la région la plus hétérogène avec des écart-types supérieurs à 50%. Ceci s'explique du fait du mélange de sous-structures très différentes en termes d'organisation neuronale telles que les sous-régions CA1-4 et le gyrus denté. Ensuite, viennent le claustrum et le putamen. Ce résultat devra être confirmé car ce sont des régions étroites et des images de 5000x5000 pixels ne contiennent en général qu'une petite partie de ces deux régions, ce qui conduit inévitablement à une contamination de la mesure par d'autres régions. Cependant, nous observons que les variabilités sont comparables entre les différentes méthodes.

Pour visualiser de façon objective et précise la variabilité des performances de notre méthode d'individualisation des neurones en fonction de toutes les régions anatomiques, nous présentons les résultats concernant le F -score pour les 50 images dans la Figure 69. Nous observons que la méthode est capable d'individualiser les neurones avec de très bons scores dans le noyau caudé, le cortex et le subiculum (F -scores supérieurs à 0,8 pour RF (b1) et à 0,85 pour RF (i1)). Cependant, il y a des régions où il existe des problèmes, par exemple, dans l'hippocampe, surtout dans les sous-régions CA2 et CA3. Quand nous traitons ces régions, la méthode ne donne pas de bons résultats (F -score). Ceci s'explique du fait que c'est une région très grande et très hétérogène qui présente des propriétés particulières. En effet, les neurones marqués comme nous l'avons vu dans la Figure 66a-b n'ont pas les propriétés correspondant à nos hypothèses (noyau plus sombre, forme circulaire, gradient d'intensité). Cela constitue une des limites de notre méthode. Il faut donc faire attention à l'interprétation des résultats produits au niveau de l'hippocampe. Toutefois, notre méthode permet d'obtenir de bons résultats dans d'autres régions d'intérêt comme le cortex. A ce stade, il est très intéressant de pouvoir identifier les régions anatomiques pour lesquelles notre méthode d'individualisation pourrait être utilisée avec un bon niveau de confiance, cela constitue une des principales contributions de notre travail. En particulier, la possibilité de traiter le cortex de façon satisfaisante est de toute première importance car c'est une région particulièrement étudiée en neurosciences et cela ouvre la possibilité de mener des études biologiques d'un grand intérêt (étude de la distribution des neurones dans les différentes aires corticales, étude des couches cellulaires spécifiques, meilleure caractérisation de l'organisation des neurones, etc.). Pour les régions présentant des résultats moins bons (quelques sous-régions de l'hippocampe, thalamus, claustrum), il est possible d'envisager deux voies d'amélioration soit en jouant sur la production des données (meilleur marqueur IHC, numérisation, etc.) soit en adaptant le protocole de traitement. Au vu des images déjà présentées (neurones en arc, cellules claires en leur centre, etc.), la première option

semble la plus prometteuse. Un dernier point à souligner est la lourdeur de ce type de validation. Si nos premiers résultats donnent des informations biologiques très riches, l'application de notre méthode à des coupes entières sur la totalité du cerveau nécessitera de poursuivre ces étapes de validation de notre méthode car il est très probable que de nouvelles conformations atypiques des neurones seront trouvées. Ceci constituera également une contribution très utile des futurs travaux de recherche envisagés.

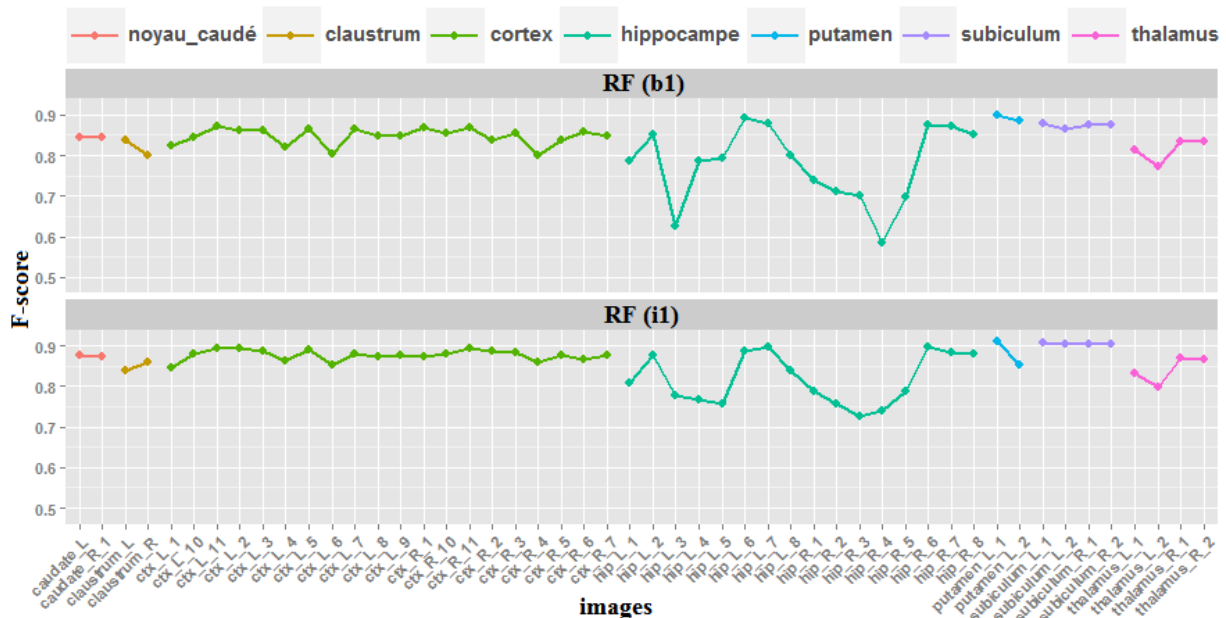


Figure 69 : *F-score* calculés en fonction des deux modèles générés par RF (b1) et RF (i1) pour 50 images issues de la base d'images pour l'individualisation.

- Etude de la distribution spatiale des neurones correctement, sous- ou sur-détectés

La Figure 70 permet de voir des exemples de la distribution des neurones en fonction de leur taille et de la qualité d'individualisation (correctement, sous- ou sur-détectés). L'idée ici n'est pas de visualiser tous les neurones mais d'avoir une vue globale nous permettant d'apprécier la proportion des neurones correctement marqués. Les neurones individualisés sont présentés sur leur centroïde par un point de couleur et de taille égale à la valeur de leur rayon moyen. La couleur rouge représente une individualisation correcte, vert représente une sur-individualisation et bleu représente une sous-individualisation. Nous observons que les résultats d'individualisation sont spatialement assez bien répartis. Nous n'avons pas identifié, dans l'ensemble des zones de travail traitées, de concentrations anormales d'erreurs (sous- ou sur-individualisation) ce qui constitue un point intéressant de notre méthode. Les résultats sur toutes les images de la base d'individualisation sont présentés dans l'annexe 4 – « Distribution spatiale des neurones selon leur taille et la qualité des résultats de l'individualisation par rapport à la détection de référence ».

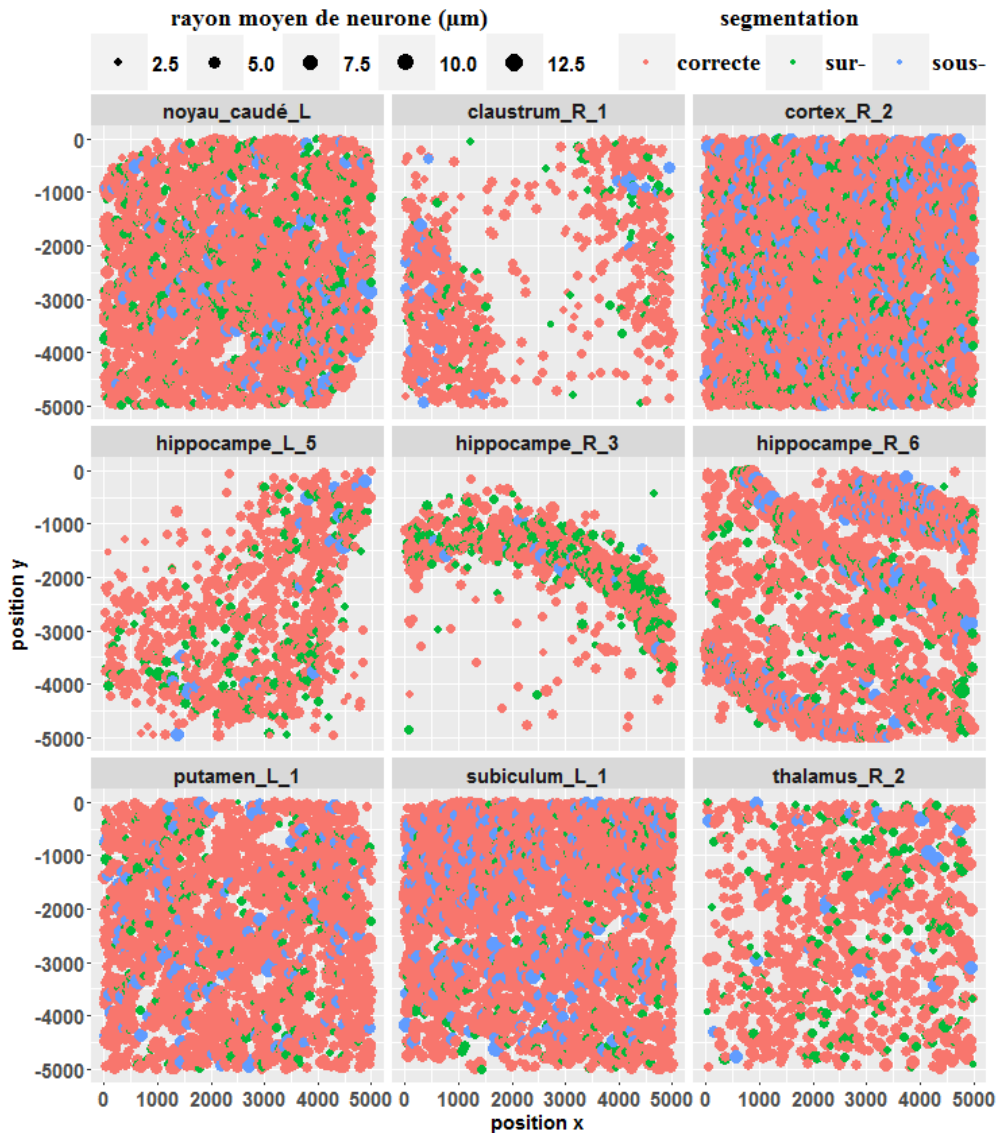


Figure 70 : Distribution spatiale des neurones en fonction de leur taille et des résultats d'individualisation. La couleur rouge représente une individualisation correcte, vert représente une sur-individualisation et bleu représente une sous-individualisation.

- Etude des rayons moyens des neurones dans les régions anatomiques

La Figure 71 illustre de manière compacte la distribution du rayon moyen des neurones individualisés dans les différentes régions du cerveau de macaque. Six statistiques récapitulatives (le minimum, 1^{er} quartile, la moyenne, la médiane, 3^{ème} quartile et le maximum) sont listées dans le Tableau 13.

Nous observons que la plupart des neurones, qui restent dans l'intervalle des 1^{er} et 3^{ème} quartiles, dans le claustrum (rayon moyen variant de 2,99 – 4,52 µm) et le putamen (rayon moyen variant de 3,31 – 4,73 µm) sont plus petits que les neurones des autres régions (rayon moyen variant de 3,34 – 6,02 µm). Le rayon moyen des neurones individualisés dans nos études varie de 1 à 14,3 µm ce qui correspond assez bien à l'étude de Andersen et al. (2016) qui ont estimé que la taille des neurones varie entre 5 et 30 µm de diamètre (5 valeurs aberrantes pour des neurones ont été exclues). La détection de très petits neurones peut être due à la production de neurones partiels lors de la coupe ou à un problème lors de la numérisation (plan focal de numérisation non centré sur le neurone, position des neurones dans la profondeur de la coupe, cellules abimées lors du marquage, etc.). Il est

important de garder en tête que les mesures morphologiques réalisées en 2D restent une estimation et qu'il serait nécessaire d'acquérir des volumes pour obtenir des mesures plus précises. Dans ce contexte, notre approche fournit de nouvelles informations qui devraient permettre d'enrichir les connaissances actuelles en s'appuyant sur les données réellement accessibles aujourd'hui dans nos études. La méthode d'individualisation donne des informations descriptives très intéressantes sur les populations de neurones pour chaque région. Sur les régions précédemment validées, cette information pourrait permettre de mieux décrire le développement, le vieillissement, la maladie voire même évaluer un traitement thérapeutique en produisant des données d'analyse à l'échelle cellulaire. Un second point fort de notre approche sera de pouvoir prochainement déployer ces technologies à très grande échelle (coupe entière, cerveau, groupes) ce qui devrait ouvrir des perspectives intéressantes dans le cadre des études biologiques menées à MIRCen.

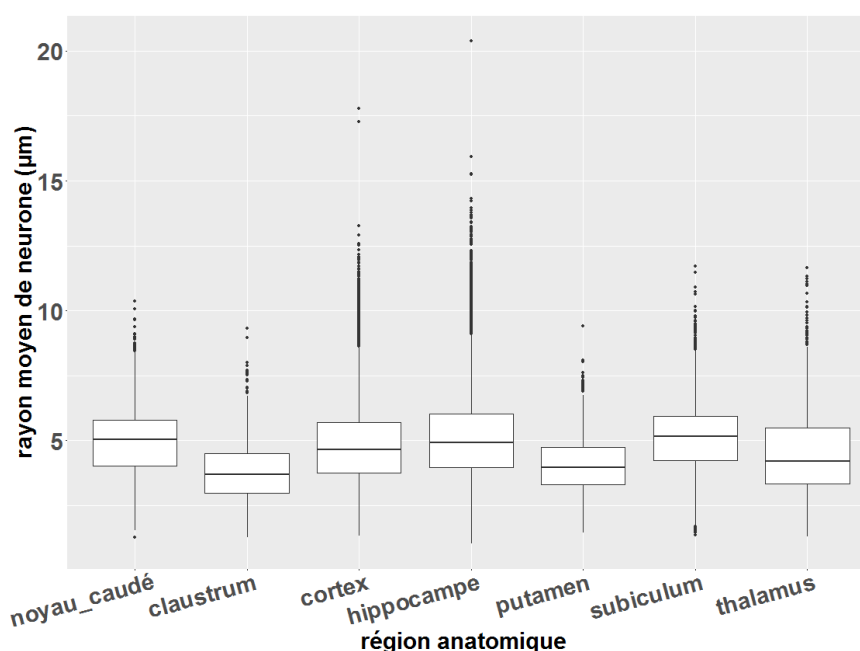


Figure 71 : Rayon moyen des neurones individualisés dans les régions anatomiques du cerveau.

Tableau 13 : Six statistiques récapitulatives concernant le rayon moyen des neurones par région anatomique. L'unité utilisée est le micromètre.

| Région | Min | 1st Qu. | Median | Mean | 3rd Qu. | Max |
|-------------|------|---------|--------|------|---------|-------|
| noyau caudé | 1,26 | 4,02 | 5,05 | 4,91 | 5,79 | 10,36 |
| claustrum | 1,28 | 2,99 | 3,69 | 3,80 | 4,52 | 9,3 |
| cortex | 1,34 | 3,74 | 4,64 | 4,77 | 5,69 | 17,80 |
| hippocampe | 1,02 | 3,96 | 4,91 | 5,08 | 6,02 | 20,38 |
| putamen | 1,44 | 3,31 | 3,98 | 4,03 | 4,73 | 9,41 |
| subiculum | 1,37 | 4,24 | 5,16 | 5,08 | 5,94 | 11,71 |
| thalamus | 1,29 | 3,34 | 4,21 | 4,49 | 5,48 | 11,64 |

- Comparaison de méthodes d'individualisation de cellules

La validation de notre méthode d'individualisation réalisée par comparaison avec le comptage d'un expert permet d'évaluer les possibilités d'application de cette approche en biologie. Il est également très intéressant de pouvoir comparer différentes méthodes ou versions de méthodes entre elles. Comme nous l'avons vu, d'autres méthodes ont été proposées dans la littérature. Nous avons en particulier identifié deux méthodes qui pourraient être testées : la méthode *iCut* (He et al., 2015) et

le *Watershed* (Cousty et al., 2009). Bien que nous ayons effectué des tests sur la méthode *iCut*, il n'a pas été possible d'intégrer ces résultats dans cette section car la méthode ne pouvait pas traiter nos images du fait de la quantité de mémoire vive nécessaire ($O(n^3)$ pour *iCut*, où n est le nombre de pixels dans l'image). Malgré des stations de travail équipées de 128 Go de mémoire vive, il n'a pas été possible d'obtenir de résultats. De plus, deux autres limitations ont été identifiées : la première est que la méthode *iCut* utilise des *a priori* sur la taille moyenne des rayons en faisant l'hypothèse de cellules de tailles similaires ce que nous avons démontré comme étant faux sur nos données et la seconde est que *iCut* se base fortement sur l'information de concavité qui ne peut être exploitée que pour des agrégations de cellules peu importantes mais serait rapidement mise en échec sur des agrégations massives que nous avons rencontrées dans le gyrus denté par exemple. Pour ce qui est de la méthode de *Watershed*, il existe un très grand nombre de versions de cette méthode. Nous avons choisi de tester une méthode particulière de *Watershed* (Cousty et al., 2009) implémentée et disponible dans notre laboratoire.

Notre étude a donc eu pour but de comparer notre méthode d'individualisation des neurones avec une autre méthode de référence – le *Watershed*. Comme ce dernier est sensible au bruit, une étape de prétraitement a été nécessaire, nous avons donc appliqué un filtrage gaussien. La taille des neurones étant variable dans le cerveau et variant entre 5 et 30 μm de diamètre, nous avons effectué deux tests avec la méthode de *Watershed* afin de comparer les deux méthodes de façon objective. Le premier test a été d'utiliser les images de σ adaptatifs calculés par notre méthode (Figure 50n) pour réduire le bruit dans les images. Nous l'appellerons « *Watershed* (σ optimal) ». Le deuxième test a été de fixer un paramètre unique de σ pour le filtrage gaussien afin de tester toutes les valeurs possibles. Nous l'appellerons « *Watershed* (23 valeurs de σ) ». Puis, l'individualisation des neurones a été réalisée par notre méthode et les deux tests concernant la méthode *Watershed* sur la base d'image créée pour valider l'individualisation.

La Figure 72 présente les résultats obtenus sur les mêmes images en utilisant notre méthode, la méthode « *Watershed* (σ optimal) » ainsi que la méthode « *Watershed* (23 valeurs de σ) » dont les σ égaux à 3 et 9 ont été choisis pour illustrer. Notre méthode et la méthode « *Watershed* (σ optimal) » fonctionnent globalement bien dans les cas simples (neurones déjà individualisés ou peu de neurones accolés). Cependant, dans les cas de figure plus compliqués (nombre de neurones accolés croissant), les résultats produits par la méthode de « *Watershed* (σ optimal) » se dégradent. Ils présentent souvent des sur- et sous-individualisations qui conduisent, dans un cas comme dans l'autre, à des contours finaux non naturels (bordures droites, grands neurones détectés). Pour la méthode « *Watershed* (23 valeurs de σ) », nous notons plus de sur-segmentations pour les petites valeurs de σ et plus de sous-segmentations pour les grandes valeurs de σ . Visuellement, notre méthode d'individualisation offre dans tous les cas un meilleur résultat.

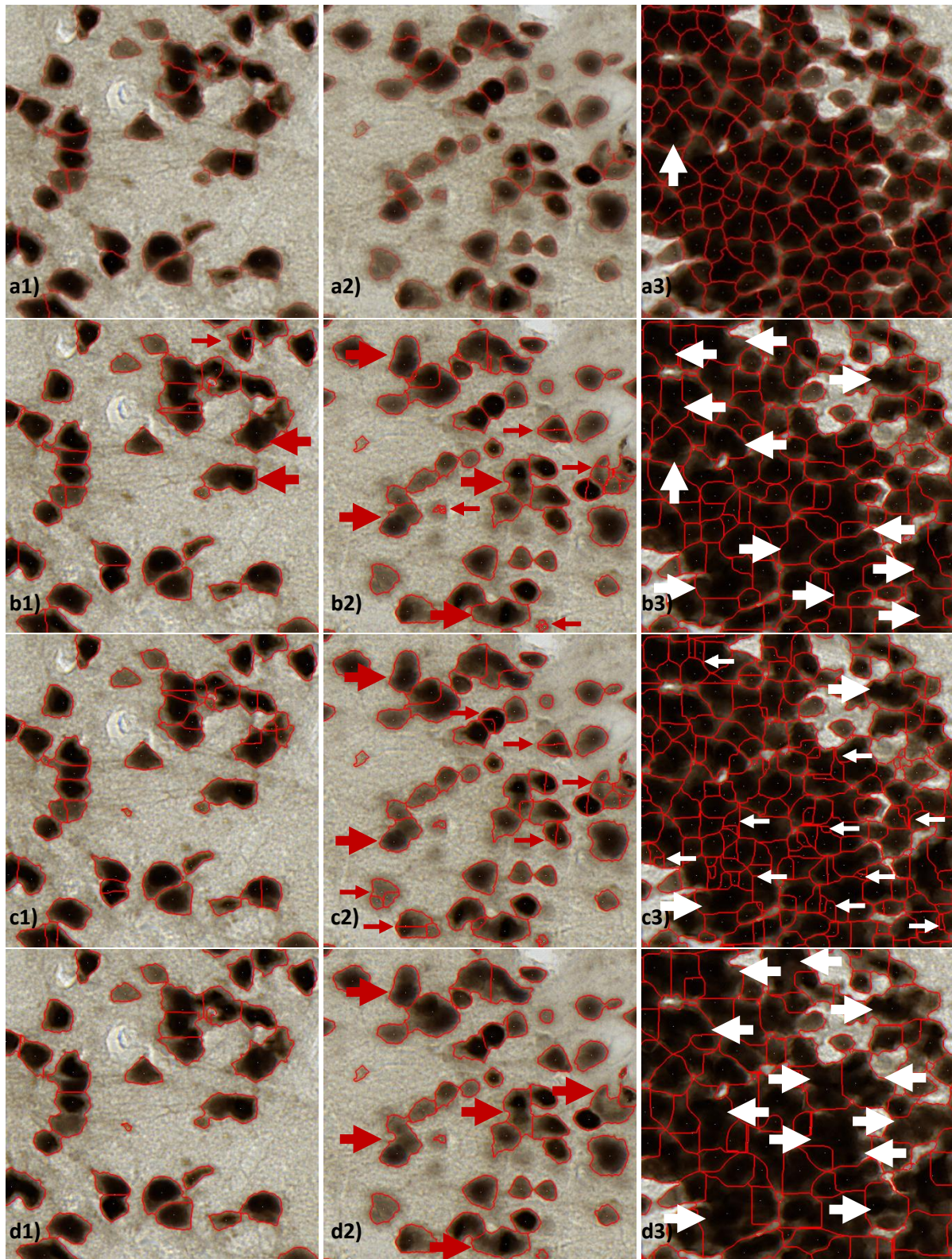


Figure 72 : Résultats d'individualisation obtenus par la méthode proposée et par deux tests concernant la méthode *Watershed* de référence. a) La méthode proposée, b) « *Watershed* (σ optimal) ». c) « *Watershed* « 23 valeurs de σ » dont le σ est fixé à 3. d) « *Watershed* « 23 valeurs de σ » dont le σ est fixé à 9. Les images 1, 2, 3 représentent différentes régions – subiculum, cortex, gyrus denté correspondant à différents niveaux de fraction de l'occupation de la surface des neurones. Les flèches fines (resp. épaisses) représentent les sur- (resp. sous-) individualisations.

En outre, une évaluation quantitative des trois méthodes a été réalisée sur la base de données pour l'individualisation et est illustrée dans la Figure 73. Nous observons que la méthode « *Watershed* (σ optimal) » détecte plus de neurones (N_a) que notre méthode dans toutes les régions. Cependant, si nous nous intéressons au nombre de neurones individualisés correctement (N_t), c'est systématiquement notre méthode qui est la plus performante. En outre, la méthode « *Watershed* (σ optimal) » présente généralement bien plus de sur- (N_{sur}) et sous-individualisations (N_{sous2}) par rapport à notre méthode. Pour N_{sur} , l'explication a deux origines. Une première cause est que la méthode RF a détecté du marquage NeuN dans des zones où l'opérateur 2 n'a pas segmenté de neurones correspondants. L'impact de ce problème est le même pour toutes les méthodes d'individualisation. La seconde cause est directement due aux méthodes d'individualisation testées. Nous savons qu'il existe de petits neurones partiels à cause du découpage de certains neurones dans l'image originale en 2D, le σ optimal calculé par notre méthode est alors petit dans l'image de σ optimal (Figure 50n). Comme la méthode *Watershed* est sensible au bruit, par conséquent, notre méthode est plus robuste pour ces petites valeurs de σ pendant le processus de la détection de petits neurones par rapport à la méthode *Watershed*. Cela conduit à une augmentation forte du nombre de neurones détectés par la méthode « *Watershed* (σ optimal) ». Pour N_{sous2} , notre méthode produit moins ce problème par rapport à la méthode « *Watershed* (σ optimal) » car notre méthode utilise deux itérations du processus de lissage gaussien suivi d'un filtrage min-max qui renforce l'information relative aux centroïdes. En considérant la méthode « *Watershed* (23 valeurs de σ) », les trois paramètres concernant N_a , N_t et N_{sur} diminuent au cours de l'augmentation de la valeur de σ tandis que N_{sous2} augmente au cours de diminution de la valeur de σ . Pour ces quatre paramètres, le changement de valeurs de N_a et N_{sur} est beaucoup plus important car les petites valeurs de σ (surtout quand σ est égal à 1 ou 2) vont conduire à un nombre très important de sur-segmentations. Le *F-score* est un critère qui évalue la performance de l'individualisation des neurones. Nous avons calculé ce critère en fonction des trois méthodes. Il montre que la méthode que nous avons développée offre toujours les meilleures performances. Pour la méthode « *Watershed* (23 valeurs de σ) », nous observons bien qu'une certaine gamme des valeurs de σ peut atteindre quasiment la même performance que celle obtenue par notre méthode. Pour les différentes régions anatomiques, les gammes des valeurs de σ qui fournissent les bonnes performances sont respectivement : 7-10 ($0,85 < F\text{-score} < 0,860$) pour le noyau caudé, 4-15 ($0,8 < F\text{-score} < 0,823$) pour le claustrum, 5-7 ($0,83 < F\text{-score} < 0,841$) pour le cortex, 6-8 ($0,73 < F\text{-score} < 0,743$) pour l'hippocampe, 4-9 ($0,85 < F\text{-score} < 0,864$) pour le putamen, 4-10 ($0,85 < F\text{-score} < 0,879$) pour le subiculum, 8-15 ($0,83 < F\text{-score} < 0,838$) pour le thalamus. Cela démontre la possibilité d'utiliser la méthode de *Watershed* avec une valeur fixe de σ , cependant, il faut adapter ce paramètre localement pour chaque région voire sous-région anatomique ce qui demande des informations *a priori* qu'il serait extrêmement difficiles d'obtenir à grande échelle. Pour la méthode « *Watershed* (σ optimal) », ces informations *a priori* ne sont pas demandées, mais les *F-score* calculés sont moins bons que ceux produits par notre méthode dans toutes les régions anatomiques. De plus, l'écart type de *F-score* par région anatomique avec notre méthode est systématiquement plus petit (les chiffres sont listés dans le Tableau 14), démontrant qu'elle est plus robuste dans tous les cas, simples et complexes. Il est important de souligner que le fait d'utiliser les images de σ adaptatifs a probablement réduit l'écart entre les deux méthodes adaptatives et que l'utilisation d'un σ fixe pour le *Watershed* sur toutes les régions anatomiques augmente le plus souvent cet écart pour certaines gammes de valeurs de σ .

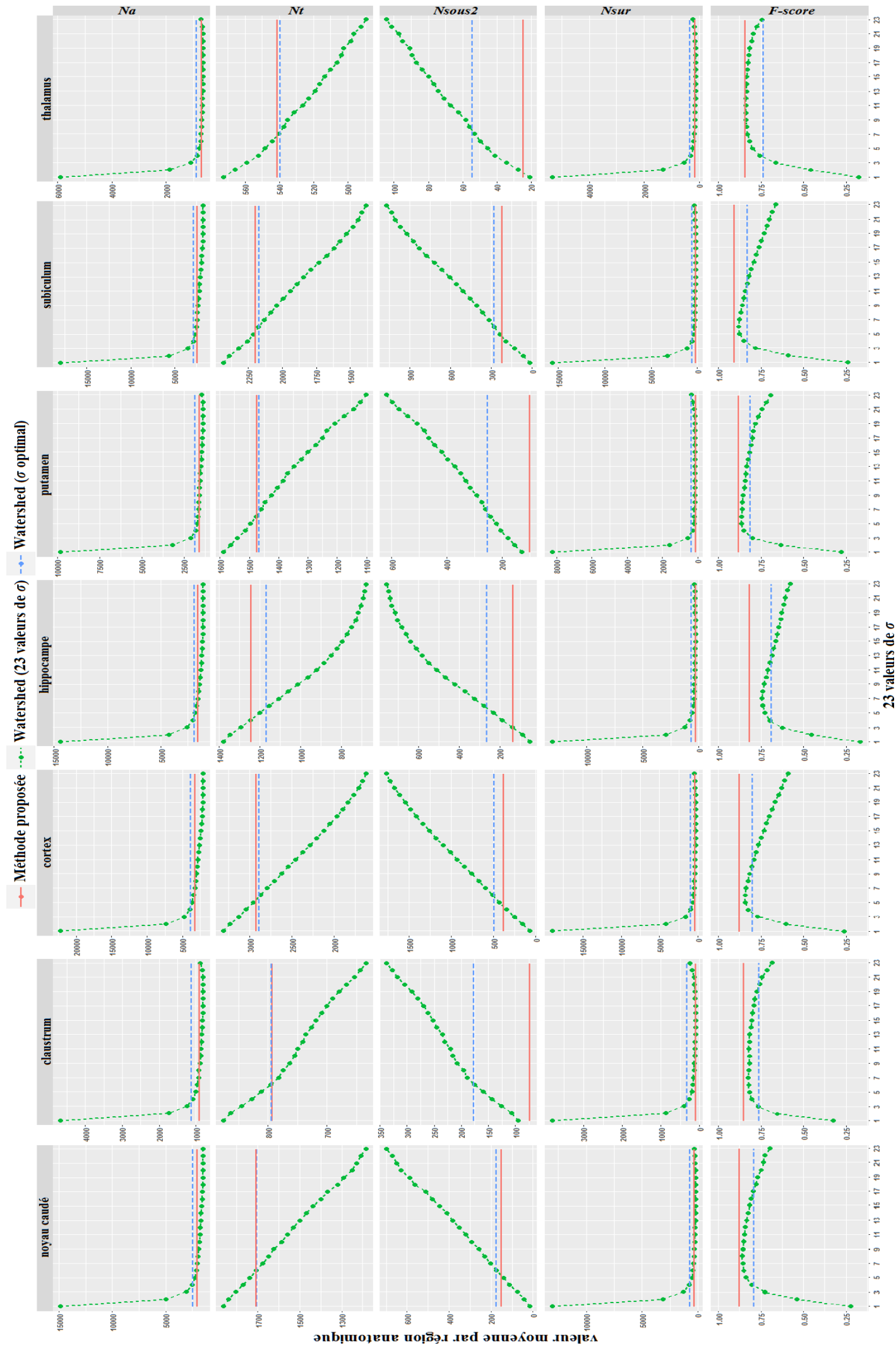


Figure 73: Comparaison des paramètres entre trois méthodes d'individualisation des neurones.

Tableau 14 : *F-score* calculés par notre méthode, par la méthode « *Watershed* (σ optimal) » et par la méthode « *Watershed* (23 valeurs de σ) » par région anatomique.

| Régions anatomiques | Méthodes d'individualisation des neurones | | | | |
|---------------------|---|--------------------------------------|--|--------------------------|----------|
| | Méthode proposée | <i>Watershed</i> (σ optimal) | <i>Watershed</i> (23 valeurs de σ) | | |
| | | | Toutes les valeurs de σ | La meilleure performance | |
| | | | | <i>F-score</i> | σ |
| noyau caudé | 0,874±0,002 | 0,791±0,004 | (0,222-0,860)±(0,001-0,009) | 0,860±0,001 | 8 |
| claustrum | 0,851±0,015 | 0,762±0,028 | (0,326-0,823)±(0,008-0,041) | 0,823±0,010 | 7 |
| cortex | 0,877±0,014 | 0,798±0,020 | (0,261-0,841)±(0,018-0,035) | 0,841±0,018 | 6 |
| hippocampe | 0,816±0,062 | 0,687±0,080 | (0,169-0,743)±(0,040-0,117) | 0,743±0,040 | 7 |
| putamen | 0,881±0,042 | 0,813±0,039 | (0,277-0,864)±(0,006-0,043) | 0,864±0,042 | 6 |
| subiculum | 0,905±0,001 | 0,831±0,006 | (0,239-0,879)±(0,004-0,035) | 0,879±0,006 | 6 |
| thalamus | 0,841±0,034 | 0,736±0,039 | (0,176-0,838)±(0,016-0,043) | 0,838±0,028 | 10 |

Les résultats qualitatifs présentés dans la Figure 72 sont particulièrement intéressants car ils mettent en lumière que même si le nombre des neurones est correctement estimé par une méthode, il n'en demeure pas moins que la qualité de l'estimation tant de la localisation des centroïdes que celle des contours de la segmentation de ces neurones individuels devraient être étudiées plus finement et également évaluées. En partant de ce constat, nous avons proposé d'introduire ces critères supplémentaires d'évaluation des méthodes d'individualisation. Nos résultats sont présentés dans la section suivante.

2.5.5. Validation des méthodes d'individualisation basée sur des caractéristiques des neurones

La base de données pour l'individualisation se compose de 58 images de 5000x5000 pixels. La segmentation manuelle des centroïdes et des contours des neurones pour toutes ces images serait une tâche qui demanderait des semaines de travail. Au lieu de cela, nous avons proposé une approche intermédiaire afin de produire des résultats préliminaires et d'évaluer la pertinence d'utiliser ces critères. Sur la base de 58 images, un sous-jeu représentatif de 9 images correspondant à 7 régions anatomiques a été choisi. Sur chaque image, 100 neurones ont été segmentés par l'expert i1 soit un centroïde et un contour pour chaque neurone (présenté dans la section 2.4.6.2). Ces données ont ensuite été utilisées pour comparer les deux méthodes automatiques – notre méthode et la méthode « *Watershed* (σ optimal) » (pour simplifier, nous le noterons *Watershed* dans les parties suivantes). Les 9 images choisies pour cette étude sont listées ci-dessous :

- CJ1301_slide91unsharp_caudate_L
- CJ1301_slide91unsharp_claustrum_L
- CJ1301_slide91unsharp_ctx_L_2
- CJ1301_slide91unsharp_ctx_L_7
- CJ1301_slide91unsharp_hip_L_6
- CJ1301_slide91unsharp_hip_L_7
- CJ1301_slide91unsharp_putamen_L_1
- CJ1301_slide91unsharp_subiculum_L_1
- CJ1301_slide91unsharp_thalamus_L_1

A partir de différentes segmentations de neurones (manuelle, *Watershed* ou méthode proposée), il est possible de calculer un certain nombre de paramètres d'intérêt pour comparer ces méthodes : distance moyenne entre les centroïdes, distance moyenne entre les contours, écart-type de la distance entre les contours, coefficient de Dice entre les surfaces des neurones segmentés. Tous les

résultats sont listés dans l'annexe 5 – « Valeurs moyennes de différentes caractéristiques des populations des neurones par région anatomique ». La Figure 74 montre les caractéristiques des populations des neurones. Nous observons que les neurones segmentés par notre méthode se superposent mieux par rapport à ceux segmentés par la méthode *Watershed*, avec les neurones segmentés manuellement (Figure 74 – coefficient de Dice). Les valeurs moyennes de coefficient de Dice sont de 0,77 pour notre méthode et de 0,72 pour la méthode *Watershed*. Les centroïdes détectés (Figure 74 – Distance_centroïde) et les contours (Figure 74 – Distance_moyenne_contour) segmentés par notre méthode, comparés au *Watershed*, sont plus proches de ceux segmentés par l'opérateur. Les valeurs moyennes de Distance_centroïde sont de 5,95 pixels pour notre méthode et de 6,61 pixels pour la méthode *Watershed*. Celles de la Distance_moyenne_contour sont de 4,46 pixels pour notre méthode et de 5,71 pixels pour la méthode *Watershed*. Nous trouvons moins de problèmes de sur- ou sous-individualisations produites par notre méthode. La même conclusion pourrait être obtenue en fonction du paramètre Ecart-type_distance_contour. Par conséquent, nous retrouvons le même résultat que précédemment, à savoir que la qualité de la segmentation des neurones par notre méthode est meilleure que celle obtenue par la méthode de *Watershed*. Nous avons vu précédemment qu'il était possible de comparer les méthodes d'individualisation en se basant sur le comptage des neurones. Les nouveaux critères de comparaison que nous avons proposés offrent des informations de comparaison supplémentaires des caractéristiques des neurones individualisés pour affiner cette comparaison. Ainsi, il serait tout à fait possible d'envisager le cas où deux méthodes détecteraient le même nombre de neurones mais où l'une des méthodes fournirait une individualisation plus rigoureuse en termes de localisation des centroïdes et des contours des neurones détectés. C'est ce que nous avons démontré avec cet exemple. Ce développement constitue une autre contribution de notre travail dans la mesure où il permet d'évaluer de manière quantitative des résultats de détection de neurones qui ne pouvait être auparavant réalisé que de façon visuelle, donc qualitative, par les experts (Figure 72).

Enfin, il faut noter que les logiciels développés à cette occasion ainsi que données produites pour cette validation sont tout à fait uniques et constituent une base de validation précieuse pour les développements méthodologiques à venir.

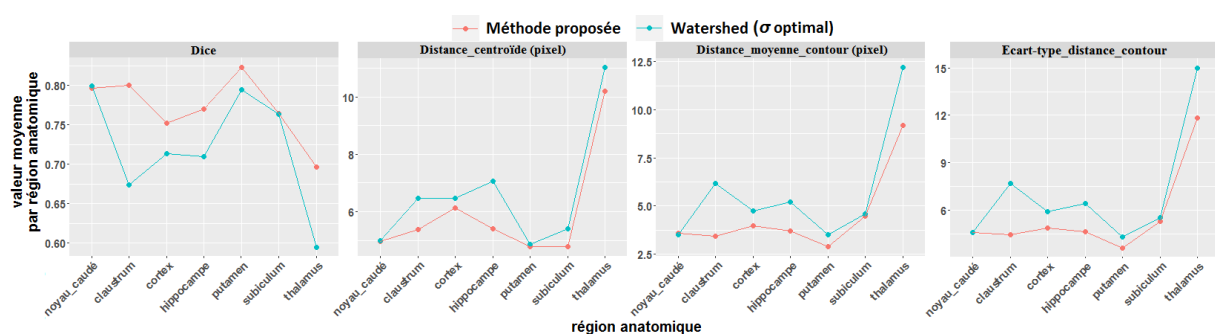


Figure 74 : Résultats obtenus pour comparer des méthodes d'individualisation des neurones – *Watershed* et notre méthode avec de nouveaux critères (coefficient de Dice, Distance_centroïde, Distance_moyenne_contour et Ecart-type_distance_contour).

Si les travaux d'évaluation de la qualité de l'individualisation des neurones ont été menés en se basant sur des techniques issues du traitement de l'image, il est également possible et fortement souhaitable d'évaluer ces méthodes avec des techniques de référence utilisées par les biologistes dans leurs études, à savoir la stéréologie.

2.5.6. Comptage stéréologique de référence

Nous avons demandé à deux biologistes (b2 et b3) de faire le comptage des neurones par stéréologie sur 8 RDIs de la coupe histologique N° 91 du cerveau de macaque au microscope. Comme la stéréologie est opérateur-dépendante, afin d'assurer la meilleure précision du comptage possible, chacune a effectué deux fois le comptage en modifiant la fraction d'échantillonnage. Le Tableau 15 liste la surface de chaque région anatomique ainsi que la fraction d'échantillonnage qui a été utilisée dans notre étude pour fournir des résultats exploitables. Ce dernier paramètre est ajusté par les biologistes pour chaque étude (taille et nombre de dissecteurs optiques).

Tableau 15 : Fractions d'échantillonnage définies par les biologistes experts pour les deux comptages stéréologiques (fraction mesurée par rapport à la surface totale des RDIs).

| Régions | S1_ cortex | S2_ noyau_ caudé | S3_ thalamus | S4_ subiculum | S5_ thalamus | S6_ gyrus_ denté | S7_ cortex | S8_ claustrum |
|-------------------------------|---------------|------------------------|-----------------|------------------|-----------------|------------------------|---------------|------------------|
| Surface (mm ²) | 5,89 | 1,56 | 4,24 | 1,24 | 1,33 | 0,13 | 2,33 | 2,31 |
| 1 ^{er} | 1/10 | 1/4 | 1/6 | 1/6 | 1/4 | 2/5 | 1/10 | 1/6 |
| 2 nd | 1/8 | 1/2 | 1/4 | 1/4 | 1/2 | 1/2 | 1/8 | 1/4 |

- Résultats de comptage produits par les deux experts biologistes

La Figure 75 montre la moyenne et l'écart-type du nombre de neurones estimés sur les 8 RDIs par les deux opérateurs. Les valeurs exactes sont listées dans le Tableau 16. Il montre que les comptages des deux opérateurs sont globalement proches (écart inférieur à 10% pour 7 régions sur 8) sauf pour la région du claustrum pour laquelle l'opérateur b3 a compté 18,4% de neurones en moins par rapport à l'opérateur b2. La figure présente les pourcentages de l'écart-type par rapport à la moyenne de chaque comptage. Ils représentent la variabilité du comptage de neurone pour chaque opérateur entre les deux mesures. Plus ils sont faibles, plus le comptage est fiable (stabilité de la mesure en fonction de l'échantillonnage). Nous observons que sur les mêmes régions anatomiques, le comptage diffère entre les deux opérateurs mais il reste relativement stable pour toutes les régions choisies pour cette étude. Pour l'opérateur b2, le comptage est généralement plus stable que pour l'opérateur b3 car l'écart-type sur les estimations des deux mesures est plus faible. Dans quelques régions comme par exemple le S5_thalamus, le comptage réalisé par les deux opérateurs varie de façon importante car c'est une région comprenant beaucoup de neurones clairs et les deux opérateurs n'ont pas effectué le comptage de la même manière. Cela met en lumière les variabilités pouvant se produire entre opérateurs du fait d'une perception différente de l'intensité des objets marqués. Par ailleurs, pour la région du gyrus denté, il est important de souligner que les biologistes ne réalisent quasiment jamais d'étude stéréologique dans cette région parce que c'est une région particulièrement difficile voire impossible à analyser car des milliers des neurones sont accolés. Malgré la difficulté de la tâche, les deux biologistes experts ont cependant essayé de compter les neurones dans cette région et les résultats semblent fiables (écart-type : 5,9% - 14,7%). Pour faciliter le travail, une partie moins dense et moins grande du gyrus denté a été ciblée.

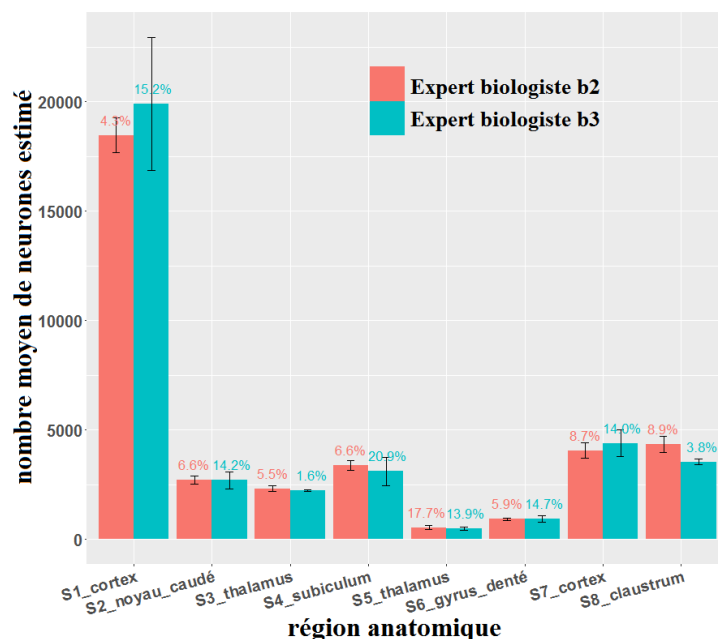


Figure 75 : Comparaison des comptages stéréologiques de neurones réalisés par deux experts.

Tableau 16 : Nombre total de neurones estimé deux fois par deux biologistes : moyenne et écart-type du nombre de neurones estimé.

| Région | Expert b2 | | | Expert b3 | | | Incrément de moyenne |
|----------------|--------------------------|---------------------------|---------------------|--------------------------|---------------------------|---------------------|----------------------|
| | 1 ^{er} comptage | 2 ^{ème} comptage | Moyenne± Ecart-type | 1 ^{er} comptage | 2 ^{ème} comptage | Moyenne± Ecart-type | |
| S1_cortex | 17912 | 19023 | 18468±786 | 22044 | 17761 | 19902±3029 | 7,8 % |
| S2_noyau_caudé | 2857 | 2602 | 2729±180 | 2981 | 2438 | 2710±384 | -0,7 % |
| S3_thalamus | 2226 | 2407 | 2316±128 | 2261 | 2212 | 2236±35 | -3,4 % |
| S4_subiculum | 3211 | 3524 | 3368±221 | 2650 | 3568 | 3109±649 | -7,7 % |
| S5_thalamus | 480 | 618 | 549±97 | 556 | 456 | 506±71 | -7,8 % |
| S6_gyrus_denté | 884 | 961 | 923±54 | 840 | 1035 | 938±138 | 1,6 % |
| S7_cortex | 4315 | 3817 | 4066±352 | 3965 | 4834 | 4400±615 | 8,2 % |
| S8_claustrum | 4619 | 4074 | 4346±385 | 3640 | 3450 | 3545±134 | -18,4 % |

- Comparaison des différentes méthodes de comptage

La Figure 76 illustre la comparaison du nombre estimé de neurones par notre approche algorithmique à partir des trois méthodes d'acquisition des images avec le scanner de lames (sur lesquelles nous avons reporté le plus précisément possible les RDIs utilisées pour la stéréologie) avec les estimations moyennes stéréologiques réalisées par les deux opérateurs. Nous observons tout d'abord que le nombre estimé de neurones par notre méthode est proche du comptage stéréologique. En ce qui concerne les différentes méthodes d'acquisition des images, les nombres de neurones calculés avec l'acquisition *unsharp* et *Z-stack* EDF sont très proches. Le nombre estimé de neurones avec l'acquisition en plan unique est inférieur dans les régions S6_gyrus_denté, S7_cortex, S8_claustrum et très inférieur dans la région S4_subiculum. Ceci s'explique du fait que la stratégie de mise au point lors de l'acquisition en plan unique peut conduire à une mauvaise focalisation sur certaines tuiles de la mosaïques et par conséquent entraîne une mauvaise segmentation des neurones à partir du tissu. Ceci a été mis en évidence par une analyse visuelle détaillée de l'ensemble des images traitées dans cette étude.

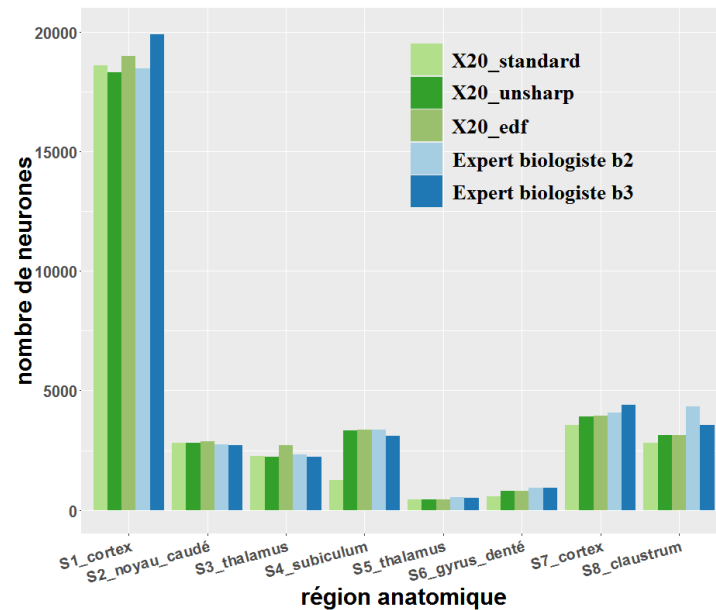


Figure 76 : Comparaison des comptages stéréologiques des neurones estimés par les deux biologistes avec les comptages produits par notre méthode avec les 3 types de numérisation.

Tableau 17 : Nombre de neurones estimé par les deux biologistes et produit par notre méthode avec les 3 types de numérisation.

| Région | X20_standard | X20_unsharp | X20_edf | Moyenne ± Ecart-type Expert |
|----------------|--------------|-------------|---------|-----------------------------|
| S1_cortex | 18589 | 18300 | 18981 | 19185±1987 |
| S2_noyau_caudé | 2805 | 2810 | 2890 | 2720±245 |
| S3_thalamus | 2247 | 2222 | 2721 | 2276±89 |
| S4_subiculum | 1250 | 3330 | 3365 | 3239±423 |
| S5_thalamus | 431 | 436 | 447 | 528±74 |
| S6_gyrus_denté | 586 | 810 | 809 | 930±86 |
| S7_cortex | 3567 | 3898 | 3930 | 4233±452 |
| S8_claustrum | 2814 | 3148 | 3137 | 3946±519 |

La Figure 77 démontre la corrélation entre le nombre estimé de neurones par notre méthode et celui estimé par stéréologie, noté $Corr_{nb}$. Ces grandes valeurs de corrélation démontrent une très forte corrélation linéaire entre les variables testées.

| | | |
|----------------------|---------------------|---------------------|
| Expert biologiste b2 | X20_standard: 0,998 | X20_standard: 0,991 |
| | X20_unsharp: 0,998 | X20_unsharp: 0,998 |
| | X20_edf: 0,998 | X20_edf: 0,997 |
| Expert biologiste b3 | | X20_standard: 0,995 |
| | | X20_unsharp: 0,999 |
| | | X20_edf: 0,999 |
| notre méthode | | |

Figure 77 : Corrélation entre le nombre de neurone calculé par notre méthode et le nombre de neurones estimés par stéréologie.

En considérant l'erreur relative dans l'estimation du nombre de neurones calculée automatiquement par rapport au comptage stéréologique, nous trouvons que les calculs automatiques basés sur l'acquisition *unsharp* et *Z-stack* EDF sont assez proches d'au moins une des estimations stéréologiques. Pour l'acquisition en plan unique, le calcul automatique est loin de l'estimation

stéréologique dans les régions du subiculum, du gyrus denté et du claustrum. Les résultats sont listés dans le Tableau 18.

Tableau 18 : Erreur relative du nombre de neurones calculée par notre méthode (3 types de numérisation) par rapport au comptage stéréologique (deux experts). Erreur < 10% en vert, 10% < erreur < 30% en orange, erreur > 30% en rouge.

| Région | X20_standard | | X20_unsharp | | X20_edf | |
|----------------|--------------|-----------|-------------|-----------|-----------|-----------|
| | Expert b2 | Expert b3 | Expert b2 | Expert b3 | Expert b2 | Expert b3 |
| S1_cortex | 0,007 | 0,066 | 0,009 | 0,081 | 0,028 | 0,046 |
| S2_noyau_caudé | 0,028 | 0,035 | 0,030 | 0,037 | 0,059 | 0,067 |
| S3_thalamus | 0,030 | 0,005 | 0,041 | 0,006 | 0,175 | 0,217 |
| S4_subiculum | 0,629 | 0,598 | 0,011 | 0,071 | 0,001 | 0,082 |
| S5_thalamus | 0,215 | 0,149 | 0,206 | 0,139 | 0,186 | 0,117 |
| S6_gyrus_denté | 0,365 | 0,375 | 0,122 | 0,136 | 0,123 | 0,137 |
| S7_cortex | 0,123 | 0,189 | 0,041 | 0,114 | 0,033 | 0,107 |
| S8_claustrum | 0,353 | 0,206 | 0,276 | 0,112 | 0,278 | 0,115 |

Les 8 images sur lesquelles nous avons travaillées appartiennent à 6 régions cérébrales d'une coupe de cerveau de macaque. Avec les résultats que nous avons obtenus, nous pouvons calculer la densité des neurones (nombre de neurones individualisés par millimètre carré, Figure 78 à droite) et la complexité (le nombre de neurones individualisés par rapport au nombre de composantes connexes, Figure 78 à gauche) pour chaque région. Nous observons que le gyrus denté est la région la plus complexe pour individualiser les neurones. Cela correspond bien à sa caractéristique - dans cette région, des milliers des neurones sont accolés. Ensuite, ce sont les régions du cortex et du subiculum qui sont les plus complexes. Du côté de la densité de neurones, c'est la région du gyrus denté qui est la plus dense. Cela correspond bien à la complexité de cette région. Ensuite, ce sont les régions du cortex puis du subiculum qui sont les plus denses.

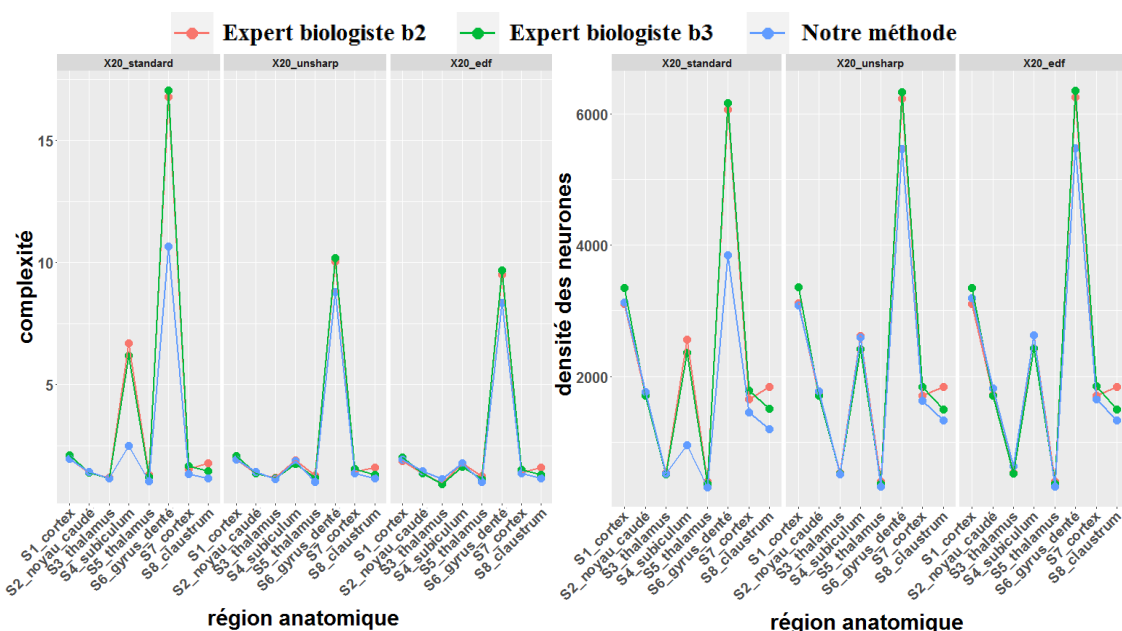


Figure 78 : Complexité correspondant à chaque RDI (à gauche) et densités de neurones estimées dans les 8 RDIs (à droite).

La Figure 79 illustre plus précisément les corrélations entre les densités de neurones estimées par les deux opérateurs et calculées automatiquement par notre méthode avec les 3 types de numérisation,

notée $Corr_{densité}$. Nous observons qu'elles sont très fortement corrélées (0,902-0,996). Cependant, pour les images numérisées avec un plan unique (standard), cette relation est plus faible à cause des biais importants dans les régions du S4_subiculum, du S6_gyrus_denté et du S8_claustrum (en raison de problèmes liés à la stratégie de mise au point par cette méthode d'acquisition), dans la région du subiculum et du gyrus denté.

| | | |
|----------------------------|---------------------|---------------------|
| Expert biologiste b2 | X20_standard: 0,995 | X20_standard: 0,902 |
| | X20_unsharp: 0,996 | X20_unsharp: 0,992 |
| | X20_edf: 0,996 | X20_edf: 0,990 |
| Expert biologiste b3 | | X20_standard: 0,927 |
| | | X20_unsharp: 0,994 |
| | | X20_edf: 0,993 |
| | | notre méthode |

Figure 79 : Corrélation entre les densités des neurones estimées par deux opérateurs et calculées automatiquement par notre méthode avec 3 types de numérisation des coupes histologiques.

Les densités de neurones estimées par les deux opérateurs et calculées par notre méthode ont été analysées sous la forme d'une régression linéaire simple (Figure 80). Le gris foncé représente l'intervalle de confiance autour de la courbe de tendance. Nous observons que l'intervalle de confiance pour l'acquisition en plan unique est plus large que pour les deux autres méthodes de numérisation. Pour les méthodes d'acquisition *unsharp* et *Z-stack EDF*, en considérant les hautes corrélations entre le nombre calculé de neurones et estimé dans la Figure 77 ($corr_{nb}=0,997-0,999$) et celles entre la densité de neurones calculée et estimée dans la Figure 79 ($corr_{densité}=0,99-0,994$), notre méthode donne des résultats très prometteurs pour estimer le nombre de neurones et de ce fait la densité de neurones dans chaque région. Il faut remarquer que la corrélation des densités des neurones est plus faible par rapport à celle du nombre de neurones. La raison en est que le paramètre de densité est normalisé par la surface de la région anatomique. Ainsi la valeur du nombre de neurones dans la région du S1_cortex est 10 fois plus grande que celle de la moyenne dans les autres régions, cela augmente donc fortement la corrélation.

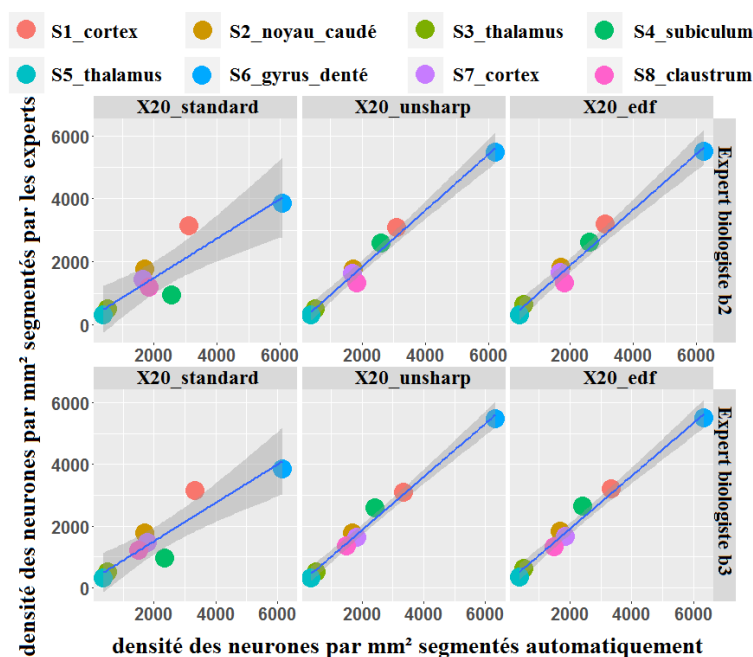


Figure 80 : Diagramme de dispersion de la densité de neurone calculée automatiquement par 3 méthodes de numérisation par rapport à l'estimation stéréologique de deux experts (ligne 1 : expert b2, ligne 2 : expert b3).

Les résultats de l'étude stéréologique sont préliminaires mais constituent, à notre connaissance, la seule étude de ce genre à avoir validé précisément une méthode automatique de comptage des neurones avec la méthode de référence utilisée en biologie. Une telle étude requiert une forte implication des experts et des études complémentaires intégrant des séries de coupes supplémentaires devront être menées pour étendre et confirmer ses résultats à plus grande échelle. Ces travaux seront nécessaires pour valider la possibilité de mettre en œuvre notre méthode de comptage des neurones dans des études biologiques menées très régulièrement à MIRCen.

2.6. Conclusion

Dans ce chapitre, nous avons présenté la méthode RF pour segmenter le marquage neuronal ainsi que notre méthode originale pour individualiser les neurones accolés. Plusieurs validations qualitatives et quantitatives ont été réalisées pour comparer les résultats d'individualisation obtenus par notre méthode, le *Watershed* et le comptage stéréologique. Pour cela, nous avons utilisé plusieurs critères de référence de la littérature ainsi que de nouveaux paramètres (centroïdes, contour des neurones). Concernant le comptage des neurones par RDIs, la performance de notre méthode montre ses limites dans un nombre limité de régions telles que CA2 et CA3 de l'hippocampe (*F-score* diminuant à 0,6). Ce sont des régions dans lesquelles la coloration des neurones est atypique (forme d'arc ou d'anneaux). Globalement, notre méthode fournit de bons résultats d'individualisation dans plusieurs RDIs importantes, en particulier le cortex (*F-score* variant de 0,85 à 0,9) qui est une région particulièrement intéressante en neurosciences. L'étude exploratoire menée dans notre travail donne une première liste des régions anatomiques dans lesquelles notre méthode de segmentation des neurones individuels pourrait être utilisée dans des études biologiques. En comparaison avec la méthode de référence de *Watershed*, le *F-score* de notre méthode est plus élevé. Grâce aux bases de validation que nous avons constituées dans ce travail, il nous sera possible de valider finement des évolutions de notre méthode ou des méthodes concurrentes. De plus, nous avons pu calculer le rayon moyen des neurones dans nos bases de travail et les résultats obtenus sont cohérents par rapport aux valeurs trouvées dans la littérature (diamètre compris entre 5 et 30 μm). En ce qui concerne les caractéristiques des populations de neurones (position des centroïdes et des contours des neurones individualisés), notre méthode fournit une meilleure qualité d'individualisation. Les centroïdes et les contours détectés automatiquement par notre méthode sont plus proches de ceux segmentés par l'expert ce qui fournit de nouveaux critères objectifs et proches de l'analyse visuelle des experts pour évaluer plus finement des méthodes de segmentation de neurones. De plus, nous avons démontré que le nombre de neurones et donc la densité de neurones calculée par notre méthode sont linéairement très fortement corrélés (0,99-0,999) avec ceux estimés par la stéréologie pour les images numérisées avec la méthode *unsharp* et *Z-stack* EDF. Ces travaux permettent également de justifier l'utilisation du mode numérisation *unsharp* qui est plus simple et plus rapide à mettre en œuvre par rapport au mode *Z-stack* EDF. Les résultats de notre travail de recherche permettent d'envisager l'utilisation de notre méthode pour individualiser automatiquement des neurones dans les études menées à MIRCen. Bien sûr, dans un premier temps, il sera nécessaire de poursuivre les comptages stéréologiques pour confirmer ces résultats. L'application de ces travaux à des coupes entières de cerveau de macaque voire à des cerveaux sera un défi majeur à relever qui nécessitera des développements et des ressources informatiques supplémentaires considérables (stockage, réseau, puissance de calcul) pour produire et analyser des quantités massives de données pouvant représenter plusieurs dizaines de téraoctets.

3. Cartographies mésoscopiques quantitatives relatives aux neurones

La première partie de ce manuscrit, en particulier la section 1.1.4, a introduit les différentes études s'intéressant aux neurones. La haute résolution spatiale des images histologiques acquises par microscopie optique permettent de produire des images à l'échelle cellulaire dans différentes régions du cerveau. Le chapitre 2 a présenté, quant à lui, notre méthode automatique permettant de segmenter et d'individualiser les neurones. Les résultats obtenus ouvrent des perspectives importantes pour utiliser ces méthodes de comptage dans des études biologiques. Après l'étape d'individualisation des neurones, il est possible d'accéder à des informations supplémentaires telles que la morphologie, la colorimétrie et la distribution des neurones dans des régions anatomiques voire dans des cerveaux entiers. Cette section décrit les méthodes de quantification et de synthèse de l'information relative aux neurones que nous proposons, sous la forme de cartes paramétriques à plus basse résolution, réalisables à l'échelle de régions d'intérêts.

3.1. Principe et intérêt de l'intégration à l'échelle mésoscopique

Les résultats de l'individualisation des neurones sont des images des régions anatomiques fournissant pour chaque neurone labélisé les coordonnées de son centroïde ainsi que son masque binaire. A partir de ces informations et de l'image couleur originale, il est possible d'accéder à des paramètres de morphologie, de colorimétrie voire d'organisation des neurones. L'idée que nous avons explorée dans cette partie de notre travail a été d'exploiter ces informations pour générer des cartographies quantitatives qui permettent de synthétiser localement les caractéristiques de groupes de neurones. L'intérêt de cette approche est de synthétiser l'information pour un nombre très important de neurones. Si la gestion de millions de neurones segmentés peut constituer une difficulté pour interpréter des résultats biologiques, cela peut clairement devenir un atout lorsque l'on utilise cette masse d'information pour créer des cartographies synthétiques par une réduction de l'espace (création de pixels de taille adaptée intégrant une population de neurones) et en ciblant des caractéristiques informatives (nombre, taille, forme, voisinage, etc.). Pour créer ces cartes paramétriques, un changement d'échelle est mis en œuvre afin de produire des cartographies présentant des résolutions spatiales à l'échelle mésoscopique : la résolution originale des coupes histologiques dans notre étude est de $0,22\ \mu\text{m}$, les cartographies créées ont une résolution spatiale adaptative pouvant aller de quelques dizaines à quelques centaines de micromètres dans notre cas. Cet intervalle d'observation est défini comme l'échelle mésoscopique qui est une échelle intermédiaire entre l'échelle microscopique, qui caractérise les atomes ou les molécules, et l'échelle macroscopique, qui caractérise les corps dans leur ensemble.

La Figure 81 présente notre méthode de génération des cartographies quantitatives à basse résolution en prenant pour exemple la région hippocampique.

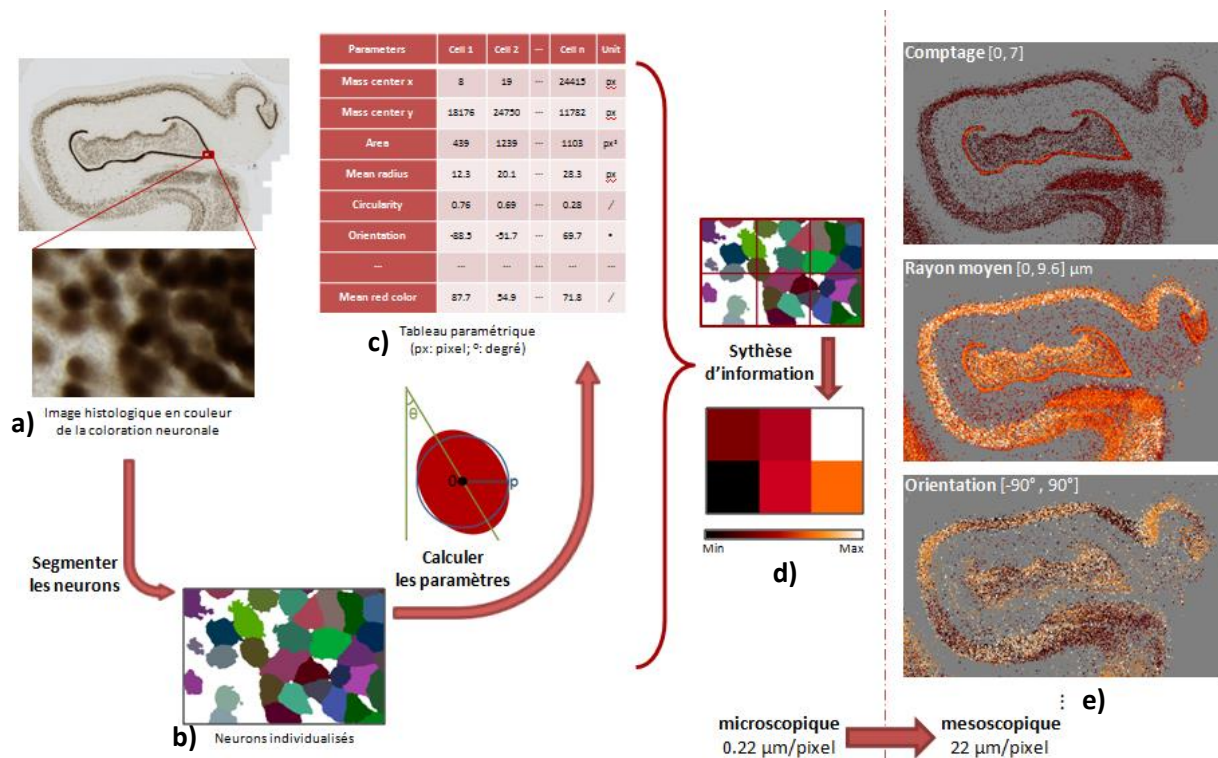


Figure 81 : Génération des cartographies paramétriques.

Dans l'hippocampe de macaque (Figure 81a), une grande quantité d'information est présente : image de 40000x23000 pixels correspondant à 2,6 Go. Pour interpréter ces informations, il faut pouvoir les quantifier. Nous réalisons donc une segmentation puis une individualisation des neurones (Figure 81b), résultant dans l'extraction et la labellisation de plusieurs dizaines de milliers de neurones. Puis pour chaque neurone, nous calculons un ensemble de paramètres d'intérêt décrivant par exemple leur morphologie et leur organisation comme leur densité, leur rayon moyen, leur orientation, etc. Nous obtenons alors pour chaque neurone des dizaines de paramètres sous la forme de tables (Figure 81c). Il devient alors possible de synthétiser ces paramètres sous la forme de cartes paramétriques de plus basse résolution. Par exemple, dans la partie de l'image en haute résolution présentée (Figure 81d), plusieurs neurones labélisés sont présents dans chaque bloc. Nous calculons alors la fraction d'occupation de la surface des neurones dans chaque bloc par rapport à la taille totale du bloc et l'attribuons comme valeur à chaque nouveau pixel d'une image en basse résolution (Figure 81d bas). Enfin, à titre d'exemple, nous présentons 3 cartes paramétriques quantitatives générées à partir de cet exemple : nombre de neurones, rayon moyen et orientation des neurones dans l'hippocampe (Figure 81e, agrégation de 100x100 pixels dans le plan).

Cette approche est intéressante car elle permet de synthétiser l'information pertinente à partir de la haute résolution vers une résolution plus faible. Les images microscopiques qualitatives sont ainsi transformées en images mésoscopiques quantitatives, plus informatives et plus simples à analyser.

Les étapes suivant l'individualisation des neurones (section 2) sont détaillées dans les sections suivantes : étape de calcul des paramètres d'intérêt (section 3.2) et étape de création des cartographies quantitatives mésoscopiques (section 3.3).

3.2. Calcul des paramètres descripteurs des neurones

Un cerveau humain possède environ 86 milliards de neurones. Ils sont de tailles et de formes très différentes et peuvent être caractérisés par un grand nombre de descripteurs qui peuvent être

colorimétriques mais aussi morphologiques. Nous avons étudié un ensemble de 49 descripteurs en utilisant l'image des neurones individualisés et labélisés (Figure 82a) ainsi qu'une image des zones de Voronoï associées (Figure 82b). L'image de partitions de Voronoï a été calculée pour estimer la zone d'influence spatiale des neurones. Dans les régions où les neurones sont clairsemés, la surface de leur partition de Voronoï correspondante est plus grande, et inversement.

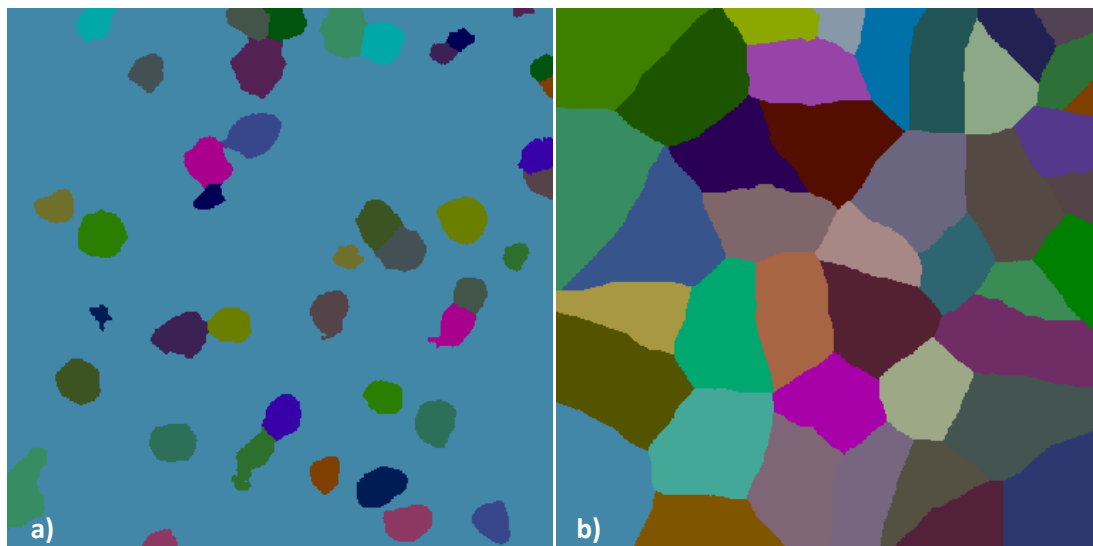


Figure 82 : a) 42 neurones individualisés et labélisés. b) Zones de Voronoï associées aux neurones labélisés.

Les descripteurs sont des caractéristiques telles que le rayon moyen, la couleur ou encore la circularité. Les 49 descripteurs utilisés dans notre recherche sont listés dans l'annexe 6 – « Descripteurs calculés pour la génération des cartographies quantitatives ». Nous pouvons regrouper ces descripteurs sous la forme de plusieurs familles de descripteurs :

- Morphologiques : surface, centre de gravité, périmètre, circularité, orientation des neurones,
- Colorimétriques : moyenne et écart-type des composantes de plusieurs espaces colorimétriques (RGB, HSV, etc.),
- De voisinage : nombre de neurones connexes et leurs caractéristiques.

Ces informations sont enregistrées sous la forme d'un « vecteur des neurones » de taille $n \times m$ (n étant le nombre de neurones segmentés et m étant le nombre de descripteurs calculés pour chaque neurone). A titre d'exemple, le Tableau 19 présente quelques descripteurs calculés sur plusieurs neurones correspondants à la Figure 82.

Tableau 19 : Descripteurs calculés sur plusieurs neurones de la Figure 82. « X_{cg} » et « Y_{cg} » sont respectivement la position X et Y du centre de gravité du neurone. « aire_px » est la surface du neurone en pixels. « p_n4_n8 » est le périmètre du neurone calculé basé en 8-connextité. « circularité » rend compte de la forme du neurone. Plus ce chiffre est proche de 1, plus la forme du neurone est circulaire. « rgb_moyen_r » est la moyenne de la composante rouge qui reflète le degré de coloration du neurone.

| label | X_{cg} | Y_{cg} | aire_px | p_n8 | rayon_moyen | circularité | orientation | ... | rgb_moyen_r |
|-------|----------|----------|---------|---------|-------------|-------------|-------------|-----|-------------|
| | pixel | pixel | pixels | pixels | pixel | / | degré | ... | / |
| 1 | 17 | 452 | 2052 | 187,233 | 28,812 | 0,736 | -63,435 | ... | 89,413 |
| 2 | 45 | 187 | 878 | 104,257 | 16,359 | 1,015 | -1,548 | ... | 120,803 |
| 3 | 67 | 351 | 1390 | 125,230 | 20,603 | 1,114 | 58,134 | ... | 84,110 |
| 4 | 82 | 14 | 854 | 107,034 | 16,283 | 0,937 | -48,945 | ... | 108,553 |
| 5 | 87 | 289 | 258 | 69,401 | 9,177 | 0,673 | 58,241 | ... | 137,597 |
| 6 | 88 | 212 | 1632 | 135,719 | 22,321 | 1,113 | -18,435 | ... | 88,278 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 42 | 500 | 164 | 562 | 95,931 | 13,445 | 0,767 | 42,797 | ... | 100,080 |

Ce tableau décrit les caractéristiques de quelques neurones, il nous aide à quantifier les informations intéressantes de chaque neurone et à synthétiser ces informations pour chaque nouveau pixel de la cartographie qui sera créée et pour chaque région anatomique qui sera étudiée. Le choix de ces paramètres s'appuie sur des paramètres référencés dans les études biologiques (nombre, circularité, etc.) mais également sur des paramètres standards utilisés en traitement de l'image. Il est possible de limiter la liste de ces paramètres en fonction des *a priori* ou encore d'enrichir cette base de nouveaux paramètres si besoin. Il est intéressant de souligner que cette étape permet de synthétiser l'information d'images de plusieurs gigaoctets sous la forme de tables de plus petites tailles. Dans l'exemple présenté (Figure 81a), l'image de l'hippocampe correspond à 2,6 Go et la table des paramètres représente seulement 15,8 Mo soit une réduction d'un facteur 165. Enfin, il est important d'avoir en tête le coût computationnel de ces calculs. 19 minutes ont été nécessaires pour calculer les paramètres dans le cas de l'hippocampe. Un code spécifique a été développé à cette occasion ainsi qu'un container dédié dans BrainVISA pour créer et accéder facilement à ce type de données.

3.3. Génération de cartographies quantitatives à l'échelle mésoscopique

La synthèse de l'information concernant les neurones sous la forme de cartographies quantitatives se base sur le « vecteur des neurones » précédemment calculé. Le premier choix à réaliser est de définir la taille des pixels de la cartographie. Ce choix n'est pas anodin car il impactera fortement l'analyse ultérieure :

- si le pixel est trop petit, le nombre de neurones sera très faible et le nombre de neurones présents à cheval sur plusieurs pixels sera plus important. Il est probable que les images produites seront également plus bruitées,
- si le pixel est trop grand alors il est probable qu'il risque d'intégrer un trop grand nombre de neurones voire des régions anatomiques hétérogènes voire différentes.

Le choix de la taille du pixel dépend de l'étude biologique, du marqueur histologique, de la taille de l'objet étudié (cellules, agrégats pathologiques, etc.) et du modèle animal. Dans cette étude, nous nous sommes basés sur une intervalle de résolutions spatiales utilisées dans nos recherches précédentes dans le domaine de la reconstruction *post mortem* 3D (anatomie : 10 μm et autoradiographie : 40 μm de résolution dans le plan, Dubois et al., 2010) ainsi que la résolution spatiale utilisée en imagerie anatomique *in vivo* en IRM sur des modèles animaux (30-300 μm , Petiet et al., 2011). Ces intervalles de valeurs sont donnés à titre indicatifs.

Pour produire les cartographies quantitatives, nous avons étudié deux modes de calcul basés sur le centre de gravité et la contribution de chaque neurone dans les pixels de cette cartographie :

- « Mode centre de gravité » : La valeur d'un paramètre calculé est attribuée au centre de gravité de chaque neurone. Lors du calcul de l'intensité du pixel dans la cartographie, les valeurs de tous les neurones sont comptabilisées si la coordonnée du centre de gravité est localisée dans le pixel. Cette méthode ne tient pas compte des effets de bord quand un neurone se trouve simultanément sur plusieurs pixels dans les cartographies.
- « Mode contribution » : La fraction de recouvrement des cellules sur plusieurs pixels de la cartographie est considérée pour l'application d'une pondération (poids de la contribution du neurone pour un pixel donné). L'importance de ces contributions est directement liée à la taille du pixel et à la densité de neurones de la zone étudiée.

La création d'une cartographie peut être vue comme l'agrégation isotrope en X et Y pixels à partir de la résolution microscopique originale. Lorsque nous générons des cartographies avec un facteur d'agrégation de $p \times q$ (agrégation de p pixels en X et q pixels en Y), chaque pixel de la cartographie synthétise la valeur d'un descripteur de la population de neurones contenue dans ce pixel. Cette synthèse peut être obtenue selon différents modes de calcul : moyenne, médiane, minimum, maximum, majoritaire ou somme des valeurs du descripteur étudié. Toutes ces possibilités ont été implémentées dans nos algorithmes pour répondre à la plupart des questions posées. Par exemple, dans la Figure 83, une image originale de taille 512x512 pixels contenant 42 neurones labélisés a été traitée pour générer une image de résolution plus faible de taille 2x2 pixels (chaque pixel de la cartographie étant l'agrégation de 256 x 256 pixels de l'image originale). Le nombre de neurones (Figure 83b) dans chaque pixel de la cartographie basse résolution sont respectivement de 9, 15, 8 et 10 en se déplaçant de gauche à droite et de haut en bas. Les figures Figure 83c et Figure 83d synthétisent l'information du rayon moyen des neurones en utilisant respectivement le « mode centre de gravité » et le « mode contribution » dans chaque pixel de l'image basse résolution. Nous observons que les cartographies générées par ces deux modes sont très similaires (cf. valeurs calculées, légende de la Figure 83). Dans les études que nous avons menées, les différences de résultats entre ces deux modes de calcul ont toujours été négligeables (inférieures à 5%). Il est intéressant de noter que le temps de calcul pour le « mode centre de gravité » est beaucoup plus faible du fait d'un coût calculatoire bien moindre (dans l'exemple présenté, Figure 81, lorsque l'agrégation est de 100x100 pixels, le temps d'exécution pour générer la cartographie de rayon moyen est de 1s en mode « centre de gravité » et de 6 minutes en mode « contribution »). De ce fait, le mode « centre de gravité » a été choisi par défaut dans la plupart de nos essais.

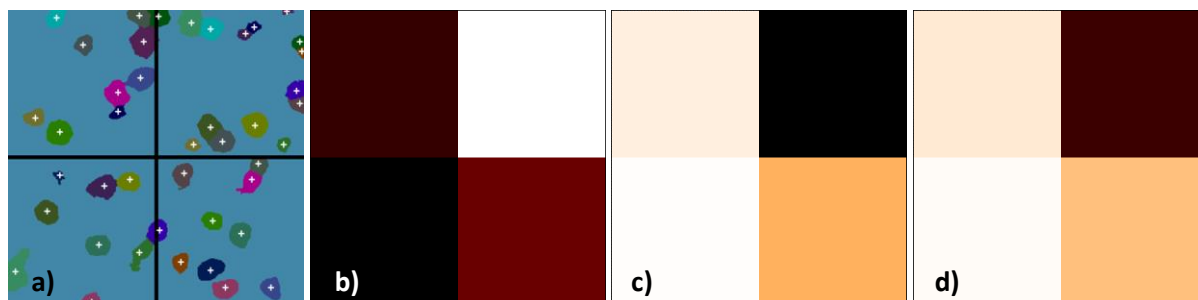


Figure 83 : a) Image originale segmentée de taille 512x512 pixels contenant des neurones individualisés et labélisés. Le centre de gravité des neurones est marqué par une croix blanche. b-d) Illustration de la création de la cartographie à partir de l'image originale vers l'image basse résolution de taille 2x2 pixels (agrégation de 256 pixels en x et 256 pixels en y). b) Cartographie du nombre de neurones. La valeur de chaque pixel de gauche à droite, de haut en bas est respectivement de 9, 15, 8 et 10 neurones. c) Cartographie du rayon moyen des neurones calculée par le « mode centre de gravité ». La valeur de chaque pixel est respectivement de 18,86, 14,85, 18,98 et 18,34 pixels. d) Cartographie du rayon moyen des neurones calculée par le « mode contribution ». La valeur de chaque pixel est respectivement de 18,78, 14,82, 18,97 et 18,34 pixels.

Il est important de rappeler que le calcul de cartographies synthétiques est un problème qui a été déjà abordé par le passé avec par exemple une approche agrégeant les pixels d'une image originale haute résolution en utilisant pour valeur d'intensité des nouveaux pixels créés la fraction d'occupation d'une segmentation. Ce type de cartographie est généralement appelée « *heatmap* ». Un exemple de ce type d'approche est illustré par Vandenberghe et ses collègues dans le cadre de la maladie d'Alzheimer sur des modèles murins (Vandenberghe et al., 2016). Notre principale contribution dans ce travail est d'étendre cette notion de « *heatmap* » au calcul de paramètres biologiques potentiellement plus informatifs que la fraction d'occupation. Pour cela, nous tirons bénéfice de la production de données massives raffinées (neurones individualisés) extrêmement difficile à produire. La synthèse de ces informations raffinées sous forme de cartographies est somme toute naturelle et permet de produire des images quantitatives avec une forte valeur ajoutée pour les biologistes (études exploratoires pour détecter des variations subtiles ou non, accès à des cartographies locales et non plus à des résultats à l'échelle de structures anatomiques). Le potentiel d'application en biologie de l'approche que nous proposons nous semble très important et sera exploré prochainement à MIRCen.

3.4. Exemples de cartographies obtenues pour des neurones

Comme détaillé dans la section précédente, plusieurs types de cartographie peuvent être générés en fonction des paramètres descripteurs. Cette section présente plusieurs types de cartographies concernant les neurones générées à partir des données utilisées pour le comptage stéréologique (section 2.2.2.4).

3.4.1. Cartographies du nombre de neurones

Le comptage du nombre de neurones dans des régions anatomiques du cerveau représente un centre d'intérêt majeur pour les biologistes. La méthode que nous proposons permet une estimation précise du nombre de neurones. Dans cet exemple, nous générons tout d'abord la cartographie du nombre de neurones par le « mode centre de gravité » en considérant la somme du nombre de neurones pour descripteur. L'agrégation des pixels a été fixée à 512x512 pixels ce qui produit une nouvelle résolution de 112,64 μm en X et Y. La Figure 84 présente les cartographies du nombre de neurones dans les 6 régions d'intérêt correspondant aux données utilisées pour le comptage stéréologique : cortex (Figure 84a/g), noyau caudé (Figure 84b), thalamus (Figure 84c/e), claustrum (Figure 84h), gyrus denté (Figure 84f) et subiculum (Figure 84d).

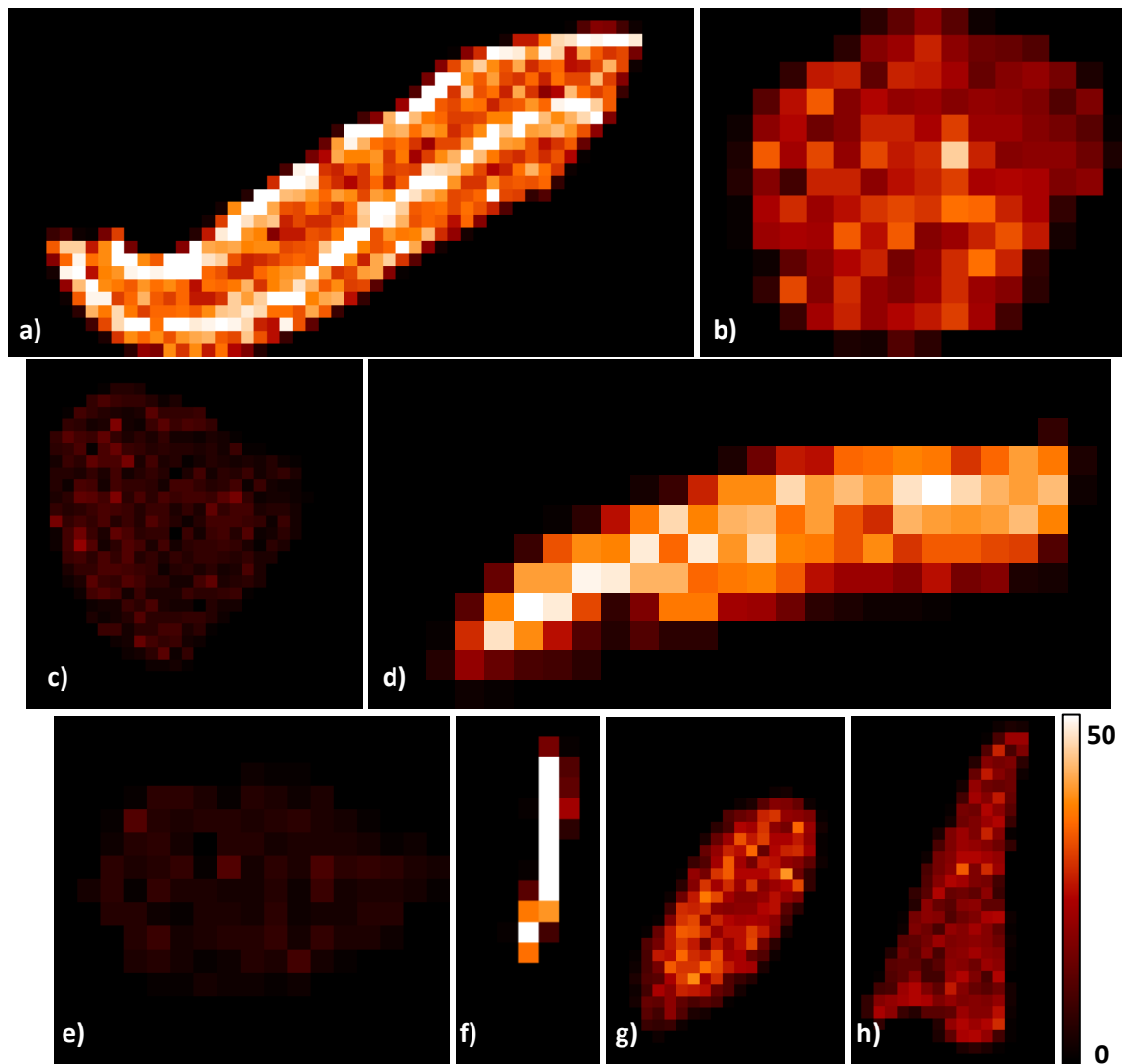


Figure 84 : Cartographies du nombre de neurones dans 6 régions d'intérêt calculées dans le « mode centre de gravité », agrégation de 512 pixels en X et 512 pixels en Y. La résolution est de 0,11264 mm/pixel. a) S1_cortex, le nombre varie de 0 à 69. b) S2_noyau_caudé, le nombre varie de 0 à 45. c) S3_thalamus, le nombre varie de 0 à 11. d) S4_subiculum, le nombre varie de 0 à 53. e) S5_thalamus, le nombre varie de 0 à 20. f) S6_gyrus_denté, le nombre varie de 0 à 88. g) S7_cortex, le nombre varie de 0 à 39. h) S8_claustrum, le nombre varie de 0 à 33.

Cette cartographie reflète la densité des neurones pour les différentes régions d'intérêt. Visuellement, les régions de plus fortes densités à moins fortes densités sont dans l'ordre : le gyrus denté, le cortex, le subiculum, le noyau caudé, le claustrum et le thalamus. Dans le cortex, nous observons également des bandes de neurones très denses qui correspondent à différentes couches connues du cortex cérébral (Figure 9). Ce type d'information pourrait par exemple être exploité pour segmenter les couches corticales qui constituent un domaine de recherche très actif en neurosciences.

3.4.2. Cartographies de la morphologie des neurones

La morphologie des neurones est également un centre d'intérêt pour les biologistes. Plusieurs caractéristiques calculées par la méthode proposée permettent de calculer des paramètres morphologiques des neurones. Ici, pour illustrer ce cas, nous avons généré des cartographies de

divers paramètres morphologiques en considérant la valeur moyenne des descripteurs étudiés. L'agrégation de pixels a été fixée à 512x512 pixels comme précédemment.

- Cartographies du rayon moyen des neurones

La Figure 85 présente les cartographies du rayon moyen des neurones pour les 6 régions d'intérêt. Nous pouvons observer les mêmes bandes que précédemment dans le cortex (Figure 85a). Dans la bande la plus dense (correspondant à la bande la plus claire dans la Figure 84a), le rayon moyen des neurones est plus petit par rapport à la bande la plus foncée, et inversement. Dans le noyau caudé (Figure 85b) et le subiculum (Figure 85d), les neurones sur les bords sont plus petits que ceux du centre. Dans S3_thalamus (Figure 85c), nous observons un gradient de la taille du rayon moyen des neurones qui sont plus petits à droite et plus grands à gauche. De plus, dans S5_thalamus (Figure 85e), les neurones présentent des tailles homogènes mais différentes de celles de S3_thalamus ce qui nous laisse penser qu'il serait possible d'utiliser ce paramètre pour segmenter des sous-régions du thalamus. La vérification sur l'image histologique à la résolution originale a permis de valider ces observations (Figure 85c/e). Dans le gyrus denté (Figure 85f), les neurones présentent un rayon moyen similaire ce qui produit une cartographie homogène. La taille des neurones est quant à elle stable dans le S7_cortex (Figure 85g) et dans le claustrum (Figure 85h).

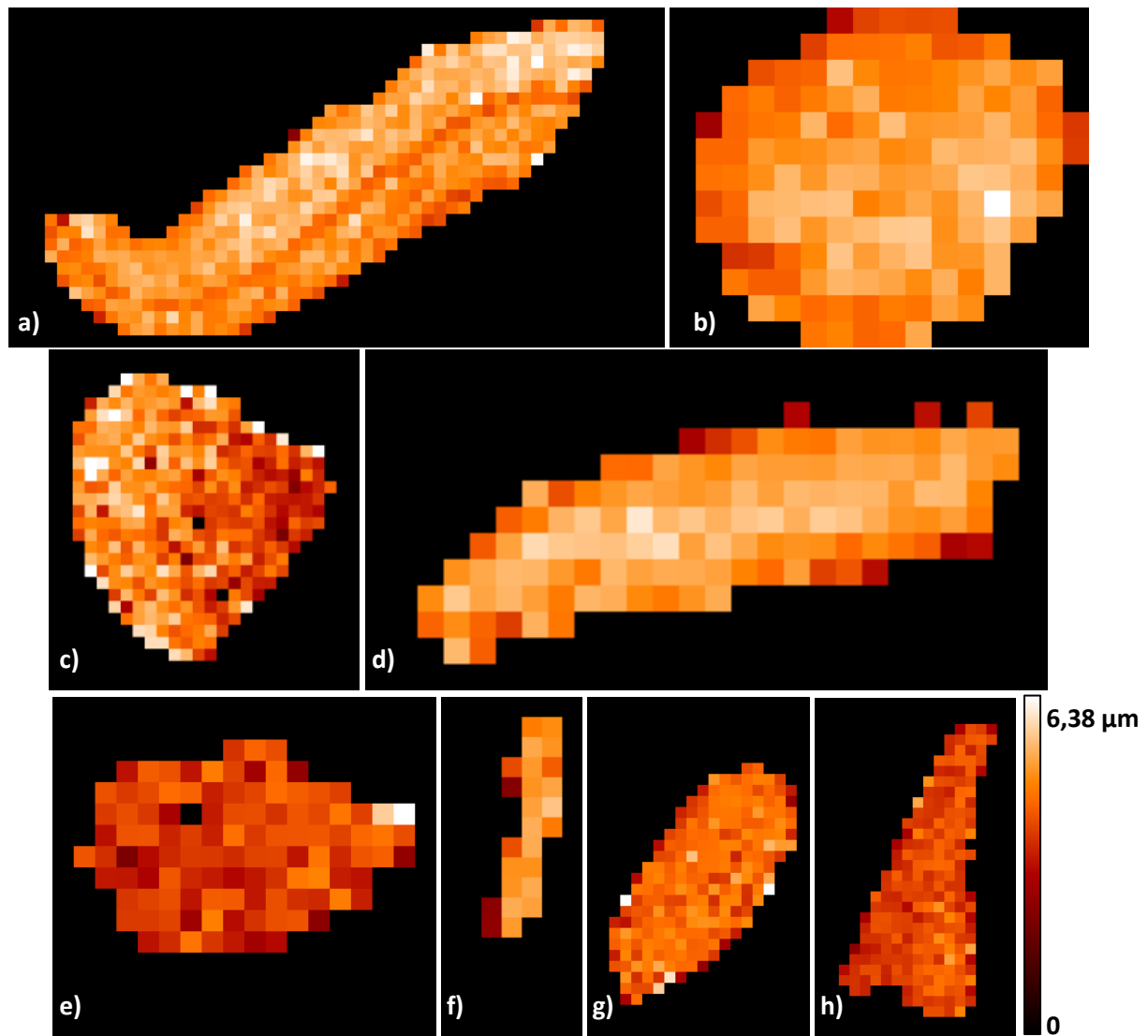


Figure 85 : Cartographies du rayon moyen des neurones dans 6 régions d'intérêt calculées dans le « mode contribution », agrégation de 512 pixels en X et 512 pixels en Y . La résolution est de 0,11264 mm/pixel. a) S1_cortex, le rayon moyen varie de 0 à 6,44 μm . b) S2_noyau_caudé, le rayon moyen varie de 0 à 6,66 μm . c) S3_thalamus, le rayon moyen varie de 0 à 6,40 μm . d) S4_subiculum, le rayon moyen varie de 0 à 6,07 μm . e) S5_thalamus, le rayon moyen varie de 0 à 7,21 μm . f) S6_gyrus_denté, le rayon moyen varie de 0 à 5,59 μm . g) S7_cortex, le rayon moyen varie de 0 à 6,79. h) S8_claustrum, le rayon moyen varie de 0 à 5,21 μm .

- Cartographies de la circularité moyenne des neurones

La Figure 86 présente des cartographies de la circularité moyenne des neurones. Les bandes observées précédemment dans le S1_cortex apparaissent mais de façon moins prononcée. Au centre du cortex, une bande de circularité plus haute, dans laquelle les neurones sont plus ronds, apparaît. Dans le noyau caudé et le subiculum, les neurones sont plus ronds à la frontière. Les neurones du thalamus, du claustrum, du gyrus denté et de la région S7_cortex présentent des cartographies plus contrastées ce qui dénote une distribution de neurones avec des formes plus hétérogènes.

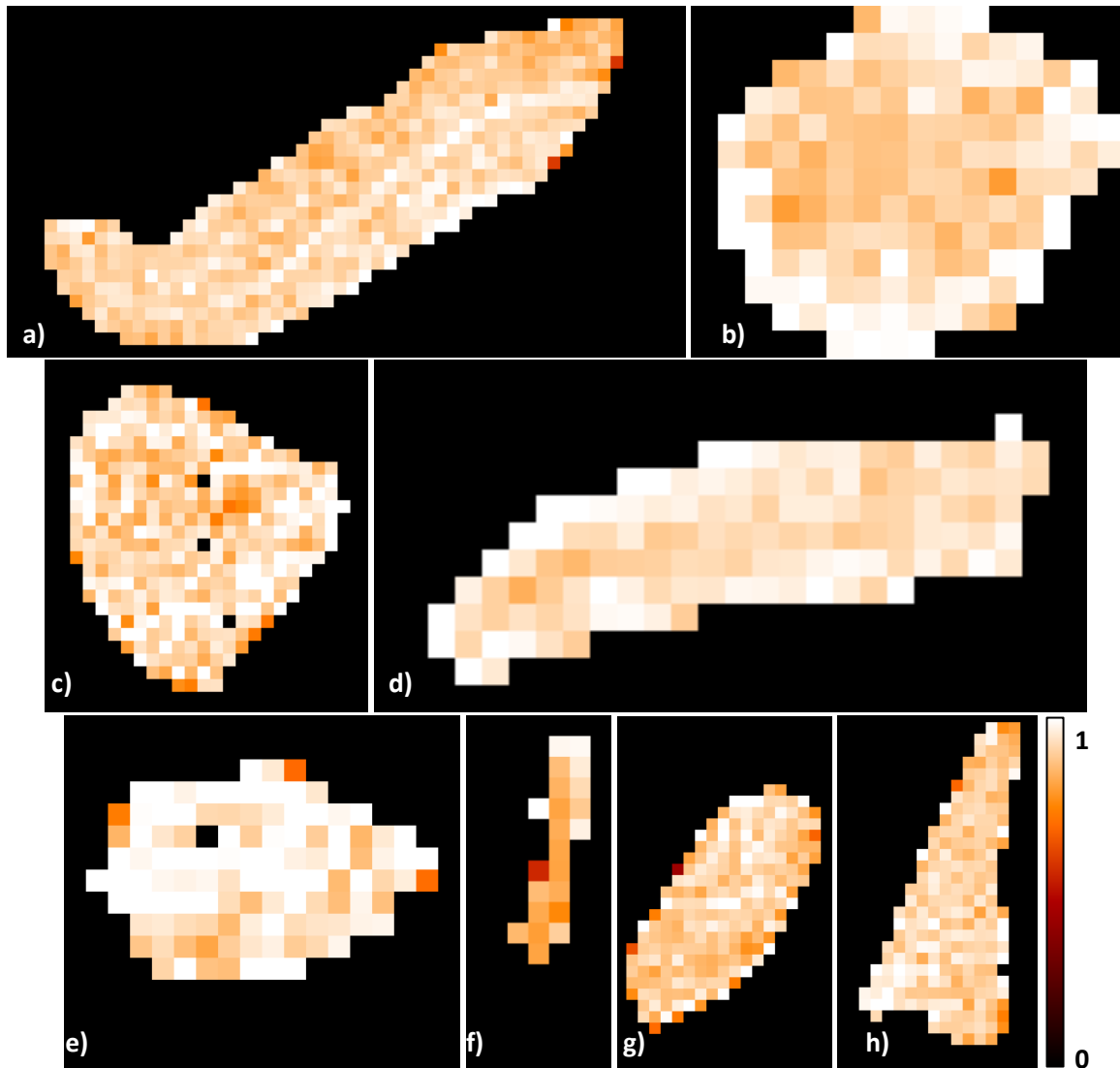


Figure 86 : Cartographies de la circularité de neurones dans 6 régions d'intérêt calculées dans le « mode centre de gravité », agrégation de 512 pixels en X et 512 pixels en Y . La résolution est 0,11264 mm/pixel. a) S1_cortex. b) S2_noyau_caudé. c) S3_thalamus. d) S4_subiculum. e) S5_thalamus. f) S6_gyrus_denté. g) S7_cortex. h) S8_claustrum.

3.4.3. Cartographies d'organisation des neurones

L'organisation des neurones dans des régions d'intérêt peut être quantifiée grâce par exemples aux cartographies d'orientation et de voisinage des neurones.

- Cartographies d'orientation moyenne des neurones

L'orientation d'un neurone est définie comme présenté dans la Figure 87. La direction verticale correspond à l'angle 0° . L'angle mesuré entre la dimension la plus grande du neurone (axe principal) et la direction verticale définit l'orientation du neurone. Le long du sens antihoraire, l'orientation des neurones augmente et le long du sens horaire, l'orientation des neurones diminue. Dans notre cas, les valeurs d'orientation moyenne varient de -90° à 90° .

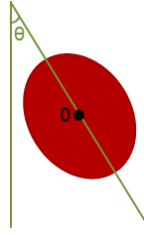


Figure 87 : Schéma du calcul de l'orientation d'un neurone.

La cartographie de l'orientation moyenne des neurones présente de grandes bandes régulières dans le cortex, le noyau caudé, le claustrum, le gyrus denté et le subiculum (Figure 88). Il est intéressant de noter que dans le cortex, nous ne retrouvons pas les bandes très fines et intenses ce qui tend à montrer que malgré des densités de neurones très différentes, l'orientation des neurones est stable et varie de façon continue. A l'inverse, l'orientation des neurones est beaucoup plus variable dans le thalamus, ce dernier est connu pour être constitué de nombreux sous-noyaux avec des organisations cellulaires différentes (Dauguet et al., 2009).

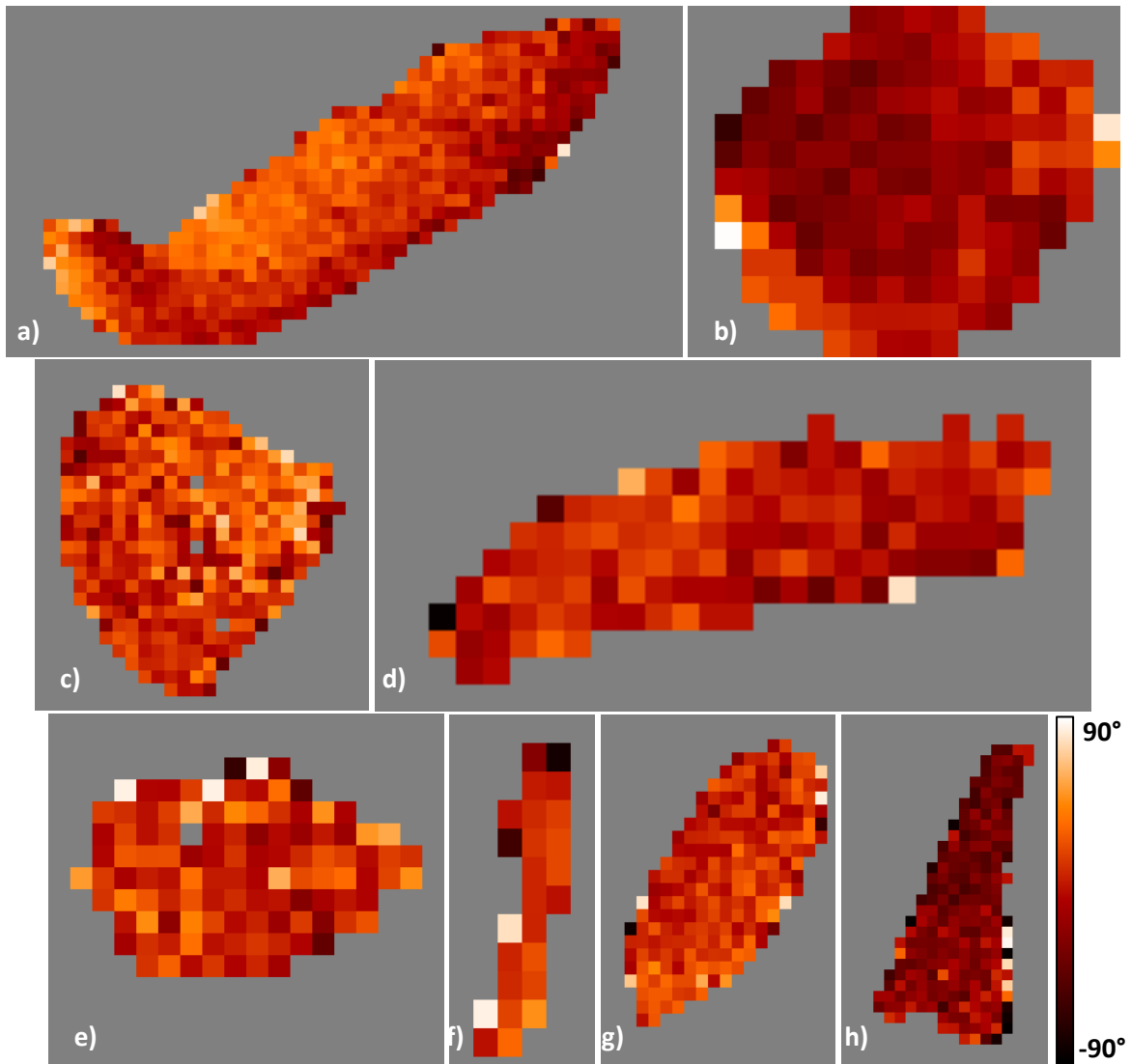


Figure 88 : Cartographies de l'orientation des neurones dans 6 régions d'intérêt calculées dans le « mode centre de gravité », agrégation de 512 pixels en X et 512 pixels en Y. La résolution est 0,11264 mm/pixel. a) S1_cortex. b) S2_noyau_caudé. c) S3_thalamus. d) S4_subiculum. e) S5_thalamus. f) S6_gyrus_denté. g) S7_cortex. h) S8_claustrum.

- Cartographies de la distance moyenne inter-neurones

La Figure 89 présente des cartographies de la distance moyenne entre une population de neurones et leurs voisins. Plus la distance est faible, plus cette région est dense et inversement. Dans le cortex, les bandes déjà repérées précédemment apparaissent dans la cartographie du nombre de neurones. Nous observons également que, dans le thalamus, les neurones sont plus espacés que dans les autres régions. Enfin, le gyrus denté est une région très dense dans laquelle les neurones sont très proches les uns des autres.

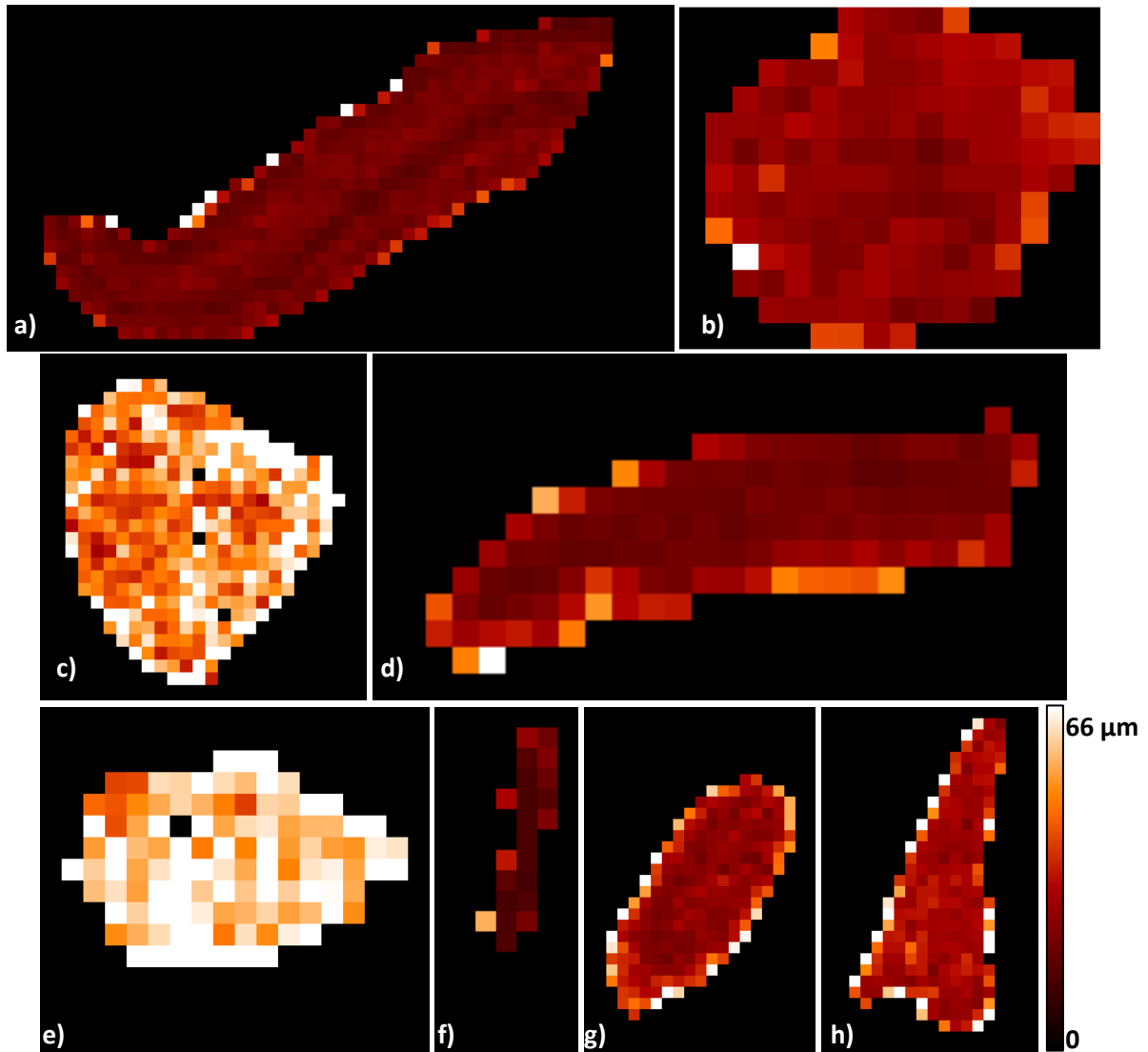


Figure 89 : Cartographies de la distance moyenne entre les neurones et leurs voisins par région d'intérêt calculées dans le « mode centre de gravité », agrégation de 512 pixels en X et 512 pixels en Y. La résolution est 0,11264 mm/pixel. a) S1_cortex, la distance moyenne varie de 0 à 149,79 μm. b) S2_noyau_caudé, la distance moyenne varie de 0 à 67,78 μm. c) S3_thalamus, la distance moyenne varie de 0 à 281,35 μm. d) S4_subiculum, la distance moyenne varie de 0 à 73,25 μm. e) S5_thalamus, la distance moyenne varie de 0 à 95,75 μm. f) S6_gyrus_denté, la distance moyenne varie de 0 à 54,93 μm. g) S7_cortex, la distance moyenne varie de 0 à 147,57 μm. h) S8_claustrum, la distance moyenne varie de 0 à 115,87 μm.

A travers les différents exemples présentés, nous avons mis en évidence la grande diversité des cartographies produites (Figure 90). Ceci traduit la très grande richesse des informations qui sont accessibles à partir des neurones individualisés. Si certaines applications dans le domaine de la biologie semblent évidentes avec ce type approche (segmentation de régions / sous-régions

anatomiques, suivi de la réorganisation des cellules dans le cadre de maladies neurodégénératives, effet de thérapies à l'échelle cellulaire, etc.) le cadre d'application pourrait être élargi à d'autres marqueurs et objets biologiques pour peu qu'ils respectent les hypothèses que nous avons présenté (objet compact, gradient d'intensité positif ou négatif du centre vers la périphérie). Un travail important reste à réaliser dans cette perspective.

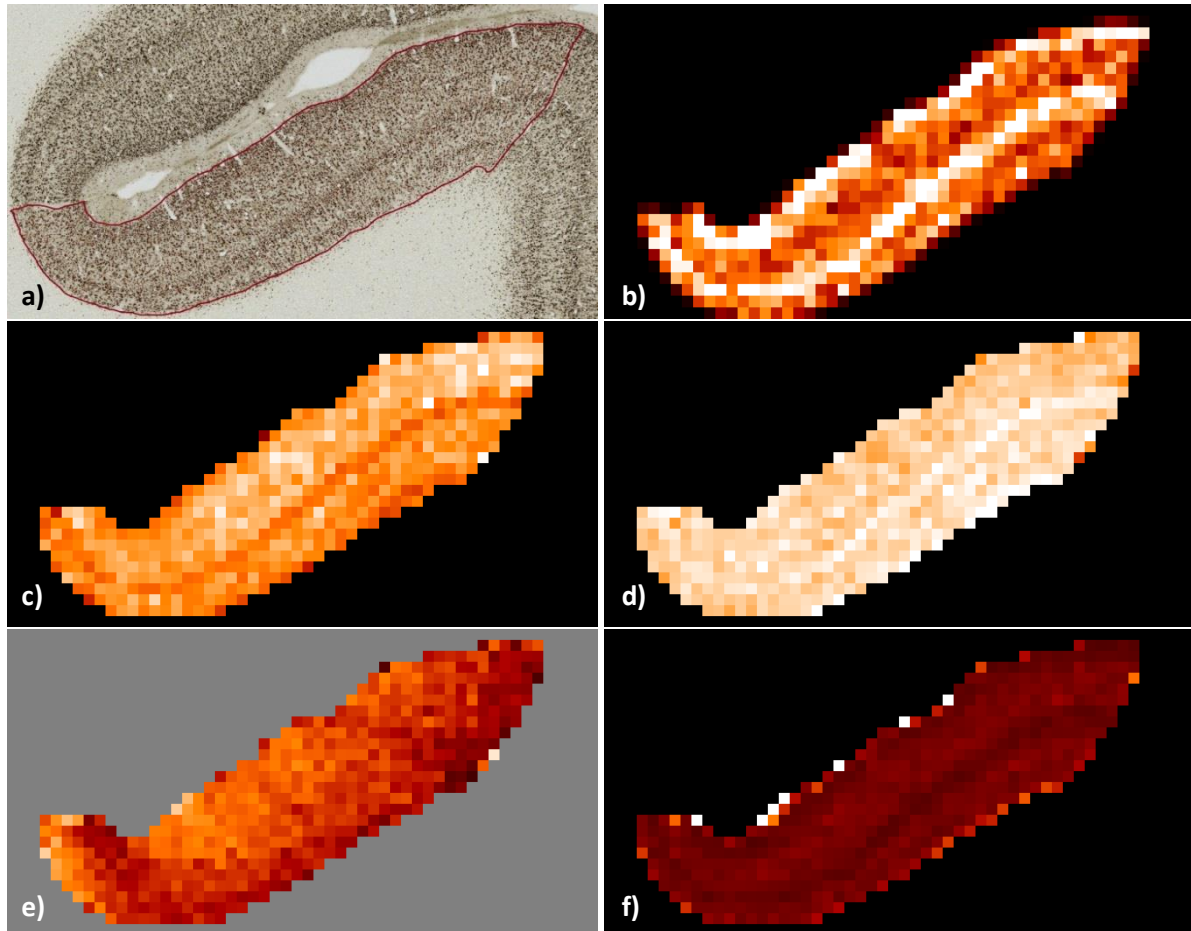


Figure 90 : Exemple de la diversité des cartographies produites pour la région du cortex. a) Image originale. b) Cartographie du nombre de neurones. c) Cartographie du rayon moyen de neurones. d) Cartographie de la circularité des neurones. e) Cartographie de l'orientation de neurones. f) Cartographie de la distance moyenne entre un neurone et ses voisins. Les cartographies b), d), e) et f) sont calculées dans le « mode centre de gravité » tandis que la cartographie c) est calculée dans le « mode contribution ».

3.5. Conclusion

L'histologie permet d'avoir une très haute résolution spatiale jusqu'au niveau cellulaire, générant également des données très lourdes. Les cartographies présentées dans cette section permettent de synthétiser l'information pertinente, de transformer ces données qualitatives massives en données quantitatives plus informatives et plus simples à analyser. La méthode proposée caractérise d'abord les neurones, puis permet d'analyser des régions d'intérêt de petite dimension. Il sera possible, à terme, d'étendre ce type d'analyse à des régions anatomiques, à des coupes voire à des cerveaux entiers. Après reconstruction 3D des cerveaux à partir de séries de cartographies quantitatives, il sera possible de mener des études de groupes par des méthodes de type *voxel-wise* comme précédemment réalisé dans notre équipe sur des données d'autoradiographies quantitatives (Dubois et al., 2008 ; Dubois et al., 2010). A partir des cartographies générées, nous avons observé que les différentes régions anatomiques présentaient pour la plupart plusieurs paramètres descripteurs

permettant de visualiser des sous-régions spécifiques (couches neuronales dans la région du cortex par exemple). Enfin, comme montré dans cette section, la combinaison de certains paramètres est intéressante et pourrait être utilisée pour étudier l'organisation voire segmenter des régions / sous-régions anatomiques. Ceci ouvre de plus la voie à des analyses combinant des images produites *post mortem* et *in vivo* pour peu que les données *post mortem* aient été reconstruites en 3D. Il sera par exemple possible de comparer les contrastes des signaux observés dans ces modalités complémentaires.

4. Discussion et perspectives

Mon projet de thèse avait deux objectifs méthodologiques principaux : le développement d'une méthodologie générale permettant de segmenter et d'individualiser les neurones marqués par l'anticorps NeuN dans des images histologiques de cerveau de primate et le développement d'une approche d'intégration des informations relatives aux neurones obtenues à l'échelle cellulaire sous la forme de cartographies quantitatives produites à une échelle mésoscopique. Tous ces travaux ont été réalisés à partir de trois coupes histologiques de cerveau d'un macaque sain, un modèle animal proche de l'Homme qui joue un rôle essentiel dans la recherche translationnelle.

Dans le **premier chapitre** de ce manuscrit, nous avons présenté le système nerveux central dont le neurone constitue l'unité fonctionnelle de base. Ensuite, des études relatives aux neurones ont été présentées pour situer le contexte. Dans ces études, les modèles animaux jouent un rôle de première importance pour la recherche cérébrale tant physiologique que physiopathologique. Ces études s'appuient sur les différentes techniques d'imagerie *in vivo* (IRM, TEP principalement) et *post mortem* (histologie). Dans le cadre de mon projet de thèse, nous nous sommes intéressés plus particulièrement aux neurones en utilisant la technique d'imagerie de la microscopie optique parce qu'elle permet d'accéder à une très grande diversité ainsi qu'un très grand nombre de marquages histologiques. De plus, cette technique d'imagerie permet d'étudier finement la structure du cerveau à l'échelle cellulaire avec une résolution pouvant atteindre quelques dixièmes de micromètres. D'autres techniques de microscopie existent et offrent de meilleures résolutions mais nous avons focalisé notre travail sur une résolution cellulaire qui constitue déjà un défi majeur à relever et nous avons utilisé des technologies somme toute standards disponibles à MIRCen. Nous avons utilisé la technique de l'immunohistochimie pour marquer les neurones à l'aide de l'anticorps NeuN sur des coupes histologiques de cerveau de macaque sain. Enfin, nous avons présenté les problématiques liées à l'étude des neurones : la variabilité de l'intensité de coloration des neurones, l'accolement de plusieurs neurones, la variabilité de la taille des neurones, le comptage précis des neurones dans différentes régions anatomiques présentant des niveaux d'hétérogénéité importants.

Dans le **deuxième chapitre**, nous avons tout d'abord présenté les méthodes existantes qui permettent de segmenter des images sous forme de classes et d'individualiser des cellules. Sur la base de cet état de l'art et de l'expérience de notre équipe dans ce domaine, nous avons choisi la méthode de *Random Forest* (RF) en déterminant les caractéristiques optimales (L, M, V, Lbp 40) pour segmenter les neurones. Il est intéressant de noter que nous retrouvons logiquement des caractéristiques de colorimétrie souvent utilisées dans ce type d'étude mais également un indice de texture. Le choix de ces caractéristiques est dépendant de l'étude biologique, des objets d'intérêt, du marqueur et une étape d'optimisation devrait être considérée pour les différents cas. Deux experts, une biologiste (b1) et un traiteur d'image (i1), ont segmenté les neurones et le tissu non marqué sur 100 imageries, afin d'obtenir une segmentation de référence utilisable pour créer les modèles de segmentation et valider notre méthode. Les résultats de la segmentation ont montré que la méthode de RF segmente mieux les neurones par rapport à la méthode bioVision, surtout pour les neurones clairs. La méthode de RF améliore la qualité de segmentation des neurones clairs d'un facteur 5 par rapport à la méthode bioVision ($F\text{-score}_{RF}=0,37$, $F\text{-score}_{bioVision}=0,07$). Cependant, les scores de segmentation obtenus sur ces neurones clairs restent insuffisants et devront être améliorés soit en modifiant la qualité des images histologiques voire la technique de numérisation soit en adaptant la méthode de segmentation. La principale difficulté résulte dans la présence de couleurs similaires dans ces neurones clairs mais aussi dans du tissu non marqué ce qui rend la segmentation parfois très difficile. Malgré cela, tenir compte au mieux de ces neurones est important pour améliorer le comptage global des neurones. Pour les neurones segmentés sous la forme d'une classe, nous

trouvons des neurones individuels mais également des amas de neurones de tailles très variables. Dans certains cas extrêmes, des milliers de neurones sont accolés comme dans le gyrus denté qui constitue le cas le plus complexe à analyser et pour lequel un traitement spécifique a été proposé. Afin de les individualiser, nous avons développé une méthode originale. Nous avons tout d'abord itéré 2 fois un processus en deux étapes (lissage gaussien et filtrage min-max) pour renforcer en même temps l'information relative aux centroïdes et aux contours des neurones qui correspondent à des informations d'intérêt pour réaliser l'individualisation. A notre connaissance, la création de cette carte des extrema est originale. Puis, un modèle basé sur des contours déformables a ensuite été utilisé pour individualiser les neurones en effectuant simultanément les calculs sur tous les neurones d'une composante connexe (processus de compétition). Afin d'individualiser les neurones de tailles variables qui sont accolés ou se chevauchent du fait d'une superposition dans l'épaisseur de la coupe histologique, nous avons proposé une méthode multi-échelle qui permet de produire d'abord tous les résultats d'individualisation possibles et ensuite de chercher la frontière la plus probable séparant des neurones accolés en fonction de l'indice de Dice qui mesure la similarité entre deux segmentations. La recherche de la solution optimale basée sur une carte de σ adaptative (détection des centroïdes) en calculant toutes les possibilités en fonction des valeurs de σ est un choix original qui introduit cependant un coût calculatoire important. Le résultat en est une combinaison locale des σ . A ce stade, le comptage des neurones par notre méthode est déjà effectif car il est conditionné strictement par le nombre de centroïdes détectés. Il serait possible de modifier cela en introduisant *a posteriori* un filtrage des neurones individualisés selon des critères de forme ou de colorimétrie. Le problème de la superposition des neurones est critique car nous ne disposons que d'une mise au point pour toute l'épaisseur de la coupe histologique (fonction de la technique de numérisation). Tout un gradient d'effets de la superposition existe : en cas de faible superposition, nous pouvons traiter le cas comme si les deux neurones étaient justes accolés mais en cas de superposition totale, notre méthode devient inopérante. La possibilité d'acquérir une série d'images selon la profondeur (*Z-stack*) est envisageable mais augmenterait significativement les temps de numérisation, le volume des données à traiter et demanderait des développements logiciels importants.

Pour **évaluer notre méthode de comptage des neurones**, nous avons demandé à un expert en traitement d'image (i2) de segmenter les centroïdes des neurones. Nous avons comparé la performance entre notre méthode et la méthode *Watershed* grâce au calcul de *F-score*. Visuellement, notre méthode individualise mieux les neurones par rapport à la méthode de *Watershed*. Cette validation qualitative est confirmée quantitativement, notre méthode, dont la valeur moyenne du *F-score* est égale à 0,86, améliore la performance de 11,6% par rapport à la méthode *Watershed* dont la valeur moyenne du *F-score* est égale à 0,77. Ces résultats sont basés sur des critères de comptage des neurones qui sont la référence dans ce type de travail. Cependant, l'analyse visuelle a pointé un problème important qui est que les contours des neurones détectés par différentes méthodes ainsi que la localisation des centroïdes détectés peuvent être très différents. Pour **évaluer notre méthode d'individualisation en tenant compte de ces informations**, nous avons demandé à l'expert i1 de segmenter les centroïdes et les contours des neurones en utilisant une interface graphique développée dans BrainVISA et spécifiquement dédiée à cette tâche. Comme il serait trop fastidieux de segmenter de dizaines voire des centaines de milliers de neurones, nous avons proposé une stratégie de tirage aléatoire d'un nombre jugé suffisant de neurones à segmenter (100 tirages pour plusieurs images de 5000x5000 pixels). Plusieurs caractéristiques de ces populations des neurones ont été calculées (distance moyenne entre les centroïdes et les contours) entre les neurones segmentés manuellement et ceux segmentés automatiquement par notre méthode et par la méthode de *Watershed*. Nous avons observé que notre méthode fournissait quantitativement la meilleure qualité de segmentation par rapport à la méthode *Watershed*. Pour ce qui est du

pourcentage de la superposition des neurones segmentés manuellement et automatiquement (coefficient de Dice), notre méthode améliore la qualité de la segmentation en moyenne de 7% par rapport au *Watershed*. Pour ce qui est de la distance entre les centroïdes segmentés manuellement et automatiquement, les centroïdes segmentés par la méthode *Watershed* sont 1,1 fois plus éloignés que ceux segmentés par notre méthode. Pour ce qui est de la distance des contours segmentés manuellement et automatiquement, c'est 1,3 fois plus éloigné pour la méthode *Watershed*. Le fait de pouvoir ainsi évaluer plus finement la qualité de la segmentation des neurones individualisés constitue une contribution originale qui repose sur des techniques plutôt issues du domaine du traitement de l'image. Afin de pousser plus loin notre validation, il nous a paru très important d'utiliser la technique de référence utilisée par les biologistes, à savoir la stéréologie. Ceci constitue **une étape indispensable pour valider la faisabilité d'utiliser nos méthodes dans des études biologiques réelles**. Nous avons donc demandé à deux biologistes (b2 et b3) d'estimer par stéréologie le nombre de neurones dans plusieurs régions anatomiques. Ces régions ont été choisies en accord avec les biologistes pour leur intérêt tout en essayant d'avoir le gradient le plus large possible de densités de neurones que l'on peut trouver dans le cerveau. Les résultats de l'individualisation des neurones obtenus par notre méthode sont très fortement corrélés ($corr_{densité}=0,99-0,994$) avec ceux estimés par stéréologie. Ces résultats préliminaires montrent le fort potentiel de notre méthode pour compter les neurones. L'utilisation de notre méthode pourrait d'ores et déjà être testée dans les études de comptage de neurones à venir à MIRCen en conservant les comptages de référence pour confirmer nos premiers résultats.

Les travaux de validation menés dans notre recherche constituent une contribution originale par la quantité de données traitées mais également par la diversité de techniques utilisées. Pas moins de 5 experts (3 biologistes et 2 traiteurs d'images) ont participé à ces tâches longues et fastidieuses. Si l'on résume, nous avons demandé à deux experts (biologiste b1 et traiteur d'image i1) de faire la **segmentation des neurones sous forme de classe sur 100 imageries** de 512x512 pixels de la base d'images dédiées à la segmentation. Il a été demandé à un autre expert (i2) de marquer les centroïdes sur 50 images représentatives de 5000x5000 pixels de la base d'images utilisée pour valider l'individualisation des neurones. Ceci a permis d'obtenir la position de **111971 centroïdes de neurones pointés manuellement**. Le traiteur d'image i1 a également segmenté complètement **les centroïdes et les contours de 900 neurones** sur 9 de ces 50 images. De plus, **approximativement 37000 neurones** ont été comptés par la technique de stéréologie par les deux biologistes (b2 et b3) sur les 8 régions anatomiques de cette base d'images. Ces **3 bases de données de référence** annotées par des experts sont précieuses et pourront être exploitées au-delà de ce travail pour mener de nouvelles études (segmentation, individualisation, comptage, etc.).

La méthode d'analyse des neurones développée pendant ma thèse donne des résultats très prometteurs. Cependant, des pistes d'amélioration subsistent.

Pour **améliorer la qualité de la segmentation**, il serait possible d'ajouter des caractéristiques supplémentaires, par exemple Lbp 10, pour garder plus d'informations sur de petits neurones. Pour les neurones clairs, il sera difficile d'améliorer significativement la performance de la segmentation. Nous avons observé que les intensités du marquage NeuN dans les neurones sont très variables : très claires dans le thalamus, moyennes dans le noyau caudé et très foncées dans le cortex, le gyrus denté et le subiculum. Il est probable que l'expression de la protéine est variable en fonction du type de neurone et de la région cérébrale ce qui oriente plutôt vers une adaptation des marqueurs en fonction des neurones ciblés.

En ce qui concerne l'individualisation des neurones, notre méthode donne de bons résultats dans l'absolu si l'on compare avec les données de référence. Cependant, il est important de souligner que

les pourcentages d'erreur calculés doivent être interprétés avec beaucoup de prudence. En effet, nous avons également observé des différences non négligeables entre les experts ce qui tend à démontrer que nos données de références ne sont pas parfaites. Il est impossible, dans ce contexte, de savoir s'il y a une accumulation ou une compensation de toutes ces erreurs. La supériorité de notre méthode par rapport à la méthode *Watershed* a également été démontrée même si nous ne pouvons garantir que l'implémentation et les paramètres utilisés à cette occasion soient optimaux pour cette méthode. La création d'une base de référence pourrait permettre à différents laboratoires travaillant dans ce domaine de comparer de façon plus objective les performances de leurs méthodes respectives. Bien que ces résultats soient satisfaisants, il serait possible d'améliorer notre méthode, par exemple en ajoutant des étapes de post-traitement pour fusionner des sur-individualisations ou parcelliser des sous-individualisations en se basant sur des critères d'asymétrie du gradient de coloration ou de texture dans chaque neurone. Dans la littérature, il existe beaucoup de méthodes relatives à l'identification de graines comme les centroïdes des neurones. L'exactitude de la segmentation des neurones dépend de la fiabilité de la détection des centroïdes. Une perspective de recherche pourrait être de réaliser la détection des centroïdes par une approche de type *deep learning*, une méthode très utilisée actuellement. Comme nous disposons des détections manuelles des centroïdes des neurones pour 50 images de 5000x5000 pixels, cette vérité terrain pourrait servir à configurer le réseau utilisé par le *deep learning*. Nous pourrions alors utiliser la liste exhaustive de tous les centroïdes détectés en faisant varier la valeur de sigma pour définir localement les meilleurs candidats. Si cette piste donnait de meilleurs résultats, il serait alors possible d'utiliser la même méthode de croissance de région pour réaliser la segmentation des neurones individuels. A partir des centroïdes détectés, de nouvelles recherches pourraient être engagées pour exploiter directement ces résultats sous la forme de nuages de points (Andrey et al., 2010; Loménie et Racoceanu, 2012).

L'individualisation des neurones est un sujet important dans le domaine des neurosciences. De ce fait, de nombreuses études et méthodes ont été proposées. En 2014, Schmitz et al. ont comparé les méthodes disponibles pour la segmentation cellulaire dans leur publication intitulée « *Current automated 3D cell detection methods are not a suitable replacement for manual stereologic cell counting* ». Ils montrent que l'individualisation des cellules est un problème très complexe et ils insistent sur la nécessité de développer de nouvelles approches dans ce domaine. Si plusieurs travaux ont été réalisés ces dernières années, la plupart des méthodes proposées sont basées sur la morphologie mathématique, la détection des concavités, la technique *Watershed* voire l'algorithme de *graph-cut*. La plupart de ces méthodes se sont concentrées sur des cas relativement simples ou très spécifiques. Lorsque nous arrivons dans des cas complexes, par exemple, le gyrus denté où des milliers de neurones sont agrégés, ces méthodes ne pourraient pas segmenter correctement les neurones.

Dans le **troisième chapitre**, nous avons présenté le travail réalisé pour synthétiser les informations relatives aux neurones individualisés à l'aide des caractéristiques morphologiques et organisationnelles sous la forme de cartographies produites à l'échelle mésoscopique. Une cinquantaine de paramètres ont été étudiés ce qui peut paraître beaucoup si l'on compare avec les mesures classiquement obtenues dans des études biologiques. Le plus souvent, ces études se limitent à effectuer un comptage dans les principales régions d'intérêt. L'approche que nous proposons est plus exploratoire et devrait permettre de détecter des phénomènes plus subtils tels des variations se produisant sur un type de cellules particulier, des réorganisations de la cyto-architecture, etc. Un autre intérêt de cette approche est de produire une information à un niveau local ce qui augmente la pertinence de l'étude par rapport à une estimation globale menée sur des régions entières. Si des cartographies de fraction d'occupation appelées « *heatmaps* » ont déjà été

proposées, la production de cartographies quantitatives basée sur des paramètres très divers augmente significativement le champ des explorations possibles par exemple pour proposer des segmentations automatiques de régions et sous-régions anatomiques ou encore pour étudier des sous-populations de neurones (localisation, organisation, évolution au cours du temps ou de la maladie, etc.).

La prochaine étape de cette recherche sera de mettre en œuvre cette méthode d'analyse des neurones sur des coupes entières de cerveaux de primates sains ou pathologiques. L'application de notre méthode à des modèles pathologiques est d'ores et déjà envisagée et ne devrait pas introduire de difficulté supplémentaire car, dans ces cas, une perte neuronale est attendue ce qui devrait simplifier le traitement. Pour cela, des limitations fortes liées à la taille des données devront tout de même être dépassées. Cela concerne aussi bien l'environnement informatique matériel à adapter (réseau, stockage, puissance de calcul, etc.) que l'environnement logiciel (lecture / écriture en accès partiels aux données, adaptation des codes de traitement de l'image, parallélisation, etc.). Une partie de ces problèmes ont d'ores et déjà été réglés mais des développements très lourds sont encore en cours. En s'appuyant sur l'expertise de l'équipe traitement de l'image de MIRCen dans le domaine de la reconstruction et de l'analyse *post mortem* 3D, de nouvelles opportunités s'ouvriront pour ce nouveau type d'imagerie *post mortem* quantitative : étudier le développement / vieillissement du cerveau, caractériser précisément des modèles animaux, tester l'effet de nouveaux médicaments (quantifier la mort neuronale) ou comparer en 3D des résultats d'imagerie produits *in vivo* et *post mortem* (interprétation des signaux physiques mesurés à partir de paramètres biologiques).

Annexe

1. Base d'images pour la segmentation des classes tissulaires

Tableau 20: Liste de la base d'images pour la segmentation. Côté: L (left) – à gauche, R (right) – à droite, C (central) – au centre. A/V (Apprentissage/Validation) : base d'apprentissage / base de validation. Les dimensions des images sont de 512x512 pixels.

| N° | Nom des images | Côté | Région | Position X | Position Y | A/V |
|----|---|------|-----------------|------------|------------|-----|
| 1 | CJ1301_slide81unsharp_BG_R | R | background | 215024 | 100457 | A |
| 2 | CJ1301_slide81unsharp_CA1_R | R | CA1 | 160066 | 146153 | A |
| 3 | CJ1301_slide81unsharp_CA2_R | R | CA2 | 155891 | 139728 | A |
| 4 | CJ1301_slide81unsharp_CA3_R | R | CA3 | 147641 | 136853 | A |
| 5 | CJ1301_slide81unsharp_CA4_R | R | CA4 | 148456 | 145163 | A |
| 6 | CJ1301_slide81unsharp_caudate_L | L | caudate | 81930 | 64713 | V |
| 7 | CJ1301_slide81unsharp_caudate_R | R | caudate | 135517 | 65280 | A |
| 8 | CJ1301_slide81unsharp_claustrum_L | L | claustrum | 43161 | 106864 | V |
| 9 | CJ1301_slide81unsharp_claustrum_R | R | claustrum | 174451 | 109137 | A |
| 10 | CJ1301_slide81unsharp_ctx_L_layer1 | L | cortex layer1 | 6538 | 114411 | V |
| 11 | CJ1301_slide81unsharp_ctx_L_layer2 | L | cortex layer2 | 7375 | 114628 | A |
| 12 | CJ1301_slide81unsharp_ctx_L_layer3 | L | cortex layer3 | 9386 | 114274 | V |
| 13 | CJ1301_slide81unsharp_ctx_L_layer4 | L | cortex layer4 | 10962 | 114238 | V |
| 14 | CJ1301_slide81unsharp_ctx_L_layer5a | L | cortex layer5a | 11864 | 113558 | A |
| 15 | CJ1301_slide81unsharp_ctx_L_layer5b | L | cortex layer5b | 12980 | 113494 | A |
| 16 | CJ1301_slide81unsharp_ctx_L_layer6a | L | cortex layer6a | 14313 | 113494 | A |
| 17 | CJ1301_slide81unsharp_ctx_L_layer6b | L | cortex layer6b | 15736 | 113172 | V |
| 18 | CJ1301_slide81unsharp_ctx_R_layer1 | R | cortex layer1 | 214234 | 92503 | A |
| 19 | CJ1301_slide81unsharp_ctx_R_layer2 | R | cortex layer2 | 213394 | 92727 | V |
| 20 | CJ1301_slide81unsharp_ctx_R_layer3 | R | cortex layer3 | 211980 | 93207 | A |
| 21 | CJ1301_slide81unsharp_ctx_R_layer4 | R | cortex layer4 | 210620 | 93159 | A |
| 22 | CJ1301_slide81unsharp_ctx_R_layer5a | R | cortex layer5a | 209532 | 93303 | V |
| 23 | CJ1301_slide81unsharp_ctx_R_layer5b | R | cortex layer5b | 208508 | 93127 | V |
| 24 | CJ1301_slide81unsharp_ctx_R_layer6a | R | cortex layer6a | 207388 | 93463 | V |
| 25 | CJ1301_slide81unsharp_ctx_R_layer6b | R | cortex layer6b | 205276 | 94103 | A |
| 26 | CJ1301_slide81unsharp_DG_R | R | dentate gyrus | 153366 | 145403 | A |
| 27 | CJ1301_slide81unsharp_globus_pallidus_ext_L | L | globus pallidus | 59954 | 101070 | V |
| 28 | CJ1301_slide81unsharp_globus_pallidus_ext_R | R | globus pallidus | 155441 | 101799 | A |
| 29 | CJ1301_slide81unsharp_globus_pallidus_int_L | L | globus pallidus | 67746 | 108801 | A |
| 30 | CJ1301_slide81unsharp_globus_pallidus_int_R | R | globus pallidus | 150434 | 103746 | V |
| 31 | CJ1301_slide81unsharp_hypothal | C | hypothal | 107307 | 128570 | A |
| 32 | CJ1301_slide81unsharp_mam_body_L | L | mammillary_body | 103854 | 120974 | A |
| 33 | CJ1301_slide81unsharp_mam_body_R | R | mammillary_body | 110739 | 121899 | V |
| 34 | CJ1301_slide81unsharp_putamen_L | L | putamen | 51147 | 84883 | A |
| 35 | CJ1301_slide81unsharp_putamen_R | R | putamen | 166812 | 91399 | V |
| 36 | CJ1301_slide81unsharp_subiculum_para_R | R | para-subiculum | 137257 | 147603 | A |
| 37 | CJ1301_slide81unsharp_subiculum_R | R | subiculum | 148216 | 150803 | A |
| 38 | CJ1301_slide81unsharp_thalamus_L | L | thalamus | 86018 | 91673 | A |

| | | | | | | |
|----|--------------------------------------|---|----------------|--------|--------|---|
| 39 | CJ1301_slide81unsharp_thalamus_R | R | thalamus | 132095 | 87392 | V |
| 40 | CJ1301_slide81unsharp_ventricle_R | R | ventricle | 124922 | 65719 | V |
| 41 | CJ1301_slide81unsharp_WM_L | L | white matter | 48104 | 114674 | V |
| 42 | CJ1301_slide81unsharp_WM_R | R | white matter | 156246 | 62111 | A |
| 43 | CJ1301_slide101unsharp_BG_R | R | background | 212245 | 52644 | V |
| 44 | CJ1301_slide101unsharp_CA1_L | L | CA1 | 49025 | 158817 | A |
| 45 | CJ1301_slide101unsharp_CA1_R | R | CA1 | 167285 | 161394 | V |
| 46 | CJ1301_slide101unsharp_CA2_L | L | CA2 | 42084 | 143295 | A |
| 47 | CJ1301_slide101unsharp_CA2_R | R | CA2 | 177325 | 144103 | V |
| 48 | CJ1301_slide101unsharp_CA3a_L | L | CA3a | 44549 | 137657 | A |
| 49 | CJ1301_slide101unsharp_CA3a_R | R | CA3a | 177356 | 140325 | V |
| 50 | CJ1301_slide101unsharp_CA3b_L | L | CA3b | 49255 | 138429 | V |
| 51 | CJ1301_slide101unsharp_CA3b_R | R | CA3b | 170600 | 143485 | A |
| 52 | CJ1301_slide101unsharp_CA4_L | L | CA4 | 49763 | 146237 | A |
| 53 | CJ1301_slide101unsharp_CA4_R | R | CA4 | 169976 | 148447 | V |
| 54 | CJ1301_slide101unsharp_caudate_L | L | caudate | 50258 | 125958 | A |
| 55 | CJ1301_slide101unsharp_caudate_R | R | caudate | 179257 | 128488 | V |
| 56 | CJ1301_slide101unsharp_ctx_L_layer1 | L | cortex layer1 | 7920 | 86958 | A |
| 57 | CJ1301_slide101unsharp_ctx_L_layer2 | L | cortex layer2 | 9075 | 80531 | V |
| 58 | CJ1301_slide101unsharp_ctx_L_layer3 | L | cortex layer3 | 10405 | 80533 | A |
| 59 | CJ1301_slide101unsharp_ctx_L_layer4 | L | cortex layer4 | 12079 | 80133 | A |
| 60 | CJ1301_slide101unsharp_ctx_L_layer5a | L | cortex layer5a | 14044 | 76592 | V |
| 61 | CJ1301_slide101unsharp_ctx_L_layer5b | L | cortex layer5b | 16993 | 74252 | V |
| 62 | CJ1301_slide101unsharp_ctx_L_layer6a | L | cortex layer6a | 18092 | 74491 | V |
| 63 | CJ1301_slide101unsharp_ctx_L_layer6b | L | cortex layer6b | 20348 | 74728 | A |
| 64 | CJ1301_slide101unsharp_ctx_R_layer1 | R | cortex layer1 | 230933 | 75485 | V |
| 65 | CJ1301_slide101unsharp_ctx_R_layer2 | R | cortex layer2 | 229580 | 73350 | A |
| 66 | CJ1301_slide101unsharp_ctx_R_layer3 | R | cortex layer3 | 228587 | 73598 | V |
| 67 | CJ1301_slide101unsharp_ctx_R_layer4 | R | cortex layer4 | 227383 | 74206 | V |
| 68 | CJ1301_slide101unsharp_ctx_R_layer5a | R | cortex layer5a | 224744 | 70457 | A |
| 69 | CJ1301_slide101unsharp_ctx_R_layer5b | R | cortex layer5b | 224692 | 72436 | A |
| 70 | CJ1301_slide101unsharp_ctx_R_layer6a | R | cortex layer6a | 223797 | 72825 | A |
| 71 | CJ1301_slide101unsharp_ctx_R_layer6b | R | cortex layer6b | 221626 | 73142 | V |
| 72 | CJ1301_slide101unsharp_DG_L | L | dentate gyrus | 48819 | 149736 | A |
| 73 | CJ1301_slide101unsharp_DG_R | R | dentate gyrus | 169343 | 151390 | V |
| 74 | CJ1301_slide101unsharp_habenula_L | L | habenula | 109521 | 89758 | V |
| 75 | CJ1301_slide101unsharp_habenula_R | R | habenula | 128654 | 90860 | A |
| 76 | CJ1301_slide101unsharp_LG_L_1 | L | LG | 68660 | 116983 | A |
| 77 | CJ1301_slide101unsharp_LG_L_2 | L | LG | 66256 | 115925 | V |
| 78 | CJ1301_slide101unsharp_LG_L_3 | L | LG | 65241 | 113883 | A |
| 79 | CJ1301_slide101unsharp_LG_L_4 | L | LG | 64272 | 112680 | V |
| 80 | CJ1301_slide101unsharp_LG_L_5 | L | LG | 60837 | 111409 | A |
| 81 | CJ1301_slide101unsharp_LG_L_6 | L | LG | 60914 | 110889 | V |
| 82 | CJ1301_slide101unsharp_LG_R_1 | R | LG | 164301 | 118741 | V |
| 83 | CJ1301_slide101unsharp_LG_R_2 | R | LG | 166026 | 118416 | A |
| 84 | CJ1301_slide101unsharp_LG_R_3 | R | LG | 167298 | 116469 | V |

| | | | | | | |
|-----|---|---|----------------|--------|--------|---|
| 85 | CJ1301_slide101unsharp_LG_R_4 | R | LG | 168376 | 118300 | A |
| 86 | CJ1301_slide101unsharp_LG_R_5 | R | LG | 170150 | 112041 | A |
| 87 | CJ1301_slide101unsharp_LG_R_6 | R | LG | 171612 | 111158 | V |
| 88 | CJ1301_slide101unsharp_PAG_L | L | PAG | 112445 | 108877 | V |
| 89 | CJ1301_slide101unsharp_PAG_R | R | PAG | 124524 | 108709 | A |
| 90 | CJ1301_slide101unsharp_pons_L | L | pons | 108715 | 140626 | A |
| 91 | CJ1301_slide101unsharp_pons_R | R | pons | 125184 | 141824 | V |
| 92 | CJ1301_slide101unsharp_subiculum_para_L | L | para-subiculum | 62704 | 138606 | A |
| 93 | CJ1301_slide101unsharp_subiculum_para_R | R | para-subiculum | 161407 | 134685 | V |
| 94 | CJ1301_slide101unsharp_subiculum_L | L | subiculum | 56393 | 153016 | A |
| 95 | CJ1301_slide101unsharp_subiculum_R | R | subiculum | 161588 | 149664 | V |
| 96 | CJ1301_slide101unsharp_thalamus_L | L | thalamus | 75488 | 91524 | V |
| 97 | CJ1301_slide101unsharp_thalamus_R | R | thalamus | 164357 | 97270 | A |
| 98 | CJ1301_slide101unsharp_ventricle_R | R | ventricle | 143068 | 77944 | A |
| 99 | CJ1301_slide101unsharp_WM_L | L | white matter | 93897 | 63118 | A |
| 100 | CJ1301_slide101unsharp_WM_R | R | white matter | 142766 | 64349 | V |

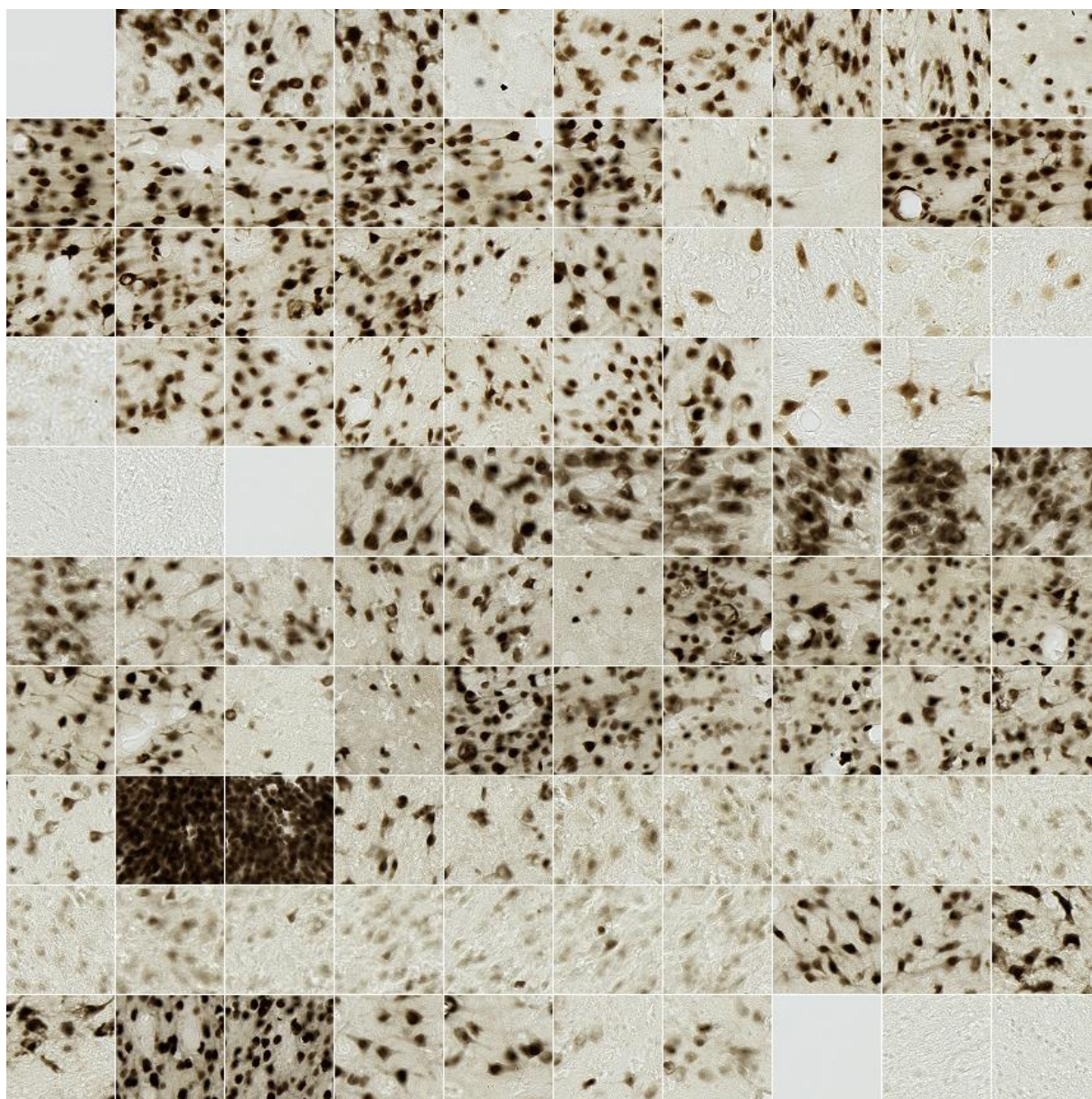


Figure 91 : Base d'images utilisées pour la segmentation. De gauche à droite, de haut à bas : images de N°1 à N° 100.

2. Base d'images pour l'individualisation des neurones

Tableau 21: Liste de la base d'images utilisées pour l'individualisation des neurones. Côté: L (left) – à gauche, R (right) – à droite, C (central) – au centre. Les dimensions des images sont de 5000x5000 pixels.

| N° | Nom des images | Côté | Région | Position X | Position Y |
|----|-----------------------------------|------|-------------|------------|------------|
| 1 | CJ1301_slide91unsharp_caudate_L | L | caudate | 82522 | 67838 |
| 2 | CJ1301_slide91unsharp_caudate_R_1 | R | caudate | 150612 | 69601 |
| 3 | CJ1301_slide91unsharp_caudate_R_2 | R | caudate | 175410 | 132107 |
| 4 | CJ1301_slide91unsharp_claustrum_L | L | claustrum | 48891 | 96789 |
| 5 | CJ1301_slide91unsharp_claustrum_R | R | claustrum | 178852 | 97930 |
| 6 | CJ1301_slide91unsharp_ctx_L_1 | L | cortex | 108174 | 61247 |
| 7 | CJ1301_slide91unsharp_ctx_L_2 | L | cortex | 111775 | 47451 |
| 8 | CJ1301_slide91unsharp_ctx_L_3 | L | cortex | 92715 | 52930 |
| 9 | CJ1301_slide91unsharp_ctx_L_4 | L | cortex | 113133 | 11625 |
| 10 | CJ1301_slide91unsharp_ctx_L_5 | L | cortex | 46576 | 34018 |
| 11 | CJ1301_slide91unsharp_ctx_L_6 | L | cortex | 30036 | 88960 |
| 12 | CJ1301_slide91unsharp_ctx_L_7 | L | cortex | 18414 | 71644 |
| 13 | CJ1301_slide91unsharp_ctx_L_8 | L | cortex | 39760 | 119015 |
| 14 | CJ1301_slide91unsharp_ctx_L_9 | L | cortex | 12737 | 129645 |
| 15 | CJ1301_slide91unsharp_ctx_L_10 | L | cortex | 13244 | 148385 |
| 16 | CJ1301_slide91unsharp_ctx_L_11 | L | cortex | 34526 | 148654 |
| 17 | CJ1301_slide91unsharp_ctx_R_1 | R | cortex | 129792 | 60821 |
| 18 | CJ1301_slide91unsharp_ctx_R_2 | R | cortex | 131056 | 44222 |
| 19 | CJ1301_slide91unsharp_ctx_R_3 | R | cortex | 144342 | 47234 |
| 20 | CJ1301_slide91unsharp_ctx_R_4 | R | cortex | 123157 | 17547 |
| 21 | CJ1301_slide91unsharp_ctx_R_5 | R | cortex | 191482 | 37605 |
| 22 | CJ1301_slide91unsharp_ctx_R_6 | R | cortex | 188387 | 88838 |
| 23 | CJ1301_slide91unsharp_ctx_R_7 | R | cortex | 215311 | 73538 |
| 24 | CJ1301_slide91unsharp_ctx_R_8 | R | cortex | 189168 | 118593 |
| 25 | CJ1301_slide91unsharp_ctx_R_9 | R | cortex | 215041 | 133395 |
| 26 | CJ1301_slide91unsharp_ctx_R_10 | R | cortex | 214521 | 153002 |
| 27 | CJ1301_slide91unsharp_ctx_R_11 | R | cortex | 199475 | 162973 |
| 28 | CJ1301_slide91unsharp_hip_L_1 | L | hippocampus | 57130 | 159664 |
| 29 | CJ1301_slide91unsharp_hip_L_2 | L | hippocampus | 50320 | 152057 |
| 30 | CJ1301_slide91unsharp_hip_L_3 | L | hippocampus | 50324 | 143779 |
| 31 | CJ1301_slide91unsharp_hip_L_4 | L | hippocampus | 61567 | 140170 |
| 32 | CJ1301_slide91unsharp_hip_L_5 | L | hippocampus | 74853 | 137205 |
| 33 | CJ1301_slide91unsharp_hip_L_6 | L | hippocampus | 81808 | 139695 |
| 34 | CJ1301_slide91unsharp_hip_L_7 | L | hippocampus | 72156 | 145516 |
| 35 | CJ1301_slide91unsharp_hip_L_8 | L | hippocampus | 58265 | 149115 |
| 36 | CJ1301_slide91unsharp_hip_R_1 | R | hippocampus | 164647 | 164902 |
| 37 | CJ1301_slide91unsharp_hip_R_2 | R | hippocampus | 174314 | 155545 |
| 38 | CJ1301_slide91unsharp_hip_R_3 | R | hippocampus | 168837 | 148803 |
| 39 | CJ1301_slide91unsharp_hip_R_4 | R | hippocampus | 154655 | 147824 |
| 40 | CJ1301_slide91unsharp_hip_R_5 | R | hippocampus | 147370 | 142851 |
| 41 | CJ1301_slide91unsharp_hip_R_6 | R | hippocampus | 141545 | 148185 |

| | | | | | |
|----|-------------------------------------|---|--------------|---------|--------|
| 42 | CJ1301_slide91unsharp_hip_R_7 | R | hippocampus | 159898 | 154049 |
| 43 | CJ1301_slide91unsharp_hip_R_8 | R | hippocampus | 165822 | 157309 |
| 44 | CJ1301_slide91unsharp_putamen_L_1 | L | putamen | 57213 | 95283 |
| 45 | CJ1301_slide91unsharp_putamen_L_2 | L | putamen | 53517 | 114512 |
| 46 | CJ1301_slide91unsharp_putamen_R | R | putamen | 175690 | 117341 |
| 47 | CJ1301_slide91unsharp_subiculum_L_1 | L | subiculum | 70588 | 153176 |
| 48 | CJ1301_slide91unsharp_subiculum_L_2 | L | subiculum | 77069 | 152620 |
| 49 | CJ1301_slide91unsharp_subiculum_R_1 | R | subiculum | 152133 | 156150 |
| 50 | CJ1301_slide91unsharp_subiculum_R_2 | R | subiculum | 145144 | 158084 |
| 51 | CJ1301_slide91unsharp_thalamus_L_1 | L | thalamus | 78962 | 95737 |
| 52 | CJ1301_slide91unsharp_thalamus_L_2 | L | thalamus | 94158 | 93285 |
| 53 | CJ1301_slide91unsharp_thalamus_R_1 | R | thalamus | 152284 | 96060 |
| 54 | CJ1301_slide91unsharp_thalamus_R_2 | R | thalamus | 135308 | 95043 |
| 55 | CJ1301_slide91unsharp_vtcl_L | L | ventricle | 92600 | 67573 |
| 56 | CJ1301_slide91unsharp_vtcl_med | C | ventricle | 1158889 | 76022 |
| 57 | CJ1301_slide91unsharp_WM_L | L | white matter | 75488 | 58650 |
| 58 | CJ1301_slide91unsharp_WM_R | R | white matter | 154657 | 60546 |

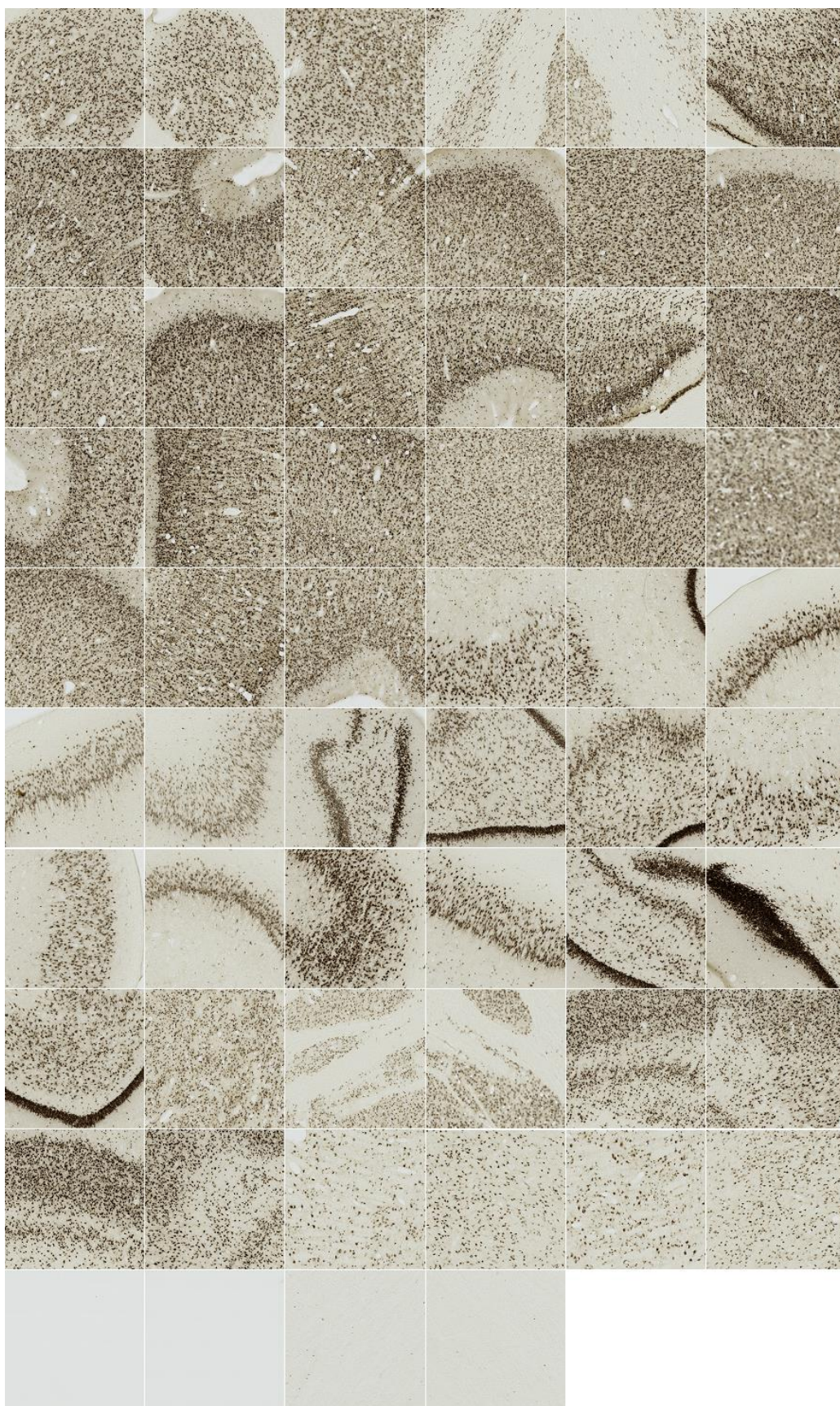


Figure 92 : Base d'images utilisées pour l'individualisation. De gauche à droite, de haut à bas : images de N°1 à N°58.

3. Valeurs moyennes des différents paramètres des données d'individualisation par région anatomique

| Région | Paramètres | Méthode (RF et bioVision) et Experts (i1 et b1) | | | |
|-------------|-------------|---|----------------|----------------|----------------|
| | | RF_i1 | RF_b1 | BV_i1 | BV_b1 |
| noyau caudé | fraction | 0,135±0,026 | 0,193±0,043 | 0,109±0,021 | 0,127±0,025 |
| | N_{cc} | 1444,5±243,952 | 1181,5±126,572 | 1391±229,103 | 1342,5±224,153 |
| | N_e | 1880,5±208,597 | | | |
| | complexité | 1,308±0,077 | 1,591±0,006 | 1,358±0,074 | 1,407±0,080 |
| | N_a | 2025±360,624 | 2181,5±398,101 | 1837,5±316,077 | 1928±335,169 |
| | N_t | 1708±254,558 | 1713±254,558 | 1656,5±256,680 | 1687,5±256,680 |
| | N_{sous2} | 147,5±48,790 | 159±50,912 | 154,5±40,305 | 157,5±44,548 |
| | N_{sur} | 317±106,066 | 468,5±143,543 | 181±59,397 | 240,5±78,489 |
| | N_m | 23±2,828 | 8,5±4,95 | 69,5±7,778 | 35,5±3,536 |
| | Rappel | 0,906±0,035 | 0,909±0,035 | 0,879±0,039 | 0,895±0,037 |
| | Précision | 0,846±0,025 | 0,788±0,027 | 0,903±0,016 | 0,877±0,019 |
| | $F-score$ | 0,874±0,003 | 0,843±0,001 | 0,890±0,012 | 0,886±0,008 |
| claustrum | fraction | 0,036±0,015 | 0,059±0,017 | 0,019±0,013 | 0,025±0,016 |
| | N_{cc} | 832±234,759 | 964±65,054 | 547±306,884 | 628±322,441 |
| | N_e | 970,5±351,432 | | | |
| | complexité | 1,153±0,097 | 0,997±0,297 | 1,892±0,419 | 1,615±0,269 |
| | N_a | 899,5±328,805 | 1103±291,328 | 557±340,825 | 679,5±371,231 |
| | N_t | 793±275,772 | 854±289,914 | 518,5±313,248 | 617±322,441 |
| | N_{sous2} | 73,5±55,861 | 78±59,397 | 65±59,397 | 73±63,640 |
| | N_{sur} | 106,5±53,033 | 249±1,414 | 38,5±27,577 | 62,5±48,790 |
| | N_m | 104±19,799 | 38,5±2,121 | 387±21,213 | 280,5±34,648 |
| | Rappel | 0,819±0,013 | 0,884±0,021 | 0,509±0,138 | 0,616±0,109 |
| | Précision | 0,885±0,017 | 0,766±0,060 | 0,934±0,009 | 0,915±0,025 |
| | $F-score$ | 0,851±0,015 | 0,820±0,026 | 0,653±0,114 | 0,733±0,070 |
| cortex | fraction | 0,212±0,037 | 0,322±0,056 | 0,168±0,036 | 0,23±0,041 |
| | N_{cc} | 2020,6±225,971 | 1391,4±311,334 | 1961,2±180,422 | 1834,4±202,976 |
| | N_e | 3386,1±320,088 | | | |
| | complexité | 1,701±0,286 | 2,573±0,693 | 1,737±0,208 | 1,872±0,312 |
| | N_a | 3286,7±353,364 | 3682,3±407,684 | 2952,3±400,724 | 3146,1±381,624 |
| | N_t | 2924,8±283,661 | 2990,3±266,605 | 2748,9±337,367 | 2865,6±303,483 |
| | N_{sous2} | 383,1±106,126 | 383,7±91,367 | 394,4±112,412 | 408,3±100,791 |
| | N_{sur} | 362±126,011 | 692±220,316 | 203,4±89,528 | 280,5±114,228 |
| | N_m | 78,3±41,615 | 12,1±4,898 | 242,8±131,302 | 112,2±72,831 |
| | Rappel | 0,864±0,024 | 0,884±0,021 | 0,81±0,041 | 0,846±0,025 |
| | Précision | 0,891±0,031 | 0,815±0,043 | 0,933±0,023 | 0,913±0,028 |
| | $F-score$ | 0,877±0,014 | 0,847±0,021 | 0,866±0,020 | 0,877±0,011 |
| hippocampe | fraction | 0,111±0,063 | 0,161±0,075 | 0,091±0,059 | 0,107±0,063 |
| | N_{cc} | 693,9±231,862 | 603,9±207,733 | 576,8±229,043 | 582,1±210,689 |
| | N_e | 1432,8±905,651 | | | |
| | complexité | 2,259±2,175 | 2,582±2,309 | 2,673±2,365 | 2,680±2,519 |
| | N_a | 1522,4±837,459 | 1792,8±924,332 | 1318,9±843,378 | 1433,1±844,194 |
| | N_t | 1242,6±794,920 | 1289,6±794,190 | 1138,6±810,827 | 1200,6±803,072 |
| | N_{sous2} | 134,8±138,715 | 132,6±122,867 | 140,1±143,457 | 138,8±139,459 |
| | N_{sur} | 279,8±180,602 | 503,2±433,022 | 180,4±108,571 | 232,4±185,490 |
| | N_m | 55,4±73,618 | 10,6±15,646 | 154,1±163,983 | 93,3±117,495 |

| | | | | | |
|-----------|----------------|----------------|----------------|----------------|----------------|
| | Rappel | 0,864±0,073 | 0,907±0,033 | 0,763±0,162 | 0,823±0,114 |
| | Précision | 0,786±0,110 | 0,698±0,149 | 0,826±0,094 | 0,806±0,106 |
| | <i>F-score</i> | 0,816±0,062 | 0,778±0,094 | 0,782±0,119 | 0,807±0,804 |
| putamen | fraction | 0,073±0,046 | 0,113±0,070 | 0,044±0,037 | 0,057±0,045 |
| | N_{cc} | 1440±452,548 | 1380,5±286,378 | 1124±627,911 | 1213,5±546,594 |
| | N_e | 1718±602,455 | | | |
| | complexité | 1,186±0,046 | 1,226±0,182 | 1,634±0,377 | 1,451±0,157 |
| | N_a | 1596±690,136 | 1807,5±710,642 | 1193±765,090 | 1337,5±792,667 |
| | N_t | 1473,5±638,517 | 1575,5±601,748 | 1157,5±741,755 | 1291,5±757,311 |
| | N_{sous2} | 99±38,184 | 108,5±30,406 | 82±52,326 | 106±60,811 |
| | N_{sur} | 122,5±51,619 | 232±108,894 | 35,5±23,335 | 46±35,355 |
| | N_m | 145,5±74,246 | 34±29,698 | 478,5±191,626 | 320,5±215,669 |
| | Rappel | 0,844±0,076 | 0,912±0,031 | 0,637±0,208 | 0,719±0,189 |
| | Précision | 0,923±0,001 | 0,874±0,011 | 0,970±0,001 | 0,968±0,007 |
| | <i>F-score</i> | 0,881±0,042 | 0,892±0,009 | 0,759±0,153 | 0,817±0,122 |
| subiculum | fraction | 0,172±0,030 | 0,246±0,045 | 0,139±0,028 | 0,162±0,032 |
| | N_{cc} | 1573,5±73,333 | 1191,5±132,314 | 1544±113,772 | 1449,5±81,037 |
| | N_e | 2453,5±437,942 | | | |
| | complexité | 1,557±0,256 | 2,107±0,629 | 1,583±0,206 | 1,694±0,306 |
| | N_a | 2409,8±319,591 | 2572,3±333,200 | 2229,3±360,187 | 2305±331,086 |
| | N_t | 2200±344,158 | 2197,3±343,631 | 2135±353,096 | 2163,5±337,305 |
| | N_{sous2} | 220±80,527 | 244,8±92,439 | 219±72,650 | 232,3±84,235 |
| | N_{sur} | 209,8±27,789 | 375±25,781 | 94,3±9,639 | 141,5±11,387 |
| | N_m | 33,5±14,731 | 11,5±4,203 | 99,5±23,530 | 57,8±22,262 |
| | Rappel | 0,899±0,019 | 0,898±0,019 | 0,872±0,012 | 0,884±0,020 |
| | Précision | 0,911±0,022 | 0,852±0,025 | 0,957±0,005 | 0,937±0,012 |
| | <i>F-score</i> | 0,905±0,001 | 0,874±0,005 | 0,912±0,006 | 0,910±0,006 |
| thalamus | fraction | 0,041±0,008 | 0,060±0,010 | 0,027±0,005 | 0,033±0,006 |
| | N_{cc} | 621,8±116,259 | 635±113,819 | 454,5±85,625 | 505±96,201 |
| | N_e | 593,5±117,730 | | | |
| | complexité | 0,953±0,012 | 0,932±0,023 | 1,306±0,061 | 1,175±0,024 |
| | N_a | 692,5±137,624 | 772,3±151,520 | 491,5±91,860 | 557,8±105,967 |
| | N_t | 541,3±113,948 | 556±114,184 | 435,8±77,745 | 480,5±92,421 |
| | N_{sous2} | 27,8±9,069 | 29,5±10,408 | 24,5±11,902 | 28,5±10,083 |
| | N_{sur} | 151,3±46,486 | 216,3±51,448 | 55,8±20,532 | 77,3±35,519 |
| | N_m | 27,5±4,203 | 8±1,414 | 133,3±48,760 | 84,5±30,946 |
| | Rappel | 0,910±0,021 | 0,936±0,019 | 0,736±0,043 | 0,810±0,030 |
| | Précision | 0,782±0,047 | 0,720±0,036 | 0,888±0,027 | 0,863±0,049 |
| | <i>F-score</i> | 0,841±0,034 | 0,814±0,029 | 0,804±0,019 | 0,835±0,022 |

4. Distribution spatiale des neurones selon leur taille et la qualité des résultats de l'individualisation par rapport à la détection de référence

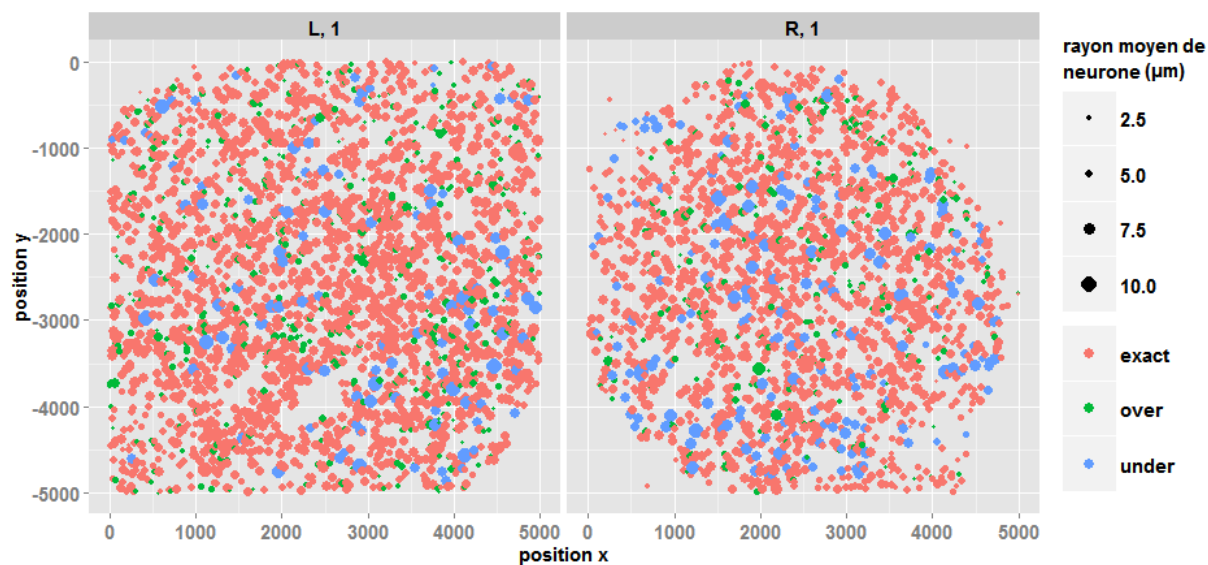


Figure 93 : Région du noyau caudé.

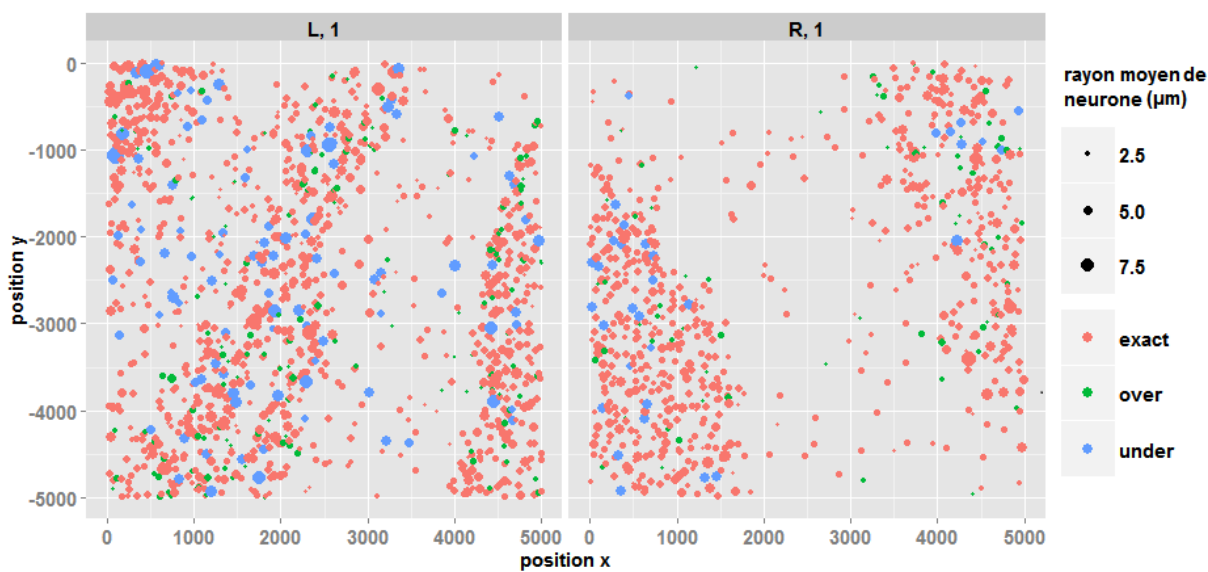


Figure 94 : Région du claustrum.

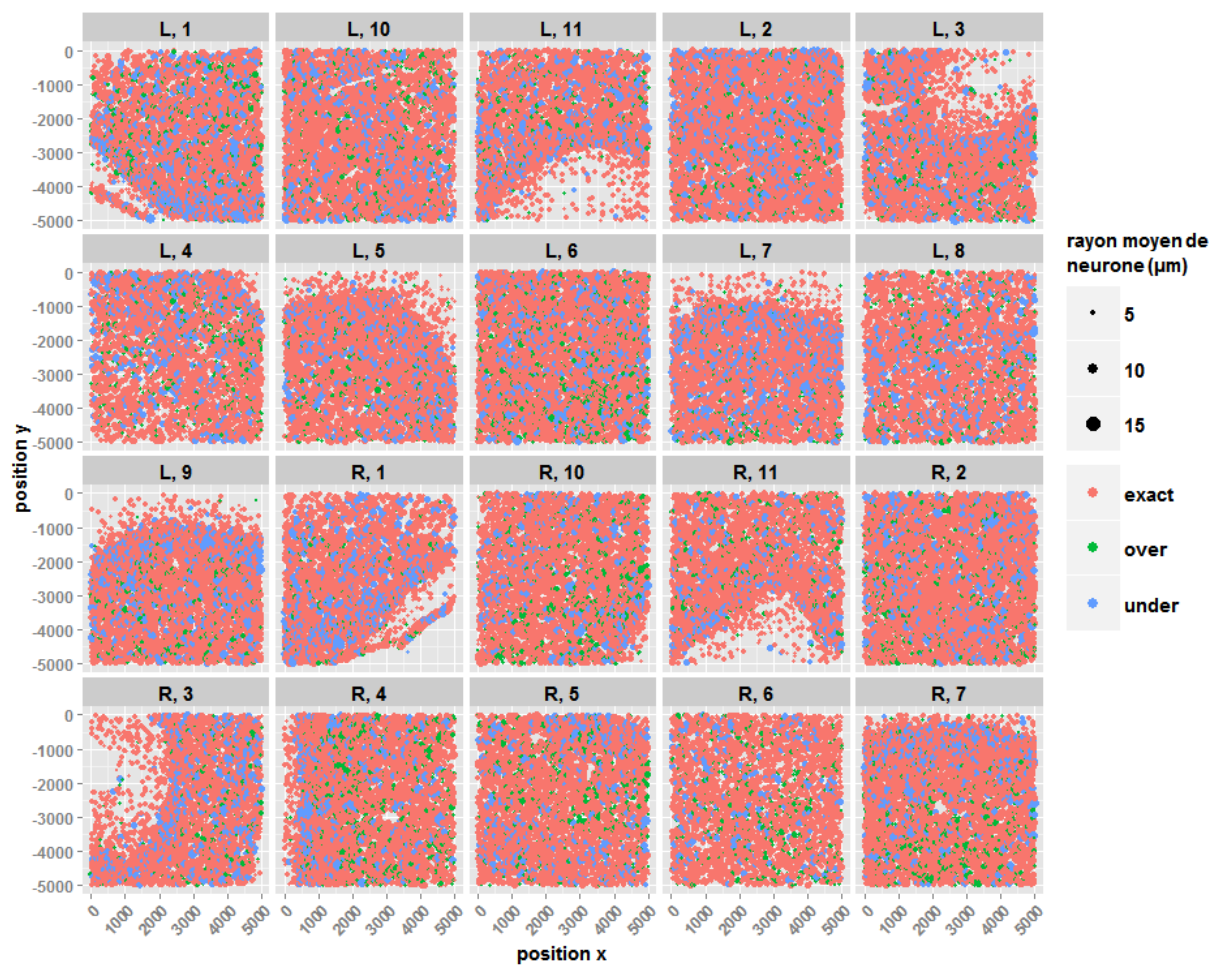


Figure 95 : Région du cortex.

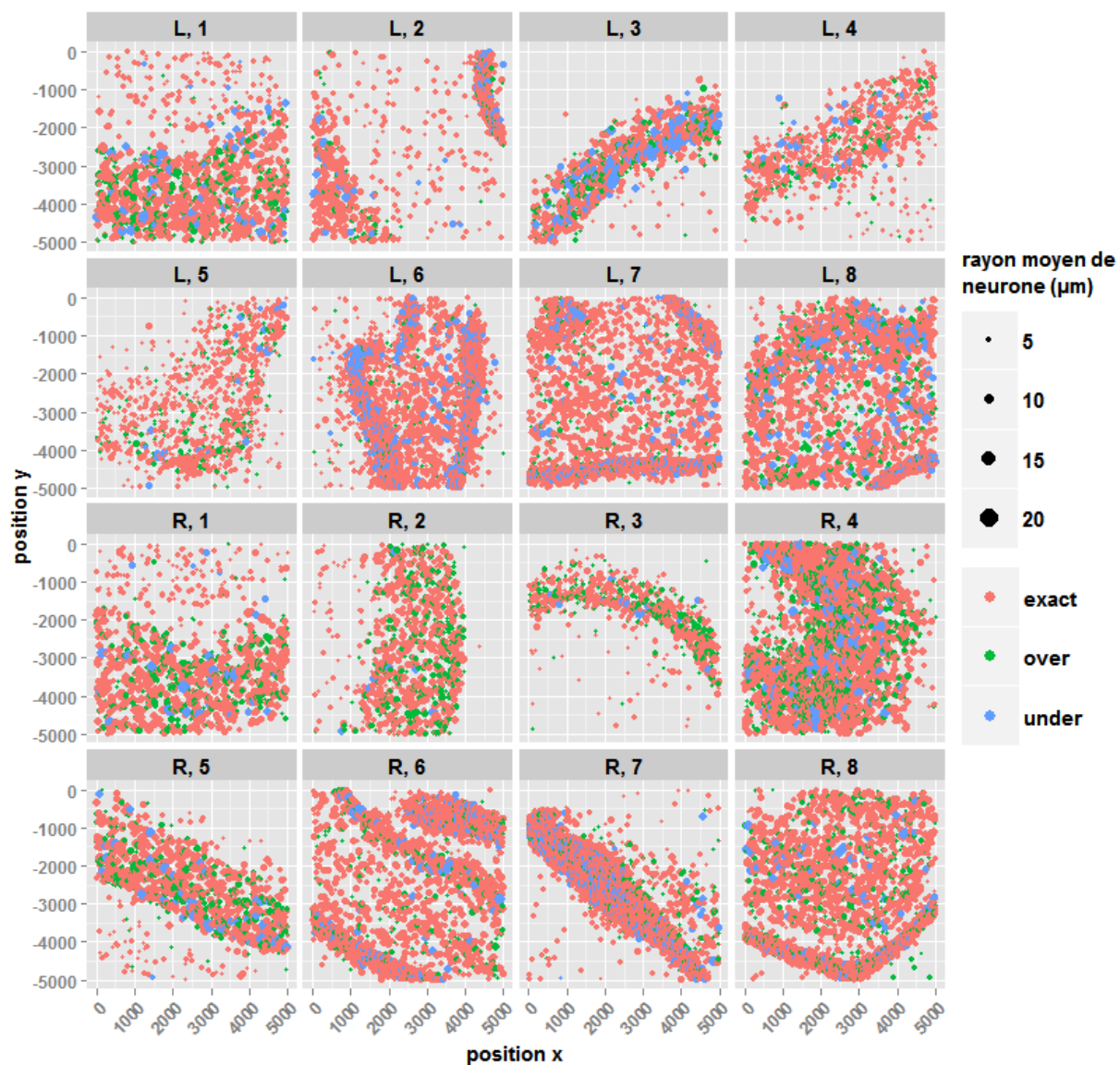


Figure 96 : Région hippocampique.

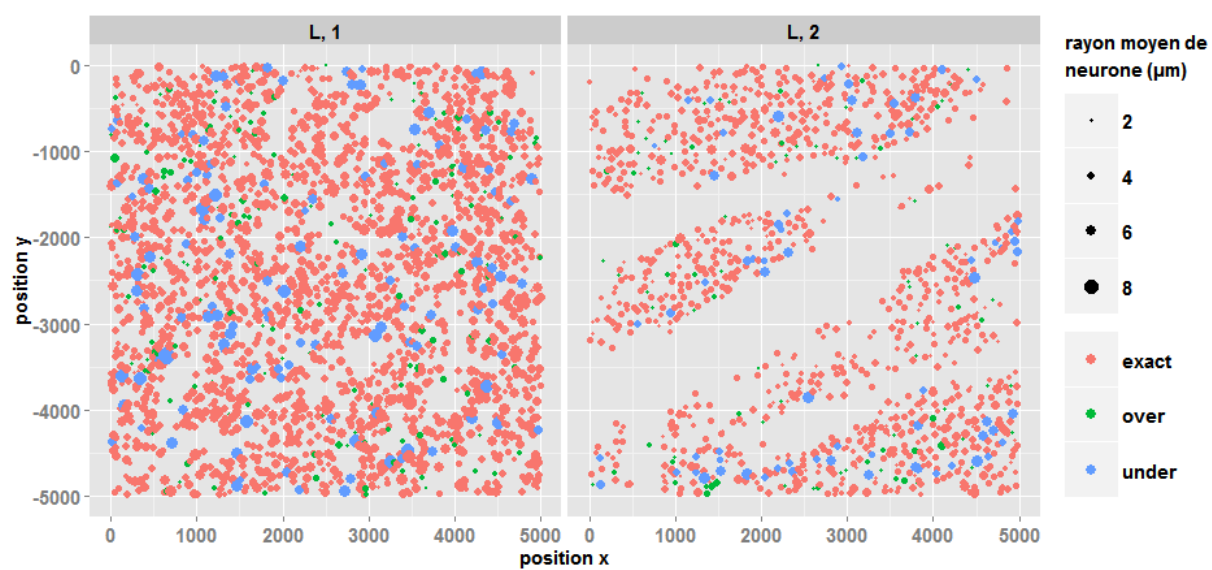


Figure 97 : Région du putamen.

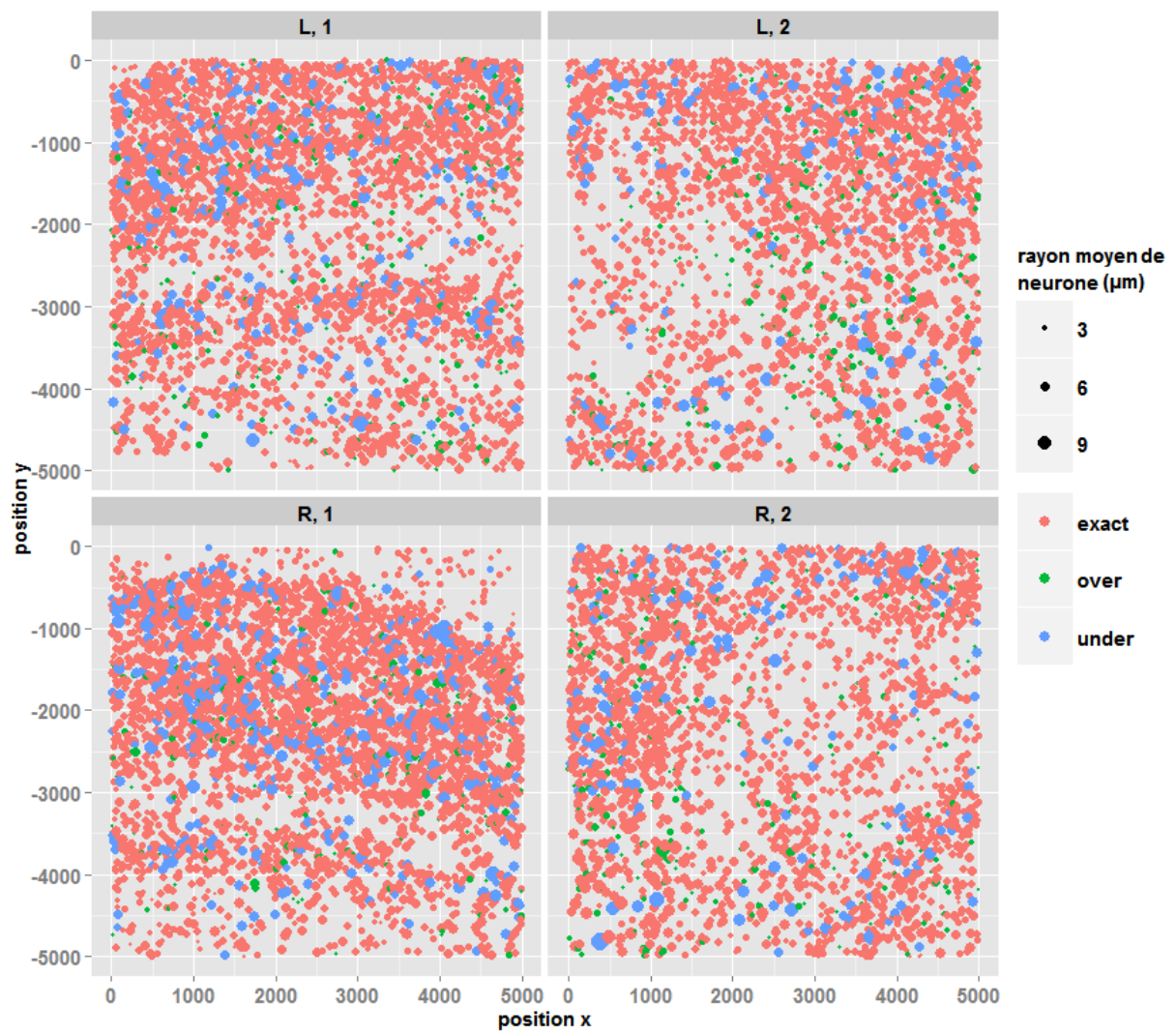


Figure 98 : Région du subiculum.

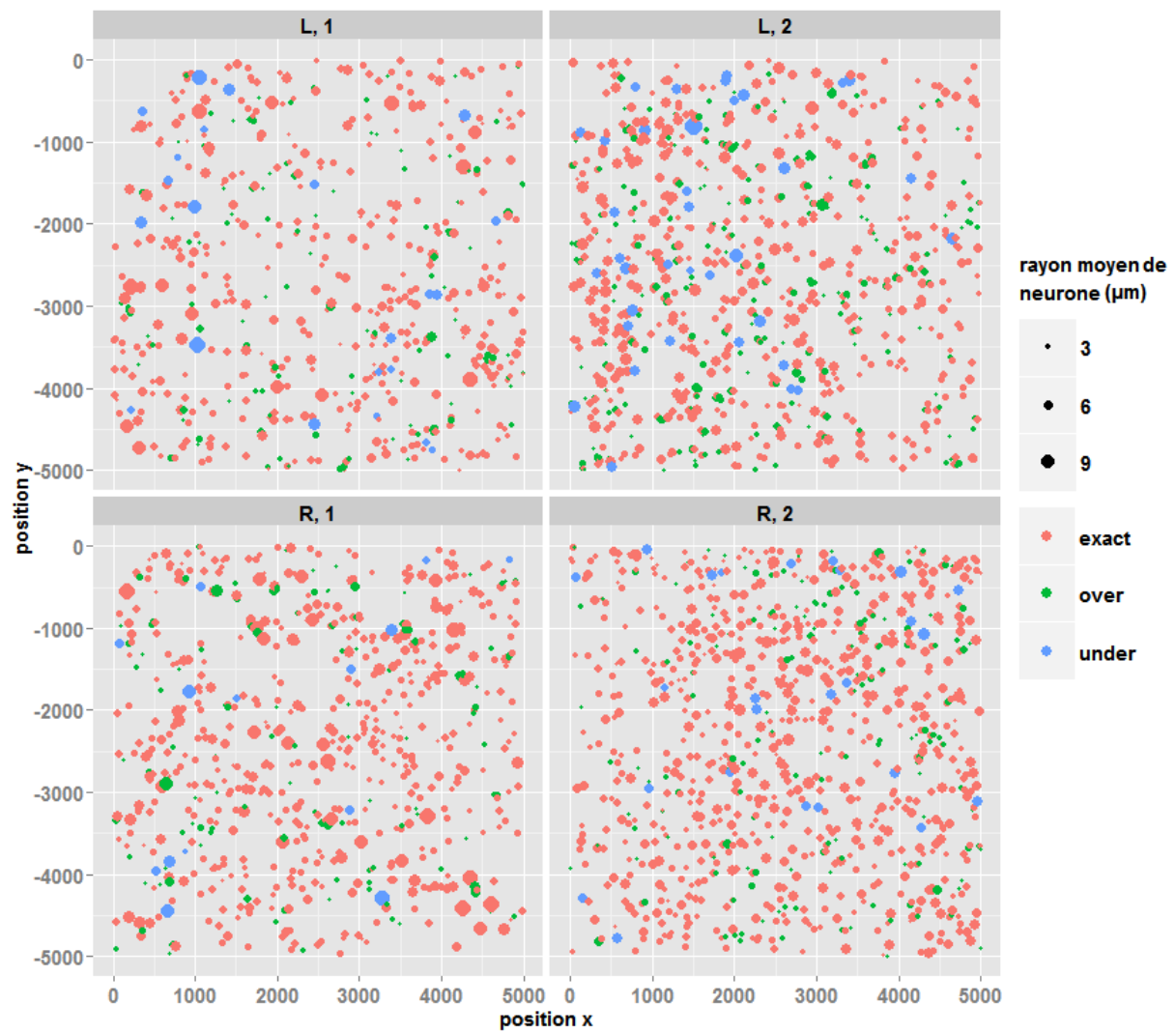


Figure 99 : Région du thalamus.

5. Valeurs moyennes de différentes caractéristiques des populations des neurones par région anatomique

| Régions | Paramètres | Notre méthode | <i>Watershed</i> (σ optimal) |
|-------------|-----------------------------|---------------|--------------------------------------|
| noyau caudé | Distance_centroïde | 4,97±5,71 | 4,98±6,12 |
| | Distance_moyenne_contour | 3,58±3,34 | 3,52±4,15 |
| | Ecart-type_distance_contour | 4,61±4,34 | 4,57±5,19 |
| | Dice | 0,80±0,17 | 0,80±0,18 |
| claustrum | Distance_centroïde | 5,38±5,44 | 6,48±8,05 |
| | Distance_moyenne_contour | 3,42±3,05 | 6,18±6,32 |
| | Ecart-type_distance_contour | 4,45±4,23 | 7,70±7,63 |
| | Dice | 0,80±0,14 | 0,67±0,25 |
| cortex | Distance_centroïde | 6,13±9,01 | 6,47±8,71 |
| | Distance_moyenne_contour | 3,98±2,75 | 4,74±3,79 |
| | Ecart-type_distance_contour | 4,89±3,37 | 5,92±4,67 |
| | Dice | 0,75±0,18 | 0,71±0,22 |
| hippocampe | Distance_centroïde | 5,41±5,85 | 7,06±7,78 |
| | Distance_moyenne_contour | 3,70±2,34 | 5,19±5,26 |
| | Ecart-type_distance_contour | 4,63±2,85 | 6,40±6,04 |
| | Dice | 0,77±0,16 | 0,71±0,23 |
| putamen | Distance_centroïde | 4,77±5,18 | 4,85±4,73 |
| | Distance_moyenne_contour | 2,87±1,89 | 3,50±2,94 |
| | Ecart-type_distance_contour | 3,62±2,36 | 4,31±3,47 |
| | Dice | 0,82±0,14 | 0,79±0,17 |
| subiculum | Distance_centroïde | 4,77±6,72 | 5,40±8,19 |
| | Distance_moyenne_contour | 4,47±2,65 | 4,60±3,55 |
| | Ecart-type_distance_contour | 5,29±3,05 | 5,54±4,20 |
| | Dice | 0,76±0,13 | 0,76±0,15 |
| thalamus | Distance_centroïde | 10,20±12,98 | 11,03±12,82 |
| | Distance_moyenne_contour | 9,19±8,86 | 12,20±11,17 |
| | Ecart-type_distance_contour | 11,84±11,15 | 14,99±13,20 |
| | Dice | 0,70±0,23 | 0,60±0,31 |

6. Descripteurs calculés pour la génération des cartographies quantitatives

| Numéro | Noms des descripteurs | Unité | Commentaire |
|--------|------------------------------------|-----------------|---|
| 1 | label_area_mm2 | mm ² | Surface en mm ² d'un neurone |
| 2 | voronoi_area_mm2 | mm ² | Surface en mm ² de zones de Voronoï associées à son neurone |
| 3 | label_area_px | pixel | Surface en pixel d'un neurone, noté S_n |
| 4 | voronoi_area_px | pixel | Surface en pixel des zones de Voronoï associées à son neurone, noté S_v |
| 5 | label_bounding_xmin | pixel | La position de boîte englobante d'un neurone |
| 6 | label_bounding_xmax | pixel | idem |
| 7 | label_bounding_ymin | pixel | idem |
| 8 | label_bounding_ymax | pixel | idem |
| 9 | voronoi_bounding_xmin | pixel | La position de boîte englobante de zones de Voronoï associées à son neurone |
| 10 | voronoi_bounding_xmax | pixel | idem |
| 11 | voronoi_bounding_ymin | pixel | idem |
| 12 | voronoi_bounding_ymax | pixel | idem |
| 13 | neighbor_count | / | Nombre de neurones auxquels les zones de Voronoï associées sont connectées avec la zone de Voronoï associée au neurone étudié |
| 14 | neighborhood_mean_distance | pixel | La distance moyenne entre le neurone et ses voisins |
| 15 | neighborhood_median_distance | pixel | La distance médiane entre le neurone et ses voisins |
| 16 | neighborhood_label_mean_area_mm2 | mm ² | Surface moyenne en mm ² des voisins d'un neurone |
| 17 | neighborhood_label_mean_area_px | pixel | Surface moyenne en pixel des voisins d'un neurone |
| 18 | neighborhood_voronoi_mean_area_mm2 | mm ² | Surface moyenne en mm ² des zones de Voronoï associées aux voisins d'un neurone |
| 19 | neighborhood_voronoi_mean_area_px | pixel | Surface moyenne en pixel des zones de Voronoï associées aux voisins d'un neurone |
| 20 | label_mass_center_x | pixel | La position X du centre de gravité d'un neurone |
| 21 | label_mass_center_y | pixel | La position Y du centre de gravité d'un neurone |
| 22 | label_perimeter_n4_px | pixel | Le périmètre approximatif calculé à partir du contour de neurone calculé en 4-connexité (N_4), noté $P_4 = N_4/1,273$ |

| | | | |
|----|--------------------------|-------|---|
| 23 | label_perimeter_n8_px | pixel | Le périmètre approximatif calculé à partir du contour de neurone calculé en 8-connexité (N_8), noté $P_8 = N_8/0,9$ |
| 24 | label_perimeter_n4_n8_px | pixel | Le périmètre approximatif calculé avec N_4 et N_8 , noté $P_{48} = (0,414P_4 + 0,586P_8)/1,055$ |
| 25 | label_feret_diameter_min | pixel | Le diamètre de Férét min, c'est la projection minimale de la cellule dans une direction, noté f_{min} |
| 26 | label_feret_diameter_max | pixel | Le diamètre de Férét max, c'est la projection maximale de la cellule dans une direction, noté f_{max} |
| 27 | label_min_radius | pixel | Le rayon minimal d'un neurone |
| 28 | label_max_radius | pixel | Le rayon maximal d'un neurone |
| 29 | label_mean_radius | pixel | Le rayon moyen d'un neurone |
| 30 | label_stddev_radius | / | L'écart-type du rayon d'un neurone autour du rayon moyen |
| 31 | label_radius_variation | / | La somme du carré des différences entre deux rayons consécutifs calculé pour le contour d'un neurone |
| 32 | label_circularity | / | La forme d'un neurone, $4\pi S_n/P_{48}^2$ |
| 33 | label_thinness | / | L'allongement d'un neurone, P_{48}^2/S_n |
| 34 | label_isotropy | / | L'isotropie, f_{min}/f_{max} |
| 35 | label_inertia_vector_x | pixel | Axe d'orientation d'un neurone |
| 36 | label_inertia_vector_y | pixel | idem |
| 37 | label_orientation_angle | degré | L'orientation d'un neurone |
| 38 | hsv_mean_hue | / | La moyenne de la teinte dans l'espace de couleur HSV |
| 39 | hsv_mean_saturation | / | La moyenne de la saturation |
| 40 | hsv_mean_value | / | La moyenne de la valeur |
| 41 | hsv_stand_dev_hue | / | L'écart-type de la teinte |
| 42 | hsv_stand_dev_saturation | / | L'écart-type de la saturation |
| 43 | hsv_stand_dev_value | / | L'écart-type de la valeur |
| 44 | rgb_mean_red | / | La moyenne de la composante rouge dans l'espace de couleur RGB |
| 45 | rgb_mean_green | / | La moyenne de la composante verte |
| 46 | rgb_mean_blue | / | La moyenne de la composante bleue |
| 47 | rgb_stand_dev_red | / | L'écart-type de la composante rouge |
| 48 | rgb_stand_dev_green | / | L'écart-type de la composante verte |
| 49 | rgb_stand_dev_blue | / | L'écart-type de la composante bleue |

Publications

Articles de Journal

- Zhenzhen, Y.*, et al.

Automated cell individualization based on a multi-scale optimization method in cerebral microscopic images.

(En préparation)

Conférences internationales

- Zhenzhen, Y.*, Vandenberghe, M. E., Balbastre, Y., Souedet, N., Hantraye, P., Jan, C., Hérard, A.-S., et Delzescaux, T.

Automated cell individualization and counting in cerebral microscopic images.

IEEE International Conference on Image Processing, Phoenix, USA, Septembre 2016.

(Communication avec acte, poster)

- Zhenzhen, Y.*, Vandenberghe, M. E., Balbastre, Y., Souedet, N., Hérard, A.-S., et Delzescaux, T.

Cellular morphometric analysis : from microscopic scale to whole mouse brains.

Micro-, meso- and macrodynamics (IPSEN foundation), Paris, France, Avril 2015.

(Communication sans acte, poster)

Bibliographie

- Adams, R., Bischof, L., 1994. Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16, 641–647. doi:10.1109/34.295913
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., 2002. *Molecular Biology of the Cell*, 4th ed. Garland Science.
- Al-Kofahi, Y., Lassoued, W., Lee, W., Roysam, B., 2010. Improved Automatic Detection and Segmentation of Cell Nuclei in Histopathology Images. *IEEE Transactions on Biomedical Engineering* 57, 841–852. doi:10.1109/TBME.2009.2035102
- Amit, Y., Geman, D., 1997. Shape Quantization and Recognition with Randomized Trees. *Neural Computation* 9, 1545–1588. doi:10.1162/neco.1997.9.7.1545
- Andersen, P., Morris, R., Amaral, D., Bliss, T., O'Keefe, J., 2016. *The Hippocampus Book*. Oxford University Press.
- Andrey, P., Kiêu, K., Kress, C., Lehmann, G., Tirichine, L., Liu, Z., Biot, E., Adenot, P.-G., Hue-Beauvais, C., Houba-Hérin, N., Duranthon, V., Devinoy, E., Beaujean, N., Gaudin, V., Maurin, Y., Debey, P., 2010. Statistical Analysis of 3D Images Detects Regular Spatial Distributions of Centromeres and Chromocenters in Animal and Plant Nuclei. *PLOS Computational Biology* 6, e1000853. doi:10.1371/journal.pcbi.1000853
- Azevedo, F.A.C., Carvalho, L.R.B., Grinberg, L.T., Farfel, J.M., Ferretti, R.E.L., Leite, R.E.P., Jacob Filho, W., Lent, R., Herculano-Houzel, S., 2009. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J. Comp. Neurol.* 513, 532–541. doi:10.1002/cne.21974
- Bailey, D.L., Townsend, D.W., Valk, P.E., Maisey, M.N., 2005. *Positron Emission Tomography - Basic Sciences* | Dale L. Bailey | Springer.
- Bai, X., Sun, C., Zhou, F., 2008. Touching Cells Splitting by Using Concave Points and Ellipse Fitting, in: *Digital Image Computing: Techniques and Applications (DICTA)*, 2008. Presented at the Digital Image Computing: Techniques and Applications (DICTA), 2008, pp. 271–278. doi:10.1109/DICTA.2008.11
- Balbastre, Y., Rivi re, D., Souedet, N., 2017. Primatologist: a modular segmentation pipeline for Macaque brain morphometry. *NeuroImage*.
- Ballard, C., Gauthier, S., Corbett, A., Brayne, C., Aarsland, D., Jones, E., 2011. Alzheimer's disease. *The Lancet* 377, 1019–1031. doi:10.1016/S0140-6736(10)61349-9
- Barranco Quintana, J.L., Allam, M.F., Del Castillo, A.S., Navajas, R.F.-C., 2009. Parkinson's disease and tea: a quantitative review. *J Am Coll Nutr* 28, 1–6.
- Bazi, Y., Bruzzone, L., Melgani, F., 2007. Image thresholding based on the EM algorithm and the generalized Gaussian distribution. *Pattern Recognition* 40, 619–634. doi:10.1016/j.patcog.2006.05.006
- Bear, M.F., Connors, B.W., Nieoullon, M.A.P.T. et adaptation française A., 2016. *Neurosciences (4e édition) : A la découverte du cerveau*.
- Bichat, X., 1799. *Traité des membranes en général et de diverses membranes en particulier*. Richard.
- Bick, K., 1987. *The early story of Alzheimer's disease*. Raven Press, New York.
- Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A., 1984. *Classification and Regression Trees*. Taylor & Francis.
- Brodmann, K., 1909. *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig : Barth.
- Burns, A., Iliffe, S., 2009. Alzheimer's disease. *BMJ* 338, b158.
- Cai, H., Yang, Z., Cao, X., Xia, W., Xu, X., 2014. A new iterative triclass thresholding technique in image segmentation. *IEEE Trans Image Process* 23, 1038–1046. doi:10.1109/TIP.2014.2298981
- Carlson, N.R., 2012. *Physiology of Behavior*, 11 edition. ed. Pearson, Boston.
- Chareyron, L.J., Banta Lavenex, P., Amaral, D.G., Lavenex, P., 2011. Stereological analysis of the rat and monkey amygdala. *J. Comp. Neurol.* 519, 3218–3239. doi:10.1002/cne.22677

- Cheng, H.D., Jiang, X.H., Sun, Y., Wang, J.L., 2001. Color image segmentation: Advances and prospects. *Pattern Recognition* 34, 2259–2281.
- Chubb, C., Inagaki, Y., Sheu, P., Cummings, B., Wasserman, A., Head, E., Cotman, C., 2006. BioVision: an application for the automated image analysis of histological sections. *Neurobiol. Aging* 27, 1462–1476. doi:10.1016/j.neurobiolaging.2005.08.023
- Coons, A.H., Creech, H.J., Jones, R.N., 1941. Immunological Properties of an Antibody Containing a Fluorescent Group. *Proceedings of the Society for Experimental Biology and Medicine* 47, 200–202. doi:10.3181/00379727-47-13084P
- Cousty, J., Bertrand, G., Najman, L., Couprie, M., 2009. Watershed Cuts: Minimum Spanning Forests and the Drop of Water Principle. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31, 1362–1374. doi:10.1109/TPAMI.2008.173
- Daněš, O., Matula, P., Ortiz-de-Solórzano, C., Muñoz-Barrutia, A., Maška, M., Kozubek, M., 2009. Segmentation of Touching Cell Nuclei Using a Two-Stage Graph Cut Model, in: *Image Analysis, Lecture Notes in Computer Science*. Presented at the Scandinavian Conference on Image Analysis, Springer, Berlin, Heidelberg, pp. 410–419. doi:10.1007/978-3-642-02230-2_42
- Dauguet, J., Condé, F., Hantraye, P., Frouin, V., Delzescaux, T., 2009. Generation of a 3D atlas of the nuclear division of the thalamus based on histological sections of primate: Intra- and intersubject atlas-to-MRI warping. *IRBM, NUMÉRO SPÉCIAL TECHNOLOGIES POUR L'AUTONOMIE* 30, 281–291. doi:10.1016/j.irbm.2009.10.004
- Davies, M., 2002. The Neuron: size comparison. *Neuroscience: A journey through the brain*, retrieved 2009-06-20.
- Dayalu, P., Albin, R.L., 2015. Huntington disease: pathogenesis and treatment. *Neurol Clin* 33, 101–114. doi:10.1016/j.ncl.2014.09.003
- Dodson, P.D., Larvin, J.T., Duffell, J.M., Garas, F.N., Doig, N.M., Kessar, N., Duguid, I.C., Bogacz, R., Butt, S.J.B., Magill, P.J., 2015. Distinct developmental origins manifest in the specialized encoding of movement by adult neurons of the external globus pallidus. *Neuron* 86, 501–513. doi:10.1016/j.neuron.2015.03.007
- Doetsch, F., 2003. The glial identity of neural stem cells. *Nat. Neurosci.* 6, 1127–1134. doi:10.1038/nn1144
- Doke, S.K., Dhawale, S.C., 2015. Alternatives to animal testing: A review. *Saudi Pharm J* 23, 223–229. doi:10.1016/j.jsps.2013.11.002
- Dong, X., Wang, Y., Qin, Z., 2009. Molecular mechanisms of excitotoxicity and their relevance to pathogenesis of neurodegenerative diseases. *Acta Pharmacol. Sin.* 30, 379–387. doi:10.1038/aps.2009.24
- Dubois, A., Hérard, A.-S., Delatour, B., Hantraye, P., Bonvento, G., Dhenain, M., Delzescaux, T., 2010. 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, 586–598. doi:10.1016/j.neuroimage.2010.02.074
- Fan, S.-K.S., Lin, Y., 2007. A multi-level thresholding approach using a hybrid optimal estimation algorithm. *Pattern Recognition Letters* 28, 662–669. doi:10.1016/j.patrec.2006.11.005
- Fine, E.J., Ionita, C.C., Lohr, L., 2002. The history of the development of the cerebellar examination. *Semin Neurol* 22, 375–384. doi:10.1055/s-2002-36759
- Flood, D.G., Coleman, P.D., 1988. Neuron numbers and sizes in aging brain: comparisons of human, monkey, and rodent data. *Neurobiol. Aging* 9, 453–463.
- Frank, S., 2014. Treatment of Huntington's Disease. *Neurotherapeutics* 11, 153–160. doi:10.1007/s13311-013-0244-z
- GBD 2015 Disease and Injury Incidence and Prevalence Collaborators, 2016. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 388, 1545–1602. doi:10.1016/S0140-6736(16)31678-6

- GBD 2015 Mortality and Causes of Death, Collaborators, 2016. Global, Regional, and National Life Expectancy, All-Cause Mortality, and Cause-Specific Mortality for 249 Causes of Death, 1980–2015: A Systematic Analysis for the Global Burden of Disease Study 2015. *Lancet* (London, England) 388, 1459–1544.
- Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. *Mach Learn* 63, 3–42. doi:10.1007/s10994-006-6226-1
- Gittins, R., Harrison, P.J., 2004. Neuronal density, size and shape in the human anterior cingulate cortex: a comparison of Nissl and NeuN staining. *Brain Research Bulletin* 63, 155–160. doi:10.1016/j.brainresbull.2004.02.005
- Gray, H., 1918. *Anatomy of the Human Body*. Lea & Febiger.
- Guadalupe, T., Mathias, S.R., vanErp, T.G.M., Whelan, C.D., Zwiers, M.P., Abe, Y., Abramovic, L., Agartz, I., Andreassen, O.A., Arias-Vásquez, A., Aribisala, B.S., Armstrong, N.J., Arolt, V., Artiges, E., Ayesa-Arriola, R., Baboyan, V.G., Banaschewski, T., Barker, G., Bastin, M.E., Baune, B.T., Blangero, J., Bokde, A.L.W., Boedhoe, P.S.W., Bose, A., Brem, S., Brodaty, H., Bromberg, U., Brooks, S., Büchel, C., Buitelaar, J., Calhoun, V.D., Cannon, D.M., Cattrell, A., Cheng, Y., Conrod, P.J., Conzelmann, A., Corvin, A., Crespo-Facorro, B., Crivello, F., Dannlowski, U., de Zubicaray, G.I., de Zwarte, S.M.C., Deary, I.J., Desrivieres, S., Doan, N.T., Donohoe, G., Dørum, E.S., Ehrlich, S., Espeseth, T., Fernández, G., Flor, H., Fouché, J.-P., Frouin, V., Fukunaga, M., Gallinat, J., Garavan, H., Gill, M., Suarez, A.G., Gowland, P., Grabe, H.J., Grotegerd, D., Gruber, O., Hagenaars, S., Hashimoto, R., Hauser, T.U., Heinz, A., Hibar, D.P., Hoekstra, P.J., Hoogman, M., Howells, F.M., Hu, H., Hulshoff Pol, H.E., Huyser, C., Ittermann, B., Jahanshad, N., Jönsson, E.G., Jurk, S., Kahn, R.S., Kelly, S., Kraemer, B., Kugel, H., Kwon, J.S., Lemaitre, H., Lesch, K.-P., Lochner, C., Luciano, M., Marquand, A.F., Martin, N.G., Martínez-Zalacáin, I., Martinot, J.-L., Mataix-Cols, D., Mather, K., McDonald, C., McMahon, K.L., Medland, S.E., Menchón, J.M., Morris, D.W., Mothersill, O., Maniega, S.M., Mwangi, B., Nakamae, T., Nakao, T., Narayanaswamy, J.C., Nees, F., Nordvik, J.E., Onnink, A.M.H., Opel, N., Ophoff, R., Paillère Martinot, M.-L., Papadopoulos Orfanos, D., Pauli, P., Paus, T., Poustka, L., Reddy, J.Y., Renteria, M.E., Roiz-Santiañez, R., Roos, A., Royle, N.A., Sachdev, P., Sánchez-Juan, P., Schmaal, L., Schumann, G., Shumskaya, E., Smolka, M.N., Soares, J.C., Soriano-Mas, C., Stein, D.J., Strike, L.T., Toro, R., Turner, J.A., Tzourio-Mazoyer, N., Uhlmann, A., Hernández, M.V., van den Heuvel, O.A., van der Meer, D., van Haren, N.E.M., Veltman, D.J., Venkatasubramanian, G., Vetter, N.C., Vuletic, D., Walitza, S., Walter, H., Walton, E., Wang, Z., Wardlaw, J., Wen, W., Westlye, L.T., Whelan, R., Wittfeld, K., Wolfers, T., Wright, M.J., Xu, J., Xu, X., Yun, J.-Y., Zhao, J., Franke, B., Thompson, P.M., Glahn, D.C., Mazoyer, B., Fisher, S.E., Francks, C., 2016. Human subcortical brain asymmetries in 15,847 people worldwide reveal effects of age and sex. *Brain Imaging Behav.* doi:10.1007/s11682-016-9629-z
- Gundersen, H.J.G., 1986. Stereology of Arbitrarily Shaped Particles: Unbiased Estimation of Number and Sizes, in: *Proceedings of the First International Symposium for Science on Form. Presented at the Proceedings of the First International Symposium for Science on Form.*
- Hanbury, A., 2008. Constructing cylindrical coordinate colour spaces. *Pattern Recognition Letters* 29, 494–500. doi:10.1016/j.patrec.2007.11.002
- Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28, 100–108. doi:10.2307/2346830
- Herculano-Houzel, S., 2009. The Human Brain in Numbers: A Linearly Scaled-up Primate Brain. *Front Hum Neurosci* 3. doi:10.3389/neuro.09.031.2009
- He, Y., Gong, H., Xiong, B., Xu, X., Li, A., Jiang, T., Sun, Q., Wang, S., Luo, Q., Chen, S., 2015. iCut: an Integrative Cut Algorithm Enables Accurate Segmentation of Touching Cells. *Sci Rep* 5. doi:10.1038/srep12089
- Holtzman, D.M., Morris, J.C., Goate, A.M., 2011. Alzheimer's disease: the challenge of the second century. *Sci Transl Med* 3, 77sr1. doi:10.1126/scitranslmed.3002369

- Hooke, R., 1667. *Micrographia: Or, Some Physiological Descriptions of Minute Bodies Made by Magnifying Glasses, with Observations and Inquiries Thereupon*. Science Heritage.
- Ishizuka, N., Cowan, W.M., Amaral, D.G., 1995. A quantitative analysis of the dendritic organization of pyramidal cells in the rat hippocampus. *J. Comp. Neurol.* 362, 17–45. doi:10.1002/cne.903620103
- Jelsing, J., Nielsen, R., Olsen, A.K., Grand, N., Hemmingsen, R., Pakkenberg, B., 2006. The postnatal development of neocortical neurons and glial cells in the Göttingen minipig and the domestic pig brain. *J. Exp. Biol.* 209, 1454–1462. doi:10.1242/jeb.02141
- Joliot, M., Tzourio-Mazoyer, N., Mazoyer, B., 2016. Intra-hemispheric intrinsic connectivity asymmetry and its relationships with handedness and language Lateralization. *Neuropsychologia* 93, 437–447. doi:10.1016/j.neuropsychologia.2016.03.013
- Kainz, P., Urschler, M., Schuster, S., Wohlfahrt, P., Lepetit, V., 2015. You Should Use Regression to Detect Cells, in: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Lecture Notes in Computer Science*. Presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Cham, pp. 276–283. doi:10.1007/978-3-319-24574-4_33
- Kalia, L.V., Lang, A.E., 2015. Parkinson's disease. *Lancet* 386, 896–912. doi:10.1016/S0140-6736(14)61393-3
- Karlsen, A.S., Pakkenberg, B., 2011. Total numbers of neurons and glial cells in cortex and basal ganglia of aged brains with Down syndrome--a stereological study. *Cereb. Cortex* 21, 2519–2524. doi:10.1093/cercor/bhr033
- Kass, M., Witkin, A., Terzopoulos, D., 1988. Snakes: Active Contour Models. *International Journal of Computer Vision* 321–331.
- Kolb, H., Fernandez, E., Nelson, R. (Eds.), 1995. *Webvision: The Organization of the Retina and Visual System*. University of Utah Health Sciences Center, Salt Lake City (UT).
- Kothari, S., Chaudry, Q., Wang, M.D., 2009. Automated cell counting and cluster segmentation using concavity detection and ellipse fitting techniques, in: *IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009. ISBI '09*. Presented at the IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009. ISBI '09, pp. 795–798. doi:10.1109/ISBI.2009.5193169
- Larsen, C.C., Bonde Larsen, K., Bogdanovic, N., Laursen, H., Graem, N., Samuelsen, G.B., Pakkenberg, B., 2006. Total number of cells in the human newborn telencephalic wall. *Neuroscience* 139, 999–1003. doi:10.1016/j.neuroscience.2006.01.005
- Li, G., Liu, T., Nie, J., Guo, L., Chen, J., Zhu, J., Xia, W., Mara, A., Holley, S., Wong, S.T.C., 2008. Segmentation of touching cell nuclei using gradient flow tracking. *J Microsc* 231, 47–58. doi:10.1111/j.1365-2818.2008.02016.x
- Loménie, N., Racocanu, D., 2012. Point set morphological filtering and semantic spatial configuration modeling: Application to microscopic image and bio-structure analysis. *Pattern Recognition* 45, 2894–2911. doi:10.1016/j.patcog.2012.01.021
- Lou, X., Koethe, U., Wittbrodt, J., Hamprecht, F.A., 2012. Learning to segment dense cell nuclei with shape prior, in: *2012 IEEE Conference on Computer Vision and Pattern Recognition*. Presented at the 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1012–1018. doi:10.1109/CVPR.2012.6247778
- Maingault, S., Tzourio-Mazoyer, N., Mazoyer, B., Crivello, F., 2016. Regional correlations between cortical thickness and surface area asymmetries: A surface-based morphometry study of 250 adults. *Neuropsychologia* 93, 350–364. doi:10.1016/j.neuropsychologia.2016.03.025
- Mao, X., Ou, M.T., Karuppagounder, S.S., Kam, T.-I., Yin, X., Xiong, Y., Ge, P., Umanah, G.E., Brahmachari, S., Shin, J.-H., Kang, H.C., Zhang, J., Xu, J., Chen, R., Park, H., Andrabi, S.A., Kang, S.U., Gonçalves, R.A., Liang, Y., Zhang, S., Qi, C., Lam, S., Keiler, J.A., Tyson, J., Kim, D., Panicker, N., Yun, S.P., Workman, C.J., Vignali, D.A.A., Dawson, V.L., Ko, H.S., Dawson, T.M.,

2016. Pathological α -synuclein transmission initiated by binding lymphocyte-activation gene 3. *Science* 353. doi:10.1126/science.aah3374
- Mayhew, T.M., 1992. A review of recent advances in stereology for quantifying neural structure. *J. Neurocytol.* 21, 313–328.
- Molnar, C., Jermyn, I.H., Kato, Z., Rahkama, V., Östling, P., Mikkonen, P., Pietiäinen, V., Horvath, P., 2016. Accurate Morphology Preserving Segmentation of Overlapping Cells based on Active Contours. *Sci Rep* 6, 32412. doi:10.1038/srep32412
- Mosconi, L., 2005. Brain glucose metabolism in the early and specific diagnosis of Alzheimer's disease. FDG-PET studies in MCI and AD. *Eur. J. Nucl. Med. Mol. Imaging* 32, 486–510. doi:10.1007/s00259-005-1762-7
- Mullen, R.J., Buck, C.R., Smith, A.M., 1992. NeuN, a neuronal specific nuclear protein in vertebrates. *Development* 116, 201–211.
- Nedzved, A., Ablameyko, S., Pitas, I., 2000. Morphological segmentation of histology cell images, in: 15th International Conference on Pattern Recognition. pp. 500–503. doi:10.1109/ICPR.2000.905385
- Nestler, E.J., Hyman, S.E., Malenka, R.C., 2008. *Molecular Neuropharmacology: A Foundation for Clinical Neuroscience*, Second Edition. McGraw Hill Professional.
- Nishiyama, A., Yang, Z., Butt, A., 2005. Astrocytes and NG2-glia: what's in a name? *J. Anat.* 207, 687–693. doi:10.1111/j.1469-7580.2005.00489.x
- Noctor, S.C., Martínez-Cerdeño, V., Kriegstein, A.R., 2007. Contribution of intermediate progenitor cells to cortical histogenesis. *Arch. Neurol.* 64, 639–642. doi:10.1001/archneur.64.5.639
- Ojala, T., Pietikäinen, M., Harwood, D., 1996. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition* 29, 51–59. doi:10.1016/0031-3203(95)00067-4
- Ojala, T., Pietikäinen, M., Mäenpää, T., 2002. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 971–987. doi:10.1109/TPAMI.2002.1017623
- Onos, K.D., Sukoff Rizzo, S.J., Howell, G.R., Sasner, M., 2016. Toward more predictive genetic mouse models of Alzheimer's disease. *Brain Res. Bull.* 122, 1–11. doi:10.1016/j.brainresbull.2015.12.003
- Otsu, N., 1979. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9, 62–66. doi:10.1109/TSMC.1979.4310076
- Pakkenberg, B., Gundersen, H.J., 1997. Neocortical neuron number in humans: effect of sex and age. *J. Comp. Neurol.* 384, 312–320.
- Parkinson's Disease Information Page [WWW Document], 2016. URL <https://www.ninds.nih.gov/Disorders/All-Disorders/Parkinsons-Disease-Information-Page> (accessed 4.24.17).
- Passingham, R., 2009. How good is the macaque monkey model of the human brain? *Curr Opin Neurobiol* 19, 6–11. doi:10.1016/j.conb.2009.01.002
- Pelvig, D.P., Pakkenberg, H., Stark, A.K., Pakkenberg, B., 2008. Neocortical glial cell numbers in human brains. *Neurobiol. Aging* 29, 1754–1762. doi:10.1016/j.neurobiolaging.2007.04.013
- Perona, P., Malik, J., 1990. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 629–639.
- Petiet, A., Delatour, B., Dhenain, M., 2011. Models of neurodegenerative disease – Alzheimer's Anatomical and amyloid plaque imaging. *Methods Mol Biol* 771. doi:10.1007/978-1-61779-219-9_16
- Platt, T.L., Reeves, V.L., Murphy, M.P., 2013. Transgenic Models of Alzheimer's Disease: Better Utilization of Existing Models through Viral Transgenesis. *Biochim Biophys Acta* 1832, 1437–1448. doi:10.1016/j.bbadis.2013.04.017
- Poulain, E., Prigent, S., Soubies, E., Descombes, X., 2015. Cells detection using segmentation competition, in: 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI).

- Presented at the 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), pp. 1208–1211. doi:10.1109/ISBI.2015.7164090
- Pyapali, G.K., Sik, A., Penttonen, M., Buzsaki, G., Turner, D.A., 1998. Dendritic properties of hippocampal CA1 pyramidal neurons in the rat: intracellular staining in vivo and in vitro. *J. Comp. Neurol.* 391, 335–352.
- Qi, J., 2014. Dense nuclei segmentation based on graph cut and convexity-concavity analysis. *J Microsc* 253, 42–53. doi:10.1111/jmi.12096
- Ranson, S.W., 1920. The anatomy of the nervous system, from the standpoint of development and function. Philadelphia : W. B. Saunders.
- Schmitz, C., Eastwood, B.S., Tappan, S.J., Glaser, J.R., Peterson, D.A., Hof, P.R., 2014. Current automated 3D cell detection methods are not a suitable replacement for manual stereologic cell counting. *Front Neuroanat* 8, 27. doi:10.3389/fnana.2014.00027
- Serra, J., 1982. Image Analysis and Mathematical Morphology. Presented at the ICASSP88 International Conference on Acoustics Speech and Signal Processing. doi:10.2307/2531038
- Shu, J., Fu, H., Qiu, G., Kaye, P., Ilyas, M., 2013. Segmenting overlapping cell nuclei in digital histopathology images. *Conf Proc IEEE Eng Med Biol Soc* 2013, 5445–5448. doi:10.1109/EMBC.2013.6610781
- Sørensen, T., 1948. A Method of Establishing Groups of Equal Amplitude in Plant Sociology Based on Similarity of Species Content and Ist Application to Analyses of the Vegetation on Danish Commons. Munksgaard.
- Spalding, K.L., Bergmann, O., Alkass, K., Bernard, S., Salehpour, M., Huttner, H.B., Boström, E., Westerlund, I., Vial, C., Buchholz, B.A., Possnert, G., Mash, D.C., Druid, H., Frisén, J., 2013. Dynamics of hippocampal neurogenesis in adult humans. *Cell* 153, 1219–1227. doi:10.1016/j.cell.2013.05.002
- Stucht, D., Danishad, K.A., Schulze, P., Godenschweger, F., Zaitsev, M., Speck, O., 2015. Highest Resolution In Vivo Human Brain MRI Using Prospective Motion Correction. *PLOS ONE* 10, e0133921. doi:10.1371/journal.pone.0133921
- Ullian, E.M., Sapperstein, S.K., Christopherson, K.S., Barres, B.A., 2001. Control of synapse number by glia. *Science* 291, 657–661. doi:10.1126/science.291.5504.657
- Vandenberghe, M., 2015. 3D whole-brain quantitative histopathology: methodology and applications in mouse models of Alzheimer’s disease.
- Vandenberghe, M.E., Balbastre, Y., Souedet, N., Herard, A.-S., Dhenain, M., Frouin, F., Delzescaux, T., 2015. Robust supervised segmentation of neuropathology whole-slide microscopy images, in: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Presented at the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3851–3854. doi:10.1109/EMBC.2015.7319234
- Vandenberghe, M.E., Hérard, A.-S., Souedet, N., Sadouni, E., Santin, M.D., Briet, D., Carré, D., Schulz, J., Hantraye, P., Chabrier, P.-E., Rooney, T., Debeir, T., Blanchard, V., Pradier, L., Dhenain, M., Delzescaux, T., 2016. High-throughput 3D whole-brain quantitative histopathology in rodents. *Sci Rep* 6. doi:10.1038/srep20958
- Vincent, L., Soille, P., 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13, 583–598. doi:10.1109/34.87344
- Vogt, B.A., Nimchinsky, E.A., Vogt, L.J., Hof, P.R., 1995. Human cingulate cortex: Surface features, flat maps, and cytoarchitecture. *J. Comp. Neurol.* 359, 490–506. doi:10.1002/cne.903590310
- Walløe, S., Pakkenberg, B., Fabricius, K., 2014. Stereological estimation of total cell numbers in the human cerebral and cerebellar cortex. *Front Hum Neurosci* 8, 508. doi:10.3389/fnhum.2014.00508
- Wang, L., He, D.-C., 1990. Texture classification using texture spectrum. *Pattern Recognition* 23, 905–910. doi:10.1016/0031-3203(90)90135-8

- Wang, S., Chung, F., Xiong, F., 2008. A novel image thresholding method based on Parzen window estimate. *Pattern Recognition* 41, 117–129. doi:10.1016/j.patcog.2007.03.029
- Warby, S.C., Graham, R.K., Hayden, M.R., 1993. Huntington Disease, in: Pagon, R.A., Adam, M.P., Ardinger, H.H., Wallace, S.E., Amemiya, A., Bean, L.J., Bird, T.D., Ledbetter, N., Mefford, H.C., Smith, R.J., Stephens, K. (Eds.), *GeneReviews*(®). University of Washington, Seattle, Seattle (WA).
- West, M.J., Slomianka, L., Gundersen, H.J., 1991. Unbiased stereological estimation of the total number of neurons in the subdivisions of the rat hippocampus using the optical fractionator. *Anat. Rec.* 231, 482–497. doi:10.1002/ar.1092310411
- Williams, R.W., Herrup, K., 1988. The control of neuron number. *Annu. Rev. Neurosci.* 11, 423–453. doi:10.1146/annurev.ne.11.030188.002231
- Xu, C., Prince, J.L., 1997. Gradient vector flow: a new external force for snakes, in: , 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997. Proceedings. Presented at the , 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997. Proceedings, pp. 66–71. doi:10.1109/CVPR.1997.609299
- Yang, X., Li, H., Zhou, X., 2006. Nuclei Segmentation Using Marker-Controlled Watershed, Tracking Using Mean-Shift, and Kalman Filter in Time-Lapse Microscopy. *IEEE Transactions on Circuits and Systems I: Regular Papers* 53, 2405–2414. doi:10.1109/TCSI.2006.884469
- Zhang, C., Sun, C., Pham, T.D., 2013. Segmentation of clustered nuclei based on concave curve expansion. *J Microsc* 251, 57–67. doi:10.1111/jmi.12043
- Zucker, S.W., 1976. Region Growing: childhood and adolescence. *Computer Graphics and Image Processing* 5, 382–399.

Résumé

La compréhension des mécanismes impliqués dans les cas sains, les maladies neurodégénératives ainsi que le développement de nouvelles approches thérapeutiques repose sur l'utilisation de modèles expérimentaux pertinents et de techniques d'imagerie adaptées. Dans ce contexte, la microscopie virtuelle offre la possibilité unique d'analyser ces modèles à l'échelle cellulaire avec une très grande variété de marquages histologiques.

Mon projet de thèse consiste à mettre en place et à appliquer une méthode d'analyse d'images histologiques en couleur permettant de segmenter et de synthétiser l'information relative aux neurones à l'aide de l'anticorps NeuN sur des coupes de cerveau de macaque. Dans ce travail, nous appliquons d'abord la méthode de *Random Forest* (RF) pour segmenter les neurones ainsi que le tissu et le fond. Ensuite, nous proposons une méthode originale pour séparer les neurones qui sont accolés afin de les individualiser. Cette méthode s'adapte à l'ensemble des neurones présentant une taille variable (diamètre variant entre 5 et 30 μm). Elle est également efficace non seulement pour des régions dites « simples » caractérisées par une faible densité de neurones mais aussi pour des régions dites « complexes » caractérisées par une très forte densité de plusieurs milliers de neurones. Le travail suivant se concentre sur la création de cartes paramétriques synthétisant la morphologie et la distribution des neurones individualisés. Pour cela, un changement d'échelle est mis en œuvre afin de produire des cartographies présentant des résolutions spatiales plus faibles (résolution originale de 0,22 μm et cartographies créées proposant une résolution spatiale adaptative de quelques dizaines à quelques centaines de micromètres). Plusieurs dizaines de paramètres morphologiques (rayon moyen, surface, orientation, etc.) sont d'abord calculés pour chaque neurone ainsi que des paramètres colorimétriques. Ensuite, il est possible de synthétiser ces informations sous la forme de cartes paramétriques à plus basse résolution à l'échelle de régions anatomiques, de coupes voire, à terme, de cerveaux entiers. Cette étape transforme des images microscopiques qualitatives couleur en images mésoscopiques quantitatives, plus informatives et plus simples à analyser. Ce travail permet d'analyser statistiquement de très grands volumes de données, de synthétiser l'information sous la forme de cartographies quantitatives, d'analyser des problèmes extrêmement complexes tels que la mort neuronale et à terme de tester de nouveaux médicaments voire de confronter ces informations acquises *post mortem* avec des données acquises *in vivo*.

Abstract

Understanding the mechanisms involved in healthy cases, neurodegenerative diseases and the development of new therapeutic approaches is based on the use of relevant experimental models and appropriate imaging techniques. In this context, virtual microscopy offers the unique possibility of analyzing these models at a cellular scale with a very wide variety of histological markers.

My thesis project consists in carrying out and applying a method of analyzing colored histological images that can segment and synthesize information corresponding to neurons using the NeuN antibody on sections of the macaque brain. In this work, we first apply the Random Forest (RF) method to segment neurons as well as tissue and background. Then, we propose an original method to separate the touching or overlapping neurons in order to individualize them. This method is adapted to process neurons presenting a variable size (diameter varying between 5 and 30 μm). It is also effective not only for so-called "simple" regions characterized by a low density of neurons but also for so-called "complex" regions characterized by a very high density of several thousands of neurons. The next work focuses on the creation of parametric maps synthesizing the morphology and distribution of individualized neurons. For this purpose, a multiscale approach is implemented in order to produce maps with lower spatial resolutions (0.22 μm original resolution and created maps offering adaptive spatial resolution from a few dozens to a few hundred of micrometers). Several dozens of morphological parameters (mean radius, surface, orientation, etc.) are first computed as well as colorimetric parameters. Then, it is possible to synthesize this information in the form of lower-resolution parametric maps at the level of anatomical regions, sections and even, eventually, the entire brains. This step transforms qualitative color microscopic images to quantitative mesoscopic images, more informative and easier to analyze. This work makes it possible to statistically analyze very large volumes of data, to synthesize information in the form of quantitative maps, to analyze extremely complex problems such as neuronal death, to test new drugs and to compare this acquired information *post mortem* with data acquired *in vivo*.