POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Modélisation topologique du réseau vasculaire rétinien à partir d'image de fond d'œil

GABRIEL LEPETIT-AIMON
Département de génie informatique et génie logiciel

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*
Génie informatique

Juin 2025

© Gabriel Lepetit-Aimon, 2025.

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Cette thèse intitulée :

Modélisation topologique du réseau vasculaire rétinien à partir d'image de fond d'œil

présentée par **Gabriel LEPETIT-AIMON**
en vue de l'obtention du diplôme de *Philosophiæ Doctor*
a été dûment acceptée par le jury d'examen constitué de :

Thomas HURTUT, président
Farida CHERIET, membre et directrice de recherche
Julien COHEN-ADAD, membre
Christian DESROSIERS, membre externe

DÉDICACE

À ma famille,
de part et d'autre de l'océan.

REMERCIEMENTS

Un doctorat, c'est un travail académique exigeant et obsédant. Ça, je le savais en commençant. Mais les sept dernières années que je lui ai consacrées m'ont enseigné qu'un doctorat, c'est surtout une épreuve d'endurance passée à déconstruire des algorithmes et des pratiques de recherche autant qu'à se déconstruire soi-même. C'est un long voyage embrumé de doutes, face auxquels on se sent parfois bien solitaire. Et pour traverser cette brume, les rencontres et les compagnons de voyage, nouveaux comme anciens, sont des lueurs précieuses. Je voudrais ici leur adresser une pensée...

D'abord, à Farida pour m'avoir accueilli dans la grande famille de son laboratoire, pour m'avoir fait confiance, m'avoir toujours soutenu en s'assurant que je ne manquai de rien, et pour m'avoir offert sans contrepartie la chose dont j'avais, au fond, le plus besoin : du temps ;

À Philippe pour son assistance dans les communications avec les différents services de Polytechnique, et pour sa disponibilité et sa bienveillance à corriger patiemment toutes mes productions en anglais ;

À Fantin pour m'avoir invité dans ce laboratoire et pour les discussions passionnées sur la théorie de l'apprentissage machine et l'interprétation qu'il faut en faire ;

À Lama pour sa détermination et sa carrière inspirante, mais aussi pour ses sourires et ses attentions apaisantes, et pour m'avoir offert ma première expérience de cours magistral ;

Aux membres du laboratoire avec qui j'ai eu la chance de collaborer : Lycia, Joe et Raphaël, et à tous ceux que j'ai côtoyés et qui ont égayé les repas du midi ;

Ensuite, à Clément pour m'avoir servi de tuteur sur lequel je me suis reposé à mon arrivée à Montréal, puis de complice lorsqu'il a fallu échafauder nos projets académiques, pour notre quête commune de sens, notamment par le développement de logiciels libres, et pour notre profond respect mutuel ;

À Zach, pour sa rigueur mathématique, pour son goût – que je partage – du détail, de l'optimisation et des bonnes pratiques de programmation, et pour son enthousiasme inattendu lorsque je tentais, maladroitement, d'esquisser une critique du système académique ;

À mes colocataires de Saint-André avec qui je me sens avoir fait partie d'une petite communauté : François et Pierre qui m'y ont invité, Martin et Thiébaud avec qui j'eus le plaisir d'être confiné ;

À Anto et Oli : la jeunesse enthousiaste et extravertie, pour être venus me chercher à la fin de la pandémie, pour m'avoir réservé une place à leurs côtés et une douce lueur d'admiration dans leur regard alors qu'il me semblait n'avoir rien fait pour les mériter ;

À Estelle, Moéa et Loïc, pour la relation de bienveillance ininterrompue qui nous relie depuis le début, pour la confiance qui y a fleuri et pour les discussions qui ont donné raison à Anne Sylvestre : « *en mettant bout à bout toutes nos solitudes, on s'est senti un peu moins effrayé* » ;

À Yann, là-bas au plat pays, pour m'avoir distrait de mes longues soirées solitaires d'interrogations, et pour avoir entretenu, chez lui comme chez moi, le goût de raconter des histoires ;

À Yann, ici au pays des érables, pour sa présence rassurante à mes côtés depuis huit ans, et pour m'avoir rappelé que, s'il y a un rôle dans lequel je m'épanouis et trouve du sens, c'est celui de comprendre pour vulgariser et transmettre ;

Enfin, à mes deux frères, Baptiste et François, pour les jeux, les rires et les souvenirs ;

À mon père pour son soutien et sa reconnaissance ;

À ma chère maman pour m'avoir appris à réfléchir, à grandir et à ne jamais renoncer, pour m'avoir donné le goût de la liberté et pour avoir tracé des chemins pour y accéder ;

Et finalement, à Emmanuelle pour avoir lu, relu et corrigé cette thèse, pour m'avoir soutenu dans sa rédaction, et plus que tout pour m'avoir accueilli dans ses bras et protégé de moi-même en m'enveloppant de sa tendresse, pour avoir cru en moi quand je n'y croyais pas, pour m'avoir rappelé comment cheminer, un pas après l'autre, et pour m'avoir aimé et appris à aimer.

À vous tous, du fond du cœur, merci !
Ce travail, je vous le dois aussi.

RÉSUMÉ

En 1851, Hermann von Helmholtz met au point un appareil permettant l'observation directe du fond de l'œil, révélant au monde la surface rétinienne et les vaisseaux qui l'arpentent. Le réseau vasculaire rétinien suscite alors un espoir inédit au sein du corps médical. Son observation simple et non intrusive est en effet singulière pour des vaisseaux du corps humain. Pourrait-elle permettre d'établir le lien entre les anomalies de la microvasculature rétinienne et celles des vaisseaux irriguant d'autres organes comme le cœur, le cerveau, les reins, etc. ? Il naît alors dans le corps médical un espoir et une ambition : découvrir sur la rétine des signes précurseurs de pathologies vasculaires. Afin de permettre la conduite et la comparaison d'études sur de larges cohortes, des indicateurs non subjectifs des altérations vasculaires rétiniennes sont développés comme les mesures de calibres équivalents de l'artère et de la veine centrales, ou comme la tortuosité des vaisseaux. Rapidement, apparait la volonté de mesurer ces paramètres vasculaires par des méthodes semi puis complètement automatiques.

Quatre étapes constituent ces méthodes : **1.** la segmentation des vaisseaux rétiniens ; **2.** leur classification en artérioles et veinules ; **3.** l'extraction topologique de l'arbre vasculaire ; **4.** sa paramétrisation. Ensemble, elles constituent la modélisation topologique et géométrique de l'arbre vasculaire à partir d'images de fond d'œil, qui sera le sujet de ce document.

Depuis 2015, les réseaux de neurones convolutifs ont reçu beaucoup d'attention de la communauté de chercheurs travaillant sur l'identification des artérioles et veinules rétiniennes. Si ces modèles culminent en haut des classements de l'état de l'art, ils restent limités par leur tendance au surapprentissage et par leur incapacité à apprendre la structure topologique sous-jacente à la vasculature de la rétine. À cause du faible volume de données annotées disponibles, les modèles de classification souffrent particulièrement de problèmes de généralisation et sont sujets à des incohérences topologiques : inversant régulièrement la classification d'un vaisseau au milieu d'une branche. Ces incohérences sont problématiques pour l'extraction fiable de la topologie de l'arbre vasculaire. Tant et si bien que les récents logiciels qui exploitent ces modèles pour la paramétrisation vasculaire automatique abandonnent la modélisation topologique (au détriment des paramètres vasculaires qui en dépendent) ou reposent sur des approches simplistes qui ignorent ces incohérences. *Comment mettre à profit alors, les recherches récentes appliquant les CNN à la segmentation et la classification des vaisseaux sur image de fond d'œil, pour extraire une modélisation topologique robuste de l'arbre vasculaire rétinien en vue d'en extraire des paramètres cliniques fiables ?* Pour répondre à cette problématique, je formule trois objectifs spécifiques.

Le premier objectif est la construction d'un nouveau jeu de données annotées (MAPLES-DR) pour l'entraînement et la validation de modèles de segmentation et de classification des vaisseaux rétiniens. Cet objectif fut réalisé par l'organisation d'une campagne d'annotation visant 4 structures anatomiques (dont les vaisseaux) et 6 structures pathologiques, symptomatiques de la rétinopathie diabétique sur 200 images de fond d'œil. L'ampleur du travail a nécessité la conception d'une plateforme d'annotation Web permettant la correction de cartes préannotées. L'étude de variabilité inter-observateur conduite pour l'occasion révèle d'importants désaccords entre rétinologues sur le contour exact et la détection des structures pathologiques.

Le second objectif est la formulation d'une nouvelle architecture de réseaux convolutifs pour la classification vasculaire : les *steered* CNN. Les filtres des neurones convolutifs sont reparamétrisés en des *steerables filters* afin de piloter, pour chaque pixel de l'image, la direction dans laquelle ils sont orientés avant d'être convolués. Pour la classification vasculaire, ce formalisme est utilisé pour maintenir les filtres convolutifs alignés avec la direction de chaque vaisseau de l'image. La mutualisation paramétrique qui en résulte assure une meilleure capacité de généralisation au modèle, surtout lorsque le nombre de paramètres ou d'échantillons d'entraînement est limité.

Le troisième objectif est la modélisation topologique et géométrique de l'arbre vasculaire rétinien à partir de cartes imparfaites de segmentation sémantique prédites par réseaux de neurones convolutifs (CNN). Cette ambition a conduit à l'implémentation du Fundus Vessels Toolkit (FVT) : une librairie Python qui fournit les outils pour extraire minutieusement une représentation topologique et géométrique du graphe vasculaire rétinien à partir de ces cartes imparfaites. Une nouvelle méthode de partition du graphe permet alors de contraindre le graphe vasculaire ainsi obtenu à adopter la forme d'une arborescence cliniquement plausible, et ce faisant corrige bon nombre des artefacts de segmentation et de classification initialement prédits par le CNN. Par ce simple post-traitement, FVT permet d'augmenter l'exactitude de classification et de reconnecter 75% des branches déconnectées de leurs vaisseaux. Cette représentation en arborescence permet finalement la paramétrisation du réseau vasculaire rétinien.

Dans un chapitre complémentaire, je reviens sur le paradoxe de la littérature récente concernant la segmentation des vaisseaux rétiniens : le nombre de publications s'envole depuis 2020, mais les performances stagnent. Je montre par une analyse lexicographique des 1738 papiers de la littérature que ces méthodes ne sont plus limitées par des lacunes architecturales ou d'entraînement, mais par un retard des protocoles pour les évaluer : dû à la fois à un manque de diversité des jeux de données de test, mais aussi à l'utilisation de métriques trop sensibles

au biais d'annotations.

L'ensemble de ces travaux pose les termes de la transposition des récents progrès technologiques de segmentation et de classification de la vasculature rétinienne, vers des résultats cliniques pour le dépistage des pathologies vasculaires.

ABSTRACT

In 1851, Hermann von Helmholtz developed an instrument that enabled direct observation of the ocular fundus, revealing the retinal surface and its intricate vasculature to the scientific community. The retinal vascular network quickly gained interest from the medical field, as its straightforward and non-invasive observation offers a unique window into the human circulatory system. Could abnormalities in the retinal microvasculature be indicators of vascular dysfunctions in other organs such as the heart, brain, or kidneys? Could they serve as early signs of vascular pathologies? To support large-scale cohort studies and enable objective comparisons, quantifiable and non-subjective indicators of retinal vascular abnormalities were developed, including the equivalent calibers of the central artery and vein, and vascular tortuosity. Soon, the need for semiautomatic and eventually fully automatic measurement methods became apparent.

These automatic methods typically consist of four key stages: **(1)** segmentation of retinal vessels; **(2)** classification of vessels into arterioles and venules; **(3)** topological extraction of the vascular tree; and **(4)** its geometric parameterization. Together, these stages form the basis for the central topic of this thesis, namely the topological and geometric modeling of the vascular tree from fundus images.

Since 2015, convolutional neural networks (CNNs) have attracted significant attention for the segmentation and classification of retinal vessels. Despite achieving state-of-the-art performance, such models remain constrained by their propensity to overfit and their limited ability to learn the underlying topological structure of the retinal vasculature. The scarcity of annotated data exacerbates generalization issues, leading to frequent topological inconsistencies, such as the misclassification of a vessel midway along its branch. These inconsistencies hinder the reliable extraction of vascular topology. As a result, most recent software tools either forego topological modeling altogether (at the expense of vasular parameters that depend on it), or rely on simplistic heuristics that ignore such errors. *How, then, can recent CNN-based advances in vessel segmentation and classification be leveraged to produce a robust topological model of the retinal vascular tree for the extraction of clinically meaningful parameters?* This thesis addresses this question through three specific objectives.

The **first objective** is the construction of a new annotated dataset, *MAPLES-DR*, for training and evaluating models that perform the segmentation and classification of retinal vessels. This objective was achieved through a large-scale annotation campaign during which four anatomical structures (including vessels) and six pathological structures associated with di-

abetic retinopathy were manually labeled in 200 fundus images. To facilitate this effort, a Web-based annotation platform was developed, enabling the correction of pre-annotated maps. An interobserver variability study conducted as part of this campaign revealed substantial disagreements among retinal specialists regarding the exact contours and detection of pathological structures.

The **second objective** involves the development of a novel CNN architecture for vessel classification: the *steered* CNN. In this design, convolutional filters are reparameterized into steerable filters that can be dynamically aligned with the local orientation of vessels for each pixel independently. This formalism enforces rotational consistency and reduces redundancy in parameterization by sharing parameters across all vessel orientations. The steetable filters improve the model's generalization capabilities, especially when few trainable parameters or training samples are available.

The **third objective** focuses on the topological and geometric modeling of the retinal vascular tree from imperfect semantic segmentation maps generated by CNNs. This led to the development of the *Fundus Vessels Toolkit* (FVT), a Python library that enables the extraction of a detailed topological and geometric representation of the vascular graph from imperfact segmentation maps. A novel graph partitioning algorithm constrains the extracted graph to form a clinically plausible tree structure, correcting many segmentation and classification artifacts produced by CNNs. This simple post-processing step significantly improves classification accuracy and reconnects up to 75% of previously disconnected vessel branches. The resulting tree representation enables accurate parameterization of the retinal vascular network.

In a complementary chapter, the thesis examines a paradox in the recent literature: while the number of publications on retinal vessel segmentation has surged since 2020, performance metrics have stagnated. Through a lexicographic analysis of 1,738 papers, it is demonstrated that progress is now hindered not by architectural or training limitations, but by outdated evaluation protocols. The latter problem can be ascribed both to insufficient diversity in test datasets and to using performance metrics that are overly sensitive to annotation biases.

Together, these contributions lay the foundation for translating recent advances in retinal vessel segmentation and classification into clinically meaningful outcomes for vascular disease screening.

TABLE DES MATIÈRES

DÉDICACE iii

REMERCIEMENTS iv

RÉSUMÉ vi

ABSTRACT ix

TABLE DES MATIÈRES xi

LISTE DES TABLEAUX xv

LISTE DES FIGURES xvi

LISTE DES ANNEXES xx

LISTE DES SIGLES ET ABRÉVIATIONS xxi

FORMALISME MATHÉMATIQUE xxiii

CHAPITRE 1 INTRODUCTION 1

CHAPITRE 2 REVUE DE LITTÉRATURE 6

2.1 Genèse de la modélisation du réseau vasculaire rétinien 6

2.2 Modélisation automatique du graphe vasculaire rétinien 10

2.2.1 Défis techniques à la segmentation sémantique des vaisseaux rétiniens 11

2.2.2 Segmentation des vaisseaux rétiniens pré apprentissage profond . . . 14

2.2.3 Classification traditionnelle des artérioles et veinules rétiniennes . . . 18

2.2.4 Modélisation topologique du graphe vasculaire pour la classification A/V 20

2.2.5 Segmentation sémantique des vaisseaux rétiniens par réseaux de neurones 26

2.3 Paramétrisation du réseau vasculaire rétinien à partir de segmentation sémantique automatique 37

2.3.1 Paramètres vasculaires rétiniens 37

2.3.2 Logiciel Semi-Automatique et Automatique de mesure des paramètres vasculaires rétiniens 39

2.3.3 Limites des logiciels existants 41

CHAPITRE 3 MÉTHODOLOGIE GÉNÉRALE 43
3.1 Synthèse de la littérature 43
3.2 Problématique de recherche 45
3.3 Objectifs spécifiques 45
3.4 Approche générale 46
3.4.1 MAPLES-DR : une campagne d'annotation des structures anatomiques et pathologiques de la rétine 46
3.4.2 STEERED CNN : l'équivariance par rotation au service de la classification des artères et veines rétiniennes 47
3.4.3 FUNDUS VESSELS TOOLKIT : modélisation géométrique et topologique de l'arbre vasculaire rétinien 48

CHAPITRE 4 MAPLES-DR : UNE CAMPAGNE D'ANNOTATION DES STRUCTURES ANATOMIQUES ET PATHOLOGIQUES DE LA RÉTINE 49
4.1 Motivations et travaux connexes 50
4.2 Méthode d'annotation de MAPLES-DR 52
4.2.1 Sélection des images 52
4.2.2 Sélection des structures anatomiques et pathologiques à annoter . . . 53
4.2.3 Recrutement des annotateurs 54
4.2.4 Conception du protocole d'annotation 55
4.2.5 Bilan du travail d'annotation 60
4.2.6 Préparation des données pour la publication 61
4.3 Évaluation de la fiabilité des annotations 63
4.3.1 Estimation de la variabilité inter-observateur sur les duplicatas 63
4.3.2 Variabilité inter-observateur sans préannotation de lésions 65
4.3.3 Variabilité inter-observateur avec préannotation de lésions 67
4.3.4 Biais introduit par la préannotation des lésions 68
4.3.5 Entraînement de modèles de segmentations avec MAPLES-DR 70
4.4 Discussion et Contributions 71

CHAPITRE 5 STEERED CNN : L'ÉQUIVARIANCE PAR ROTATION AU SERVICE DE LA CLASSIFICATION DES ARTÈRES ET VEINES RÉTINIENNES 73
5.1 Motivations et travaux connexes 73
5.1.1 Motivations à l'équivariance par rotation 73
5.1.2 CNN et équivariance par rotation 75
5.1.3 Steerable Filters : des filtres de convolution orientables 76
5.1.4 Steerable CNN 78

5.2 Étude préliminaire : isotropismes des CNN classifieurs de vaisseaux 78
5.3 Méthode 83
5.3.1 Objectif et vue schématique des *steered* CNN 83
5.3.2 Calcul du champ de direction $\vec{\boldsymbol{\alpha}}$ 84
5.3.3 Reparamétrisation des noyaux de convolution 85
5.3.4 Différences théoriques entre les CNN standards et les *steered* CNN . . 87
5.3.5 Implémentation optimisée des *steered* CNN 88
5.3.6 Réglage du jeu des filtres élémentaires 90
5.4 Évaluation expérimentale des *steered* CNN 92
5.4.1 Protocole expérimental 92
5.4.2 Performance générale de classification 94
5.4.3 Équivariance à la rotation et anisotropisme des steered convolutional neurons (SCN) 95
5.4.4 Efficacité de la modélisation 98
5.5 Discussion et Contributions 101

CHAPITRE 6 FUNDUS VESSELS TOOLKIT : MODÉLISATION GÉOMÉTRIQUE ET TOPOLOGIQUE DE L'ARBRE VASCULAIRE RÉTINIEN 102
6.1 Enjeux et Motivations 102
6.2 Méthode 104
6.2.1 Extraction initiale de la topologie 105
6.2.2 Extraction de la géométrie de chaque branche 107
6.2.3 Intégration de la classification A/V au graphe vasculaire 112
6.2.4 Résolution de l'arborescence propre à chaque vaisseau 115
6.3 Évaluation de la modélisation topologique 119
6.3.1 Correction des artéfacts topologiques 119
6.3.2 Exemple d'application : Étude pilote sur CLSA 122
6.4 Discussion et Contributions 123

CHAPITRE 7 CHAPITRE COMPLÉMENTAIRE : LE PARADOXE DE LA LITTÉRATURE SUR LA SEGMENTATION DES VAISSEAUX RÉTINIENS 125
7.1 Méthode de revue systématique et automatique 125
7.2 Sonder la littérature de segmentation vasculaire rétinienne 126
7.3 Pistes d'explications du paradoxe 129

CHAPITRE 8 DISCUSSION GÉNÉRALE 132
8.1 MAPLES-DR 132

8.2 *Steered* CNN 133
8.3 Fundus Vessels Toolkit 134

CHAPITRE 9 CONCLUSION 136
9.1 Synthèse des travaux 136
9.2 Recommandation et travaux futurs 137

RÉFÉRENCES 139

ANNEXES 166

LISTE DES TABLEAUX

Tableau 2.1 Jeux de données publics d'images de fond d'oeil avec annotations vasculaires. 13

Tableau 2.2 Comparaison des logiciels automatiques et semi-automatiques de mesure des paramètres vasculaires rétiniens 40

Tableau 4.1 Performance en généralisation de modèles de segmentation de lésions par ensemble d'entraînement. Les jeux de données d'entraînement sont triés en fonction de leur classement moyen sur tous les ensembles de test. 70

Tableau 5.1 Performances (aire sous la courbe ROC (AUC-ROC)) des 5 meilleurs modèles *steered* et standard 94

Tableau 5.2 Effet de la présence ou l'absence de rotation aléatoires des échantillons d'entraînement sur la métrique d'AUC calculée sur les ensembles de test. 96

Tableau 6.1 Correction par FVT des artefacts de segmentation sémantique sur les images de MAPLES-DR. 121

Tableau 6.2 Associations entre les tortuosités calculées par FVT ou Automorph, et les accidents vasculaire-cérébraux (AVC) dans la cohorte *Baseline* de CLSA. (Les p-values significatives sont indiquées en gras.) 122

LISTE DES FIGURES

Figure 1.1	Schéma des principales structures anatomiques de la rétine.	3
Figure 2.1	Nombre de publications annuelles cumulées concernant la segmentation et la classification des vaisseaux rétiniens à partir d'image de fond d'oeil.	10
Figure 2.2	Exemples de segmentations vasculaires difficiles dans la base de données HRF.	11
Figure 2.3	Colorations d'une même rétine imagée par différentes caméras. *Source : [1]*	12
Figure 2.4	Vue schématique des étapes précédant la classification des branches vasculaires.	18
Figure 2.5	Partitionnement d'un graphe de branches pour isoler chaque vaisseau.	21
Figure 2.6	Résolution de la connectivité locale d'un nœud reliant 4 branches incidentes.	22
Figure 2.7	Construction d'un graphe de ligne à partir de la segmentation vasculaire.	23
Figure 2.8	Variété d'arborescence induisant un même graphe vasculaire non-orienté.	24
Figure 2.9	Typologie des grandeurs géométriques à une bifurcation.	38
Figure 4.1	Résumé graphique du protocole d'annotation de MAPLES-DR.	49
Figure 4.2	Différences de prévalence de rétinopathie diabétique (DR) et œdème maculaire (ME) entre MAPLES-DR (violet), MESSIDOR au complet (turquoise) et une population dépistée par téléophtalmologie [2] à Toronto (gris).	52
Figure 4.3	Structures anatomiques annotées dans MAPLES-DR (Bleu sombre : disque optique, bleu clair : papille, violet sombre : macula, violet clair : vaisseaux).	53
Figure 4.4	Exemples d'annotations de structures pathologiques	55
Figure 4.5	Prétraitements disponible sur la plateforme d'annotation pour améliorer la lecture des image de fond d'oeil	56
Figure 4.6	Interface de sélection d'une tâche parmi celles assignées.	58
Figure 4.7	Interface d'annotation des structures rétiniennes.	58
Figure 4.8	Proportion d'images annotées par chaque rétinologue.	60

Figure 4.9 Vue quantitative du travail d'annotation pour chaque catégorie de structures rétiniennes. Les barres sombres correspondent aux segmentations ajoutées manuellement, celles claires correspondent aux préannotations effacées. Chacune est annotée du nombre de pixels corrigés par image. Le pourcentage indique l'aire occupée par chaque structure rapportée à l'aire des images. 60
Figure 4.10 Préannotations corrigées par les rétinologues. (*Blanc* : préannotation inchangée ; *Turquoise* : segmentation ajoutée manuellement, *Violet* : préannotation effacée manuellement.) 61
Figure 4.11 Page d'accueil de la documentation du paquet `maples-dr`. 62
Figure 4.12 Visualisation comparative des annotations vasculaires pour les images dupliquées de MAPLES-DR. (*Blanc* : Pixels communs aux deux annotations, *Violet* : pixels inclus uniquement dans le jeu d'entraînement, *Bleu* : pixels inclus comme doublons) 63
Figure 4.13 Visualisation comparative des lésions rouges (en haut) et lésions claires (en bas) pour les images dupliquées de MAPLES-DR. (Même légende que la Figure 4.12.) 64
Figure 4.14 Variabilité inter-observateur lors de segmentation de lésions sans préannotation. (En haut : variabilité de détection ; en bas : variabilité de segmentation. Chaque paire d'annotateurs est évaluée individuellement. La ligne pointillée correspond à la moyenne des paires.) 66
Figure 4.15 Variabilité Inter-observateur lors de segmentation de lésions **avec** préannotation. (À gauche : variabilité de détection ; à droite : variabilité de segmentation.) 68
Figure 4.16 Distribution de la taille et du nombre de lésions par images avec et sans préannotation. (*O1*, *O2*, *O3* : annotations manuelles sans préannotations ; *CP* : annotation par correction de préannotations ; *IA* : préannotations avant corrections. La barre verticale indique la valeur médiane.) 69
Figure 5.1 Variations de la classification A/V prédite par un même CNN, pour différentes rotations initiales de l'image d'entrée 74
Figure 5.2 Schéma de la méthode de calcul de $\mathbf{\Phi}(\theta)$. 79
Figure 5.3 Exemples de décomposition polaire de la densité spectrale de puissance (DSP) pour des harmoniques polaires de rang 0, 1, 2 et 3 ; et pour deux noyaux de convolution symétrique et anti-symétrique. 80
Figure 5.4 Comparaison du tropisme de la prédiction de CNN entraîné avec et sans augmentation de données géométriques. 81

Figure 5.5 Effet de l'augmentation de données sur les composantes polaires de la Densité Spectrale de Puissance pour chaque couche convolutive d'un CNN classifieur d'A/V. 82
Figure 5.6 Champ d'orientation $\vec{\boldsymbol{\alpha}}$ calculé à partir de la squelettisation de la segmentation vasculaire. *À gauche* : Champ d'influence généré par 4 pixels individuels du squelette (dessinés en bleu). *À droite* : Champ complet, superposé au squelette et à la segmentation vasculaire. (Pour des raisons de lisibilité, les champs affichés sont sous-échantillonnés.) 84
Figure 5.7 Le jeu de filtres élémentaires $\Psi_{k,r}$ échantillonné en coordonnées cartésiennes. 86
Figure 5.8 Précision moyenne d'un *steered* CNN sur son ensemble de test en fonction du rang harmonique maximum K de ces filtres élémentaires $\boldsymbol{\Psi}_{k,r}$. 90
Figure 5.9 Architecture de modèle U-Net retenue pour évaluer les *steered* CNN. 92
Figure 5.10 Classification A/V par un *steered* CNN (gauche) et un CNN standard (droite). 95
Figure 5.11 Distribution des poids $\omega_{k,r}$ d'un *steered* CNN regroupés par harmonique polaire k et par couche convolutive. 97
Figure 5.12 Performances comparées du *steered* CNN avec un CNN standard en fonction du nombre de paramètres. 98
Figure 5.13 Performances comparées du *steered* CNN avec un CNN standard en fonction du volume d'entraînement. 99
Figure 5.14 Exemple de regression par une courbe exponentielle de l'évolution de l'AUC-ROC de validation ; et de mesure géométrique de la demi-vie λ. 100
Figure 5.15 Rapidité de convergence lors de l'entraînement de CNN *steered* et standard. 100
Figure 6.1 Artefacts de segmentation sémantique des vaisseaux rétiniens par CNN. 103
Figure 6.2 Vue d'ensemble de l'extraction de l'arbre vasculaire dans FVT. 104
Figure 6.3 Masques de détecteurs *hit-or-miss* pour la détection des noeuds du graphe. 105
Figure 6.4 Illustration de la correction des croix creuses (CC). 106
Figure 6.5 Nettoyage du squelette aux extrémités des branches. 108
Figure 6.6 Informations géométriques extraites pour une branche vasculaire. . . 109
Figure 6.7 Découpage itératif de la courbe d'une branche en une B-spline cubique. 111

Figure 6.8 Correction des artefacts topologiques. **Gauche :** Graphe colorié par les classes artère/veine (A/V) (*rouge : **ART** ; bleu : **VEI** ; violet : **BOTH** ; vert : **UNK***). **Droite :** Graphe simplifié. 113
Figure 6.9 Illustration de la recherche des reconnections possibles. 114
Figure 6.10 Illustration de la construction du graphe de ligne. **a.** Graphe vasculaire $G(V, B)$; **b.** Exemple d'arborescence que G peut générer ; **c.** Graphe de ligne listant toutes les arborescences possibles. 116
Figure 6.11 Illustration de l'optimisation d'arborescence **a.** Graphe vasculaire G ; **b.** Graphe de ligne orienté $L(G)$; **c.** Graphe de ligne simplifié $L^*(G)$. (Les arcs de lignes correspondants aux hypothèses rejetées par la arborescence couvrante maximale (MSA) sont semi-transparents.) 118
Figure 6.12 Arborescence extraite par RMHAS (gauche) et par FVT (droite). . . 120
Figure 7.1 Nombre de publications par an et par pays traitant de la segmentation des vaisseaux rétiniens sur images de fond d'oeil. 125
Figure 7.2 Nuages de mots des titres d'articles sur la segmentation des vaisseaux rétiniens. 127
Figure 7.3 Occurence entre 2005 et 2024 des termes méthodologiques dans le titre et les sections de méthodes des articles de segmentation des vaisseaux rétiniens. 128
Figure 7.4 Occurence des métriques et jeu de données annotées pour la segmentation des vaisseaux rétiniens dans les sections de validation et les tableaux des articles du corpus. 129

LISTE DES ANNEXES

Annexe A Variance du signal d'inférence dans un Steered CNN 166

LISTE DES SIGLES ET ABRÉVIATIONS

A/V	Artère/Veine
AUC-ROC	Aire sous la courbe ROC
AVC	Accident Vasculaire Cérébral
AVR	Ratio du calibre Artère/Veine (Artery/Vein Ratio)
CNN	Réseau de neurones convolutifs (Convolutional Neural Network)
CRAE	Calibre Équivalent de l'Artère Centrale Rétinienne (Central Retinal Artery Equivalent)
CRF	Champ Aléatoire Conditionnel (Conditional Random Field)
CRVE	Calibre Équivalent de la Veine Centrale Rétinienne (Central Retinal Vein Equivalent)
CWS	Nodules Cotonneux (Cotton Wool Spots)
DL	Apprentissage Profond (Deep Learning)
DR	Rétinopathie Diabétique (Diabetic Retinopathy)
DSP	Densité Spectrale de Puissance
DMLA	Dégénérescence Maculaire Liée à l'Âge
FCN	Réseau de neurones complètement convolutif (Fully-Convolutional Neural Network)
FVT	Fundus Vessels Toolkit (le paquet Python implémentant la méthode du chapitre 6)
GAN	Réseau Adversarial Génératif (Generative Adversarial Network)
GNN	Réseau de neurones par graphe (Graph Neural Network)
HSL	Espace chromatique : Teinte Saturation Luminosité (Hue Saturation Luminosity)
IA	Intelligence Artificielle
IoU	Intersection sur Union (Intersection over Union)
IRMA	Anomalies Microvasculaires Intrarétiniennes (Intra-Retinal Microvascular Abnormality)
k-NN	Classifieur par k plus proches voisins (k-Nearest Neighbour)
LAB	Espace chromatique : Luminance A B
LDA	Classifieur par analyse discriminante linéaire (Linear Discriminant Analysis)
LLM	Large Modèle de Langage (Large Language Model)

ME	Œdème Maculaire (Macular Edema)
morphomat	Morphologie Mathématique
MRF	Matched Response Filters
MSA	Arborescence Couvrante Maximale (Maximum Spanning Arborescence)
PCA	Analyse en composantes principales (Principal Component Analysis)
ROC	Fonction d'efficacité du récepteur ou courbe ROC (Receiver Operating Characteristic)
SCN	Neurone convolutif dont l'orientation est orientable (Steered Convolutional Neuron)
SVM	Machines à vecteurs de support (Support-Vector Machine)
ViT	Modèle transformeur pour la vision (Vision Transformer)

FORMALISME MATHÉMATIQUE

Formatage général des équations

Ce document utilise des conventions de formatage pour ses équations, afin d'en simplifier la lecture : les variables vectorielles ou matricielles seront notées en gras, pour les distinguer des variables scalaires. Les vecteurs à deux dimensions (indiquant une direction et une norme) seront indiqués par une flèche : $\vec{\alpha}$; et leurs champs seront notés en gras : $\vec{\boldsymbol{\alpha}}$.

Nombres complexes

$z = x + \mathrm{i}y$	$\in \mathbb{C}$	Nombre complexe. La variable imaginaire est notée : i (à ne pas confondre avec l'indice : i)
$\mathrm{Re}(z)$	$= x$	Partie réelle d'un nombre complexe
$\mathrm{Im}(z)$	$= y$	Partie imaginaire d'un nombre complexe
$\lvert z \rvert$	$= \sqrt{x^2 + y^2}$	Module d'un nombre complexe
$\arg(z)$	$= \arctan(y/x)$	Argument d'un nombre complexe

Système de coordonnées, repères et indices

(x, y)	$\in \mathbb{R}^2$	Coordonnées cartésiennes continues
(i, j)	$\in \mathbb{N}^2$	Coordonnées cartésiennes discrètes positives, utilisées comme repère pour les images. Par convention l'origine du repère est placée en haut à gauche de l'image, i est orienté vers la droite et j vers le bas.
(p, q)	$\in \mathbb{Z}^2$	Coordonnées cartésiennes discrètes positives ou négatives. Cette variante de (i, j) est utilisée pour les filtres de convolution. L'origine du repère $(p = 0, q = 0)$ correspond au centre du filtre.
(ρ, θ)	$\in (\mathbb{R}^+, [0; 2\pi[)$	Coordonnées polaires où ρ décrit la composante radiale et θ la composante angulaire.

Transformées

$\mathcal{F}\Big[\mathbf{X}\Big]$	Transformée de Fourier de $\mathbf{X}$
$\mathrm{Rot}_\alpha\Big[\mathrm{X}(\rho,\theta)\Big] = \mathrm{X}(\rho,\theta-\alpha)$	Rotation de $\mathbf{X}$ par un angle α
$\mathbf{X} \underset{(i,j)}{*} \mathbf{W} = \sum_{p,q} \mathrm{X}(i-p, j-q) \cdot \mathrm{W}(p,q)$	Convolution de $\mathbf{X}$ par $\mathbf{W}$, calculée autour des coordonnées (i,j) de $\mathbf{X}$
$\mathbf{X} \underset{(i,j)}{\circledast} \mathbf{W} = \sum_{p,q} \mathrm{X}(i+p, j+q) \cdot \mathrm{W}(p,q)$	Corrélation croisée de $\mathbf{X}$ par $\mathbf{W}$, calculée autour des coordonnées (i,j) de $\mathbf{X}$

Graphe

$G(V,B)$	Graphe vasculaire non-dirigé défini par un ensemble de noeuds V et un ensemble de branches B.
$v \in V$	Noeud du graphe vasculaire (*vertex* en anglais). Il peut-être de types : - *orphelin* : s'il n'est connecté à aucun autre noeud ; - *terminaison* : s'il n'est connecté qu'à un seul noeud ; - *jonctions* : s'il est connecté à plus d'un noeud. On distingue alors l'*intersection* où deux vaisseaux se croise et la *bifurcation* où un vaisseau se ramifie.
b $\in B$	Arc du graphe reliant deux noeuds $(u,v) \in V^2$. Il représente une branche c'est à dire une section d'un vaisseau comprise entre deux jonctions ou terminaisons. Une branche est dite *terminale* si u ou v sont des terminaisons.
$\mathrm{Adj}(v) \subset V$	Ensemble des noeuds voisins (adjacents) à v.

CHAPITRE 1 INTRODUCTION

«Nowhere else in the living body are we able to study the blood vessels of the size of the central artery and its branches to the finest arterioles than in the fundus. The man with the ophthalmoscope is impressed by its importance in the recognition of the early signs of general disease and of its general diagnostic and its prognostic value. But its greatest value lies in the possibility it affords to help solve the problems that have so long baffled medical science, in the search for ultimate causes of conditions that are so common, so varied in their manifestations and so serious in their course.»

H. Friedenwald, Oxford 1930 [3]

La présence de vaisseaux sanguins à la surface du fond d'œil est connue dès le 18^{e} siècle : Jean Méry les aurait remarqués pour la première fois en 1704 dans les yeux de son chat, lorsqu'il lui maintenait la tête sous l'eau. Il faut cependant attendre 1851 pour que l'Allemand Hermann von Helmholtz mette au point un appareil permettant l'observation de la rétine d'un individu sans qu'il soit nécessaire de l'immerger. L'*ophtalmoscope* est de conception rudimentaire : une source de lumière est reflétée par une surface semi-refléchissante vers la pupille du patient. Un observateur plongeant alors son regard au travers de la surface voit sa ligne de vue s'aligner avec les rayons illuminant le fond d'œil, et y découvre le réseau vasculaire rétinien.

La simplicité de l'appareil conduit à son succès immédiat. Il faut dire que l'observation directe et non intrusive de vaisseaux sanguins est inédite dans le corps humain. Pourrait-elle permettre d'établir le lien entre les anomalies de la microvasculature rétinienne et celles des vaisseaux irriguant d'autres organes comme le cœur, le cerveau, les reins ? Il naît alors dans le corps médical un espoir et une ambition : découvrir sur la rétine des signes précurseurs de pathologies vasculaires. Autrement dit, utiliser les yeux comme fenêtre, non pas de l'âme comme le disait Georges Rodenbach, mais du cœur et du système vasculaire en général pour mieux comprendre et anticiper ses défaillances.

Aussi, pendant près de trois siècles, les ophtalmologues ont regardé leurs patients dans le noir de la pupille, d'abord par des observations directes via un ophtalmoscope puis par *imagerie de fond d'œil* ("fundus" en anglais). Cette famille d'imagerie désigne tout système d'acquisition capable de capturer une image des tissus semi-transparents rétiniens par réflexion de lumière sur la rétine [4]. Depuis la première commercialisation par Zeiss en 1926 de photographies monochromes de la rétine, les technologies d'acquisition ont bien évolué : les caméras fond d'œil contemporaines capturent en couleur et à haute résolution jusqu'à un angle de 133 ° de l'intérieur du globe oculaire. Dans cette thèse, on s'intéressera plus précisément aux photographies de champ standard (30 ° à 50 °) car cette modalité non invasive, rapide

et bon marché est largement répandue dans les services ophtalmologiques du monde entier. Elle a ainsi largement contribué aux études conduites sur de grandes cohortes de patients pour établir le lien entre altérations de la micro-vasculature rétinienne et d'autres pathologies vasculaires. Dans un premier temps, les ophtalmologues recensent ces altérations : rétrécissement des artères, augmentation de la tortuosité, etc. Puis, dans un objectif de standardisation des études, des indicateurs géométriques, morphologiques et topologiques de la vasculature rétinienne sont développés. Mais la mesure manuelle de ces indicateurs souffre d'une grande variabilité et est chronophage.

Depuis 30 ans, des yeux d'un autre genre s'ouvrent et scrutent avec ceux des cliniciens les images de la rétine, des yeux rapides, infatigables et constants : les yeux des algorithmes. Ces dernières années, la segmentation et la classification automatique du réseau vasculaire rétinien ont en effet suscité l'intérêt de la communauté de chercheurs en vision par ordinateur. Ces algorithmes, les réseaux de neurones particulièrement, ont une précision croissante et constituent une première étape vers l'automatisation et la standardisation de l'extraction d'indicateurs vasculaires cliniques. Cependant, pour compléter cette tâche, une seconde étape est nécessaire : la modélisation topologique de l'arbre vasculaire à partir des segmentations de vaisseaux. De plus, la définition de la topologie de l'arbre vasculaire rétinien nécessite l'extraction de deux autres structures qui constituent des "points cardinaux" permettant de s'y repérer et d'en estimer l'échelle : la *fovea* et le *disque optique*.

Structures anatomiques d'intérêt sur images de fond d'oeil

La surface extérieure du globe oculaire est une paroi blanche, rigide et opaque : la *sclère*. À l'intérieur de cette enveloppe protectrice, tapissant le fond de l'œil, deux tissus se superposent : la *rétine* et la *choroïde*. La choroïde est un tissu hautement vascularisé chargé d'alimenter en sang la surface externe de la rétine, dont la surface interne abrite les photorécepteurs qui nous donnent la vue. C'est cette dernière surface qui est scrutée par les ophtalmologues, initialement à l'aide d'un ophtalmoscope et aujourd'hui par *imagerie du fond d'œil* dont un exemple est donné à droite de la figure 1.1. Et c'est la vasculature de cette surface qui va nous intéresser au cours des 136 prochaines pages de ce projet de doctorat.

La fovea Notre acuité visuelle est maximale au centre de notre champ de vision. En effet, c'est au centre de la rétine que la densité de photorécepteurs – particulièrement celle des cônes responsables de la vision en couleurs – est la plus élevée. Cette zone absorbe donc plus la lumière, teintant le fond de la rétine et y apparaissant comme une tache sombre : la *macula*. Cette zone diffuse de 5 à 6 mm de diamètre contient en son cœur un point sombre bien défini : la *fovea*, mesurant environ 150 μm.

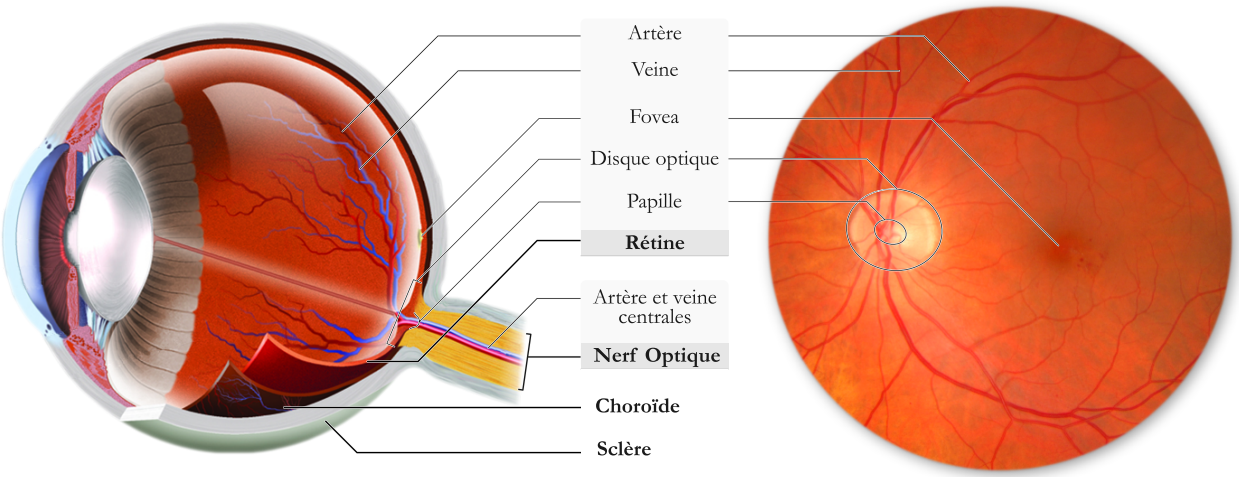


FIGURE 1.1 Schéma des principales structures anatomiques de la rétine.

Gauche : Vue en coupe du globe oculaire ; **Droite :** Image du fond d'œil centrée sur la fovea.

Le disque optique Le globe oculaire est connecté à son orbite par le nerf optique qui assure la liaison nerveuse entre les photorécepteurs rétiniens et le cortex visuel et qui alimente l'œil en sang par l'artère et la veine centrales. Le *disque optique* est la tête de ce nerf, c'est le point de convergence des axones (les terminaisons nerveuses) rétiniennes. Le disque optique se situe du côté nasal de la macula. Lorsqu'il est sain, son diamètre vertical mesure environ 1,7 mm et celui horizontal 1,5 mm. Sur l'imagerie de fond d'œil, il apparaît comme une zone claire dénuée de capteurs (causant d'ailleurs un point aveugle dans notre champ de vision). Sa surface n'est pas plane : en son centre, elle prend la forme d'un entonnoir d'où émerge l'artère et la veine centrales. Ce renfoncement nommé excavation papillaire ou *papille* ("cup" en anglais) se présente comme une ellipse claire au cœur du disque optique.

Les vaisseaux rétiniens À la sortie de la papille, l'artère et la veine centrales se ramifient en artérioles et veinules. Les artérioles acheminent le sang chargé d'oxygène vers les capillaires qui sillonnent la surface interne de la rétine et dont la paroi endothéliale permet des échanges avec les cellules photoréceptrices (notamment leur ravitaillement en oxygène et la récupération du dioxyde de carbone). Le sang appauvri en oxygène est alors collecté par les veinules puis par la veine centrale. Cette différence de composition sanguine se traduit par une couleur plus foncée des veinules [1]. La paroi des artérioles rétiniennes est – comme celle des artérioles cérébrales et coronaires – dotée de muscles lisses qui permettent à ces vaisseaux d'adapter leur calibre (leur diamètre) et donc leur débit sanguin aux besoins des cellules photoréceptrices qui varient avec la luminosité. Cette caractéristique fait parfois apparaître sur la ligne médiane du vaisseau un liseré blanc propre aux artérioles. Enfin, une dernière propriété distingue les deux types de vaisseaux : le diamètre des veinules est globalement plus large que celui des artérioles (respectivement 300 μm et 200 μm en moyenne).

1. La protéine d'hémoglobine oxyde ses atômes de fer pour transporter la molécule de dioxygène. Cette oxydation se traduit, dans le sang comme dans la rouille, par une teinte rouge vif.

Objectif de recherche

Avec l'avènement des réseaux de neurones entraînés par apprentissage profond, le nombre de publications traitant de la segmentation et de la classification des vaisseaux rétiniens explose : depuis 2020, il dépasse les 150 publications chaque année (plus de 200 en 2022) ! Mais, si ces articles motivent généralement leur recherche par des enjeux cliniques, peu de travaux évaluent les retombées réelles de ce foisonnement d'articles. La segmentation sémantique du réseau vasculaire n'est d'ailleurs que la première étape vers l'extraction des indicateurs cliniques : elle doit être suivie d'une extraction des informations géométriques, morphologiques et topologiques. Or, cette seconde étape n'est pas triviale lorsqu'elle doit travailler à partir de segmentations automatiques qui introduisent des artefacts topologiques inhérents aux approches basées uniquement sur une analyse des intensités de l'image.

Comment mettre ces progrès récents des techniques de segmentation sémantique des vaisseaux au service de la recherche clinique, en automatisant l'extraction des indicateurs vasculaires existants et potentiellement en permettant le développement de nouveaux ? Mon doctorat s'inscrit dans l'effort clinique de compréhension et de dépistage des rétinopathies diabétiques et de pathologies vasculaires. Il vise au développement et à la validation d'un outil automatique de modélisation topologique de la vasculature rétinienne, afin d'en permettre une analyse fiable et reproductible sur les images de fond d'œil issues de grandes cohortes.

Plan de la thèse

La présente thèse expose les travaux conduits pendant mon doctorat et leurs conclusions. Elle est structurée en neuf chapitres dont le *premier* est cette introduction. Le *second* est une revue de littérature traitant de la modélisation topologique du réseau vasculaire rétinien : ses motivations, l'évolution de ses algorithmes et ses limitations. Le *troisième* énonce les objectifs spécifiques de mon projet de recherche et sa méthodologie générale. Les trois chapitres suivants détaillent la méthode proposée pour atteindre chaque objectif spécifique. Ils sont inspirés des trois articles publiés durant mon doctorat, mais explorent plus en profondeur chaque sujet. Le *quatrième* est consacré à MAPLES-DR : une campagne d'annotation des structures anatomiques et pathologiques de la rétine ayant abouti à la publication d'un nouveau jeu de données annoté. Le *cinquième* expose un nouveau formalisme de neurones convolutifs : les *steered* CNN, spécifiquement conçus pour la classification des artérioles et veinules rétiniennes. Le *sixième* détaille l'élaboration du Fundus Vessels Toolkit, un outil de modélisation automatique de la vasculature rétinienne robuste aux artefacts de segmentation et de classification. Le *septième* chapitre introduit un paradoxe intéressant sur la littérature actuelle et le dépeint par une analyse lexicographique, employant un modèle de langage large (LLM) pour le triage de milliers d'articles scientifiques. Le *huitième*, la discussion, montre comment

les méthodes proposées aux chapitres 4, 5 et 6 répondent aux objectifs spécifiques et quelles limites persistent. Le *neuvième* et dernier chapitre de cette thèse en résume les contributions et formule des recommandations pour des travaux futurs.

CHAPITRE 2 REVUE DE LITTÉRATURE

Ce chapitre revisite les motivations cliniques à la modélisation topologique du réseau vasculaire rétinien. Puis, il propose un tour d'horizon de la littérature sur la segmentation et la classification automatique des vaisseaux, et sur la reconstruction de l'arbre vasculaire. Il se conclut par une revue des outils semi- et complètement automatiques de modélisation et de paramétrisation de l'arbre vasculaire rétinien et une analyse de leurs limites.

2.1 Genèse de la modélisation du réseau vasculaire rétinien

On verra ici comment la paramétrisation du réseau vasculaire rétinien s'est imposée comme un formidable outil pour analyser les effets des pathologies vasculaires systémiques sur les artérioles et veinules de la rétine.

Premières observations d'anomalies rétiniennes dues aux défaillances vasculaires

Dès 1898, Marcus Gunn recense et décrit les anomalies de la vasculature rétinienne chez des patients souffrant d'hypertension [5]. Il observe un rétrécissement généralisé et irrégulier du calibre (c'est-à-dire du diamètre) des artérioles, accompagné de perturbations des reflets lumineux le long de leur surface. Il note aussi l'apparition de lésions vasculaires : hémorragies, œdèmes et anomalies des capillaires (aujourd'hui appelées néovascularisation). Cette description évolue peu durant la première moitié du 20^{e} siècle, mais leur compréhension étiologique progresse. On associe la contraction généralisée des artérioles à l'élévation de la pression artérielle et on lui impute l'apparition de lésions vasculaires. Quant aux irrégularités de calibres, on les relie plutôt à une hypertension contrôlée ou à des cas de sclérose artériolaire [6].

Afin de permettre la conduite et la comparaison d'études sur de plus larges populations, des systèmes de classification standardisés de ces anomalies sont développés. Ainsi, en 1939, Keith, Wagener et Barker classent les symptômes de rétinopathie hypertensive selon 4 stades de gravité : **1.** rétrécissement généralisé ou focal des artérioles ; **2.** entailles artérioveineuses ; **3.** hémorragies ou exsudats ; **4.** œdème de la papille [7]. Ils proposent aussi pour chacun une définition précise. Mais en 1966, Kagan et al. montrent que malgré cet effort de standardisation, la détection de ces symptômes souffre d'une grande variabilité inter-observateur [8]. Cette variabilité se manifeste particulièrement lors de l'évaluation du rétrécissement des artérioles, alors que ce signe est peut-être le plus prometteur pour étudier les effets précoces des pathologies vasculaires sur la rétine.

Vers des indicateurs non-subjectifs des altérations vasculaires rétiniennes

Pour pallier cette variabilité, apparaît la nécessité d'une mesure quantitative, non subjective de la contraction généralisée des artérioles rétiniennes. Comme il n'est pas envisageable de faire mesurer à un clinicien le diamètre de toutes les artérioles du fond d'œil (en tout cas dans un temps raisonnable), Parr et al. [9] proposent en 1974 un protocole de mesure estimant le calibre de l'artère centrale de l'œil par une mesure des quelques artérioles à la périphérie du disque optique. La zone d'intérêt est limitée de 0.5 à 1 diamètre de disque autour de la papille : de sorte à en être suffisamment éloignée pour éviter les sections de transitions artères/artérioles, tout en se concentrant uniquement sur les premiers embranchements artériolaires Les diamètres de toutes les artérioles traversant cette zone sont mesurés individuellement par un expert puis agrégés en remontant la topologie de l'arbre artérien jusqu'à l'artère centrale [1]. Le calibre obtenu est nommé calibre équivalent de l'artère rétinienne centrale (CRAE) et peut être comparé d'un individu à l'autre. Sa méthode de mesure sera simplifiée en 2003 par Knudtson et al. [11] qui limitent le nombre d'artérioles considérées aux 6 plus importantes, après avoir remarqué que le nombre d'artérioles biaisait le calibre équivalent. Le CRAE est toujours utilisé sous cette forme aujourd'hui pour estimer le calibre global des artérioles rétiniennes.

En 1992, Hubbard et al. [12] adaptent la formule de Parr et Spears pour calculer le calibre équivalent de la veine rétinienne centrale (CRVE) et en déduisent le ratio du calibre artérioles/veinules (AVR) défini par $\text{AVR} = \frac{\text{CRAE}}{\text{CRVE}}$. En utilisant le CRVE comme référence, l'AVR est bien moins sensible à la variabilité naturelle du calibre des vaisseaux au sein d'une population. De ce fait, il capte rapidement l'attention de la recherche clinique qui projette de l'évaluer sur de larges cohortes. Il deviendra ainsi rapidement le porte-étendard des indicateurs quantitatifs de l'altération des vaisseaux rétiniens et d'autres émergeront dans son sillage, comme la tortuosité ou la dimension fractale (cf. section 2.3.1).

1. Pour chaque paire d'artérioles issues d'une même bifurcation, le calibre $\hat{W}$ de l'artère parent est dérivé par : $\hat{W} = \sqrt{0.87w_a^2 + 1.01w_b^2 - 0.22w_a w_b - 10.76}$ où w_a et w_b sont respectivement le calibre de l'artériole la plus petite et la plus large en μm. Cette agrégation tient ainsi compte du nombre d'embranchements artériolaires, puisque la surface totale du système vasculaire artériolaire augmente avec ce dernier et qu'il est très variable d'un individu à l'autre. Elle est aussi fonction du carré des calibres artériolaires de sortes à mieux représenter la capacité sanguine de l'artère centrale (qui est proportionnelle à sa section et non à son rayon) [10].

Premier protocole de mesure de l'AVR assistée par ordinateur

C'est dans ce contexte qu'en décembre 1999, Hubbard et al. [12] publient la première étude qui quantifie l'AVR sur une large cohorte de 9040 participants âgés de 48 à 73 ans. Ils montrent que l'AVR suit une distribution normale au sein de la population, que sa moyenne varie selon le sexe, l'ethnicité, le statut de fumeur et décroît avec l'âge. Par ailleurs, après avoir corrigé pour ces facteurs, ils observent bien une corrélation entre l'AVR et la pression artérielle et confirment quantitativement un rétrécissement généralisé des artérioles rétiniennes.

Mais le vrai intérêt de cette publication, pour nous, réside dans la méthode utilisée pour mesurer l'AVR sur une population de cette ampleur. Ce papier est en effet le premier (en tous cas à notre connaissance) à proposer l'utilisation d'un «*image processor*» pour standardiser l'évaluation d'un paramètre vasculaire rétinien. Étant donné la singularité d'un tel appareil pour l'époque, les auteurs décrivent leur protocole avec force détails. Le scan rétinien monochrome (canal vert) est tout d'abord numérisé par un scanner de film 35 mm et transféré au processeur d'image Sun Microsystems affublé d'un écran "haute résolution" de 19 pouces. L'annotateur sélectionne ensuite la région d'intérêt dont le contraste et la netteté sont améliorés numériquement. Il peut alors marquer les contours d'un vaisseau à l'aide d'un outil suffisamment saugrenu pour être mentionné par les auteurs : « *A mouse was used to mark the two edges of the vessel* » !

Du point de vue du traitement d'images, la méthode est rudimentaire : elle se limite à une amélioration des contrastes et une magnification. Mais les auteurs annoncent qu'ils travaillent à un algorithme pour automatiser la sélection manuelle des bords des vaisseaux et espèrent ainsi améliorer la reproductibilité et l'efficacité des mesures de calibre. Dès lors, l'histoire de la paramétrisation du réseau vasculaire est définitivement liée à celle de l'informatique biomédicale et du traitement d'image. Et les motivations invoquées en 1999 restent inchangées aujourd'hui : **1.** la standardisation des protocoles de mesures pour éviter la variabilité inter-observateur et comparer les résultats entre études ; **2.** leur automatisation pour rendre possible l'analyse de cohortes importantes et améliorer la force statistique des résultats.

L'utilisation d'une telle « débauche » d'outils technologiques (un écran 19 pouces et une souris !) contraint cependant les auteurs à souligner une limitation de leur méthode : « *Measurement of A/V ratio using a computerized image processor is practical in a research setting, but currently could not be easily done in a clinical setting.* »

Cette limitation sera très éphémère...

Premiers logiciels de mesure semi-automatique de l'AVR

À peine 3 ans plus tard, en 2002, la même équipe présente le logiciel Retinal Analysis [13]. Un outil capable de mesurer automatiquement le calibre des vaisseaux en détectant leurs contours sur des profils d'intensité extraits le long de leurs sections. Néanmoins, l'utilisateur doit toujours placer manuellement la grille de la zone de mesure, choisir les sections de vaisseaux à évaluer, valider les mesures de calibre et calculer le CRAE, le CRVE et l'AVR. Ainsi, si la reproductibilité du résultat est améliorée par la détection des contours, le protocole nécessite toujours 25 minutes pour extraire l'AVR d'une image.

Ce logiciel est remplacé quelques années plus tard par IVAN qui réduit de 5 minutes le temps de mesure par une pré-segmentation des vaisseaux. L'algorithme qui réalise cette tâche n'est pas documenté, on sait simplement qu'il s'appuie sur une carte de super-pixels parmi lesquels sont identifiés les segments vasculaires. Durant une décennie, son utilisation se répand dans le monde entier pour quantifier l'AVR sur des cohortes et investiguer l'effet des maladies cardiovasculaires sur la vasculature rétinienne de plusieurs milliers de patients. Pour ne donner que quelques exemples, il sera utilisé aux États-Unis : avec la *Beaver Dam Eye Study* [14,15] (diabète) ou la *Multi-Ethnic Study of Atherosclerosis* [16] (athérosclérose), en Australie : avec la *Blue Montains Eye Study* [17,18] (hypertension), ou encore en Europe avec la *Thessaloniki Eye Study* [19] (facteurs de risques cardiovasculaires).

En 2010, Cheung et al. [20] proposent d'étendre la région de mesure de l'AVR à deux diamètres du disque optique, alors qu'elle était jusque-là limitée à 1. Ils montrent que l'AVR ainsi obtenu est plus reproductible et a une meilleure corrélation statistique avec les facteurs de risques cardiovasculaires. Mais pour réussir cet exploit sans augmenter le temps d'annotation, leur logiciel SIVA automatise la segmentation du disque optique et des vaisseaux et identifie même les artérioles et veinules !

2.2 Modélisation automatique du graphe vasculaire rétinien

On l'a vu, le réseau vasculaire rétinien, singulier par la simplicité de son observation, est scruté et mesuré depuis le XIX^{e} siècle avec l'espoir de mieux comprendre et anticiper les pathologies cardiovasculaires et rétiniennes. Mais les protocoles de mesure de calibre développés à la fin du XX^{e} siècle se restreignent, par pragmatisme, à une petite région autour du disque optique. Pour étendre leur définition à une portion plus large de la rétine et pour standardiser la mesure d'autres paramètres vasculaires (comme la tortuosité ou la géométrie des embranchements), la segmentation, la classification et la modélisation topologique automatique des vaisseaux rétiniens s'imposent comme une nécessité. Leur objectif est double : réduire le temps d'annotation pour permettre l'analyse de cohortes plus nombreuses et standardiser la mesure des paramètres vasculaires pour comparer leurs valeurs d'une étude à l'autre sans variabilité inter-observateur.

Lorsqu'en 2010, Cheung et al. développent SIVA, le premier logiciel de mesure de l'AVR à intégrer un algorithme de segmentation et de classification de la vasculature rétinienne, ils bénéficient des travaux de recherche déjà conduits sur ces algorithmes depuis plusieurs décennies. En préambule d'une analyse de SIVA et de ses successeurs semi et complètement automatiques, cette section passe en revue l'évolution des algorithmes de segmentation et de classification automatique des vaisseaux rétiniens : depuis les méthodes non supervisées jusqu'aux réseaux de neurones à plusieurs millions de paramètres. Cependant, ce domaine connaît un tel intérêt depuis 30 ans qu'il est impossible de le synthétiser en une revue exhaustive : le nombre de papiers concernant la segmentation seule avoisine les 1700 depuis l'an 2000, dont la moitié a été publiée après le début de mon doctorat en 2019 (cf. Fig. 2.1). Je prends donc le parti d'une revue narrative s'appuyant sur des articles de revues récents [21–26]. Je tenterai néanmoins une approche exhaustive (bien qu'expérimentale) dans le chapitre 7.

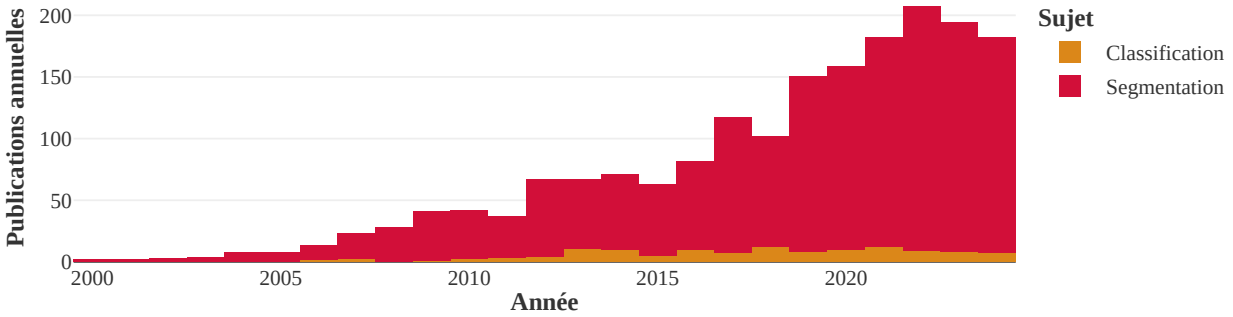


FIGURE 2.1 Nombre de publications annuelles cumulées concernant la segmentation et la classification des vaisseaux rétiniens à partir d'image de fond d'oeil.

Les publications regroupées dans ce graphique ont été identifiées selon la méthode de la section 7.

2.2.1 Défis techniques à la segmentation sémantique des vaisseaux rétiniens

Pour comprendre les méthodes de segmentation et de classification des vaisseaux rétiniens sur images de fond d'œil, commençons par relever les difficultés auxquelles elles doivent répondre.

Singularité morphologique et variabilité des structures vasculaires rétiniennes

La morphologie singulière des vaisseaux, et a fortiori des vaisseaux rétiniens, pose plusieurs défis aux algorithmes de segmentation. Leur structure longiligne et leur présence dans l'ensemble de l'image interdisent les approches par bounding-box ou détection de centroïde. À cause de leur finesse, et bien qu'ils sillonnent la quasi-totalité de l'image, les vaisseaux ne représentent que 15% des pixels de la région d'intérêt : soit un débalancement de classe non négligeable. Cette finesse est en outre hétérogène : une grande diversité de calibres coexiste dans une même image avec un ratio de l'ordre de 1 à 20. Autre défi, la vasculature se déploie en 3D sur la surface rétinienne mais est projetée en 2D par l'imagerie de fond d'œil. Il est donc fréquent qu'un vaisseau soit occulté par un autre lorsqu'ils se croisent ou évoluent parallèlement. Dans ces deux cas, il est parfois difficile d'identifier de quel vaisseau émanent certaines ramifications (voir Figure 2.2c). Pour cette même raison de perspective, les artères et veines de la cupule optique sont difficiles à distinguer car vues de face.

Les algorithmes de segmentation doivent aussi être robustes à la variabilité morphologique de la vasculature rétinienne d'un individu à l'autre. Celle-ci s'applique au calibre moyen dont les variations sont liées à l'âge, au sexe, à l'éthnicité, etc. Hao et al. [27] ont même mesuré une fluctuation du calibre intra-individu au cours du cycle cardiaque (de l'ordre de 6% pour

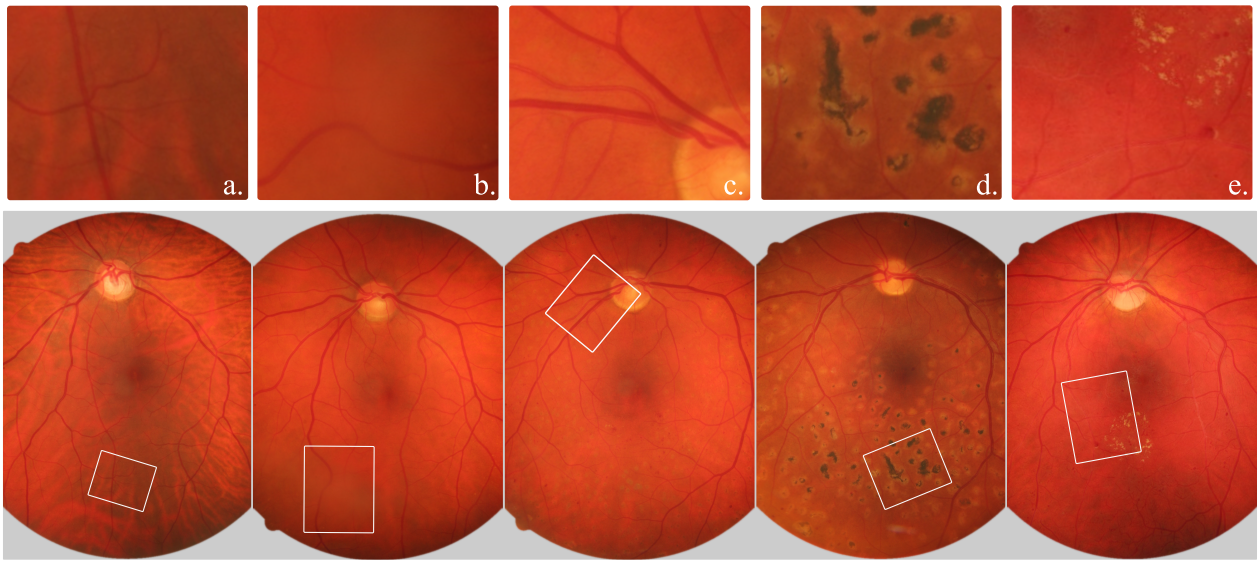


FIGURE 2.2 Exemples de segmentations vasculaires difficiles dans la base de données HRF.
a. & b. Vaisseaux flous et contrastant peu avec le fond ; **c.** Vaisseaux enchevêtrés et de tailles diverses ;
d. & e. Vaisseaux occultés par des lésions de la rétine ou du vitré.

les artérioles et 4% pour les veinules). La tortuosité et la topologie des vaisseaux connaissent aussi une variabilité importante entre individus, au point qu'elle puisse être utilisée comme descripteur pour valider l'identité d'une personne [28] !

Le contraste entre vaisseaux et fond est normalement un atout important pour la segmentation des vaisseaux, comme celui entre artérioles et veinules l'est pour leur classification. Mais si ce contraste est bien marqué près du disque optique et au centre de l'image, il s'atténue en périphérie. En outre, sur certaines images, les vaisseaux de la choroïde (couche vasculaire et nerveuse sous la rétine) transparaissent et induisent des variations importantes de luminosité du fond rétinien, affectant son contraste avec les vaisseaux (cf. Figure 2.2a).

Les lésions du vitré ou de la rétine – particulièrement les hémorragies et les exsudats – peuvent aussi occulter des portions de vaisseaux (cf. Fig. 2.2d & e), altérer leurs couleurs (cf. Fig. 2.2e) ou rendre imprécises leurs bordures (cf. Fig. 2.2b). Tous ces cas sont propices à des faux négatifs dans la segmentation automatique. Par ailleurs, chez certains patients, la membrane externe des vaisseaux reflète, en son sommet, la lumière émise par la caméra. Une ligne blanche apparaît alors au centre du vaisseau et semble le scinder en deux dans le sens de la longueur (cf. Fig. 2.2c). Ce cas se produit plus fréquemment à la surface des artérioles. L'algorithme doit alors être capable de le distinguer du cas où deux vaisseaux sont limitrophes et parallèles : dans le premier, la ligne claire appartient au vaisseau, dans le second, elle appartient au fond.

Artefacts et variabilité d'acquisition

Une mauvaise qualité d'acquisition de l'image de fond d'œil peut aussi être source d'erreurs pour les algorithmes de segmentation et de classification des vaisseaux rétiniens. Ainsi, une mauvaise mise au point rendra flous les contours des petits vaisseaux ; des poussières oubliées sur l'objectif prendront la forme d'une tache sombre sur la rétine, ou un mauvais réglage de la source lumineuse plongera des pans entiers du fond d'œil dans l'ombre.

En outre, il faut ajouter à la variabilité anatomique naturelle de la rétine celle due au dispositif d'imagerie. En effet, depuis la première commercialisation d'une caméra fundus par Zeiss-Littman en 1955, les technologies de photographie du fond d'œil ont beaucoup évolué.

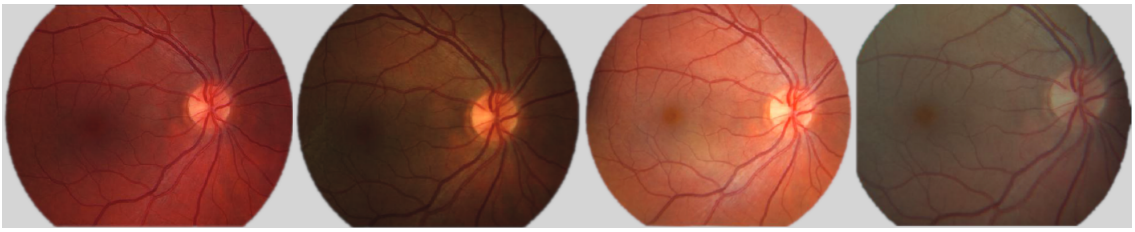


FIGURE 2.3 Colorations d'une même rétine imagée par différentes caméras. *Source : [1]*

Ainsi, l'angle de FOV, c'est-à-dire l'angle correspondant à la section imagée de l'intérieur du globe oculaire, était initialement limité à 25° mais atteint aujourd'hui 133° pour une caméra fundus couleur ou même 200° pour une caméra laser [29]. La résolution des capteurs, facteur déterminant pour distinguer les vaisseaux les plus fins, a aussi grandement progressée : de 500×500 à plus de 3000×3000 pixels. D'une caméra à l'autre, la coloration de l'image peut aussi grandement varier [1], comme démontré par la Figure 2.3 (la couleur de la rétine dépend aussi de la pigmentation rétinienne [30]). Toutes ces variations présentent un défi de généralisation aux algorithmes de segmentation et de classification vasculaire.

Volume limité de données publiques pour entraîner et valider les modèles

Cet enjeu de généralisation est d'autant plus préoccupant que le volume de données publiques annotées pour entraîner et valider les algorithmes de segmentation et de classification est très réduit, en tout cas jusqu'à ces dernières années. Il faut dire que l'annotation manuelle de ces données est particulièrement fastidieuse : la segmentation seule nécessite 2 heures par image [31] ! Aussi, au début de mon doctorat en 2019, ces jeux de données publics totalisaient moins de 200 images pour la segmentation et 107 images pour la classification. Et pour parvenir à ces chiffres, il fallait agréger des images de résolutions et de FOV hétéroclites, certaines centrées sur la macula et d'autres sur le disque optique. Pour cette raison, les algorithmes que nous allons présenter ont souvent été validés sur le jeu qui a servi à leur entraînement, quitte à produire une version de l'algorithme par jeu de données.

Tableau 2.1 Jeux de données publics d'images de fond d'oeil avec annotations vasculaires.
† Ces base de données ne sont pas réellement publiques mais sont disponibles sur demande.

Nom (année)		# Images	Dimensions	Camera (FOV)	Population
Annotées pour la segmentation uniquement					
STARE	(2000 [32])	20	700×650	TopCon TRV-50 (35°)	Californie, US
DRIVE	(2004 [31])	40	565×594	Canon CR5 (45°)	Pays-Bas (25-90 ans)
CHASE_DB1	(2009 [33])	28	1280×960	NM-200-D (30°)	UK (10 ans)
HRF	(2013 [34])	45	3504×2336	Canon CF-60 UVi (60°)	République Tchèque
IOSTAR†	(2015 [35])	30	1024×1024	EasyScan SLO (45°)	-
FIVES	(2022 [36])	800	2048×2048	TRC-NW8 (50°)	Zhejiang, Chine
Annotées pour la segmentation et la classification					
AV-DRIVE	(2013 [37])	40	*voir DRIVE*		
LES-AV	(2018 [38])	22	1958×2196	Visucam ZK-5 (30°/45°)	- (53-90 ans)
HRF (AV)	(2019 [39])	45	*voir HRF*		
RETA	(2022 [40])	81	4288×2848	Kowa VX-10α (50°)	Inde
Leuven-Haifa†	(2024 [41])	220	1444×1444	Visucam 500 (30°)	Belgique (18-90 ans)

La suite de cette section 2.2 est organisée en quatre parties. Les sections 2.2.2 et 2.2.3 traitent des méthodes de segmentation puis de classification automatique des vaisseaux rétiniens n'impliquant pas l'entraînement d'un modèle par apprentissage profond (DL) ; elles visent à décrire la nature de ces tâches par un tour d'horizon de leurs algorithmes conçus manuellement. La section 2.2.4 présente les méthodes qui améliorent la classification A/V par une modélisation topologique des vaisseaux. Enfin, la section 2.2.5 présente les modèles de réseaux de neurones et leurs évolutions spécifiques à la segmentation sémantique de la vasculature rétinienne.

2.2.2 Segmentation des vaisseaux rétiniens pré apprentissage profond

Compte tenu de la faible quantité de données publiques disponibles au début des années 2000, les premières méthodes de segmentation vasculaire sont non supervisées. Elles reposent sur la connaissance a priori de la morphologie des vaisseaux pour les distinguer de l'arrière-plan, en particulier sur leur forme longiligne et leur contraste avec le fond de la rétine. Ce contraste étant particulièrement prononcé sur le canal vert, il sera généralement le seul exploité par une majorité des méthodes citées ci-dessous qui travaillent en niveaux de gris.

Méthodes non supervisées

Filtrage convolutif Dès 1989, Chaudhuri et al. [42] constatent que le profil d'intensité le long d'une section des vaisseaux rétiniens est analogue à une courbe gaussienne. Ils proposent une modélisation par un *matched response filters (MRF)* de taille 32×32 pour détecter les vaisseaux verticaux, puis déclinent ce filtre en 12 variantes pour couvrir toutes les orientations avec un pas de 15°. Cette approche consistant à estimer la "vascularité" de chaque pixel (c'est-à-dire la vraisemblance qu'il appartienne à un vaisseau) par filtrage convolutif fera de nombreux émules. Hoover et al. (2000 [32]) lui ajoutent des caractéristiques extraites sur l'ensemble de l'image. Al Rawi et al. (2007 [43]) affinent les paramètres du filtre pour maximiser la segmentation sur DRIVE et atteignent 94,7% d'exactitude (accuracy) sur cette base. Puis une série de travaux explore d'autres familles de filtres : comme les filtres de Gabor (Meng et al., 2015 [44]), les ondelettes (Wang et al., 2013 [45]) ou les filtres COSFIRE (Azzopardi et al., 2015 [46]). Les bonnes performances de toutes ces méthodes témoignent de l'efficacité du filtrage convolutif pour la segmentation vasculaire. Il sera au cœur de nombreuses méthodes supervisées ou non et on peut même considérer les CNN comme des lointains héritiers. Cependant, l'approche initiale est limitée par son absence de considérations pour la dimension longitudinale des vaisseaux, qui favorise la prédiction des segmentations bruitées et discontinues. Pour répondre à cette limitation, des méthodes par suivi (tracking) de vaisseaux ou par évaluation de la Hessienne virent le jour.

Suivi (tracking) des vaisseaux En 1998, Chutatape et al. [47] combinent les MRF de Chaudhuri et al. avec des filtres de Kalman chargés d'estimer, à partir du dernier pixel connu d'un vaisseau, la position du pixel suivant le long de sa ligne centrale. La procédure débute par l'identification, à la frontière du disque optique, des pixels appartenant aux vaisseaux ; puis elle suit itérativement ces vaisseaux jusqu'à atteindre une terminaison vasculaire. La même année, Tolias et al. [48] proposent une méthode similaire où les MRF sont remplacés par un Fuzzy C-mean Clustering (FCM). Ces approches dites de "suivi" des vaisseaux assurent des résultats topologiquement cohérents, et plusieurs travaux s'en inspireront. Lalonde et al. (2000 [49]) proposeront de plutôt suivre les contours des vaisseaux détectés par un algorithme de Canny. Puis, des variations de l'algorithme de suivi seront explorées : notamment par ajustement d'un modèle physique de vaisseaux (Cree et al., 2005 [50]), par approche bayésienne (Adel et al., 2009 [51]) ou par filtre à particules (Nayebifar et al., 2013 [52]). Mais cette famille de méthodes présente deux défauts : **1.** les embranchements – où plusieurs directions sont "suivables" – sont généralement gérés par des solutions *ad hoc* ; **2.** ces méthodes sont vulnérables aux artefacts d'acquisition ou aux lésions qui causent parfois des discontinuités le long des vaisseaux et interrompent prématurément leur découverte.

Valeurs propres de la Hessienne En 1999, Martinez-Perez et al. (1999 [53]) adaptent au fond d'œil les travaux de Frangi et al. (1998 [54]) sur la segmentation de vaisseaux sur images rayons X, et proposent un nouvel estimateur de "vascularité" : l'écart entre les deux valeurs propres de la Hessienne. En effet, puisque la Hessienne mesure la courbure locale de la variation d'intensité, un fort écart entre ses deux valeurs propres $\lambda_1 << \lambda_2$ indique une structure tubulaire, dont l'orientation est celle du vecteur propre associé à λ_1. Contrairement aux filtres MRF, la Hessienne ne s'appuie pas uniquement sur un profil orthogonal au vaisseau, mais évalue aussi sa stabilité dans la direction tangente. Elle est aussi naturellement équivariante à la rotation et ôte la nécessité de répéter l'analyse dans plusieurs orientations. Vingt ans après, la Hessienne est toujours utilisée pour développer des méthodes non supervisées de segmentation des vaisseaux rétiniens (Zhang et al., 2010 [55] ; Alhussein, 2020 [56]).

Approches multi-échelles Dans leur article initial sur les images rayons X, Frangi et al. notaient que, pour obtenir les meilleurs résultats, l'échelle à laquelle est calculée la Hessienne doit correspondre au calibre du vaisseau analysé. Ils évaluaient donc la Hessienne pour une série d'échelles de sorte à couvrir l'ensemble des calibres présents dans l'image, puis combinaient les cartes de réponses. Cette approche multi-échelle sera une des clés pour permettre la segmentation de tous les vaisseaux rétiniens, petits comme larges. Et elle ne se limite pas au calcul de la Hessienne : en 2002, Gang et al. [57] montrent que les algorithmes MRF

s'améliorent aussi lorsque la taille du filtre s'adapte à celle des vaisseaux rétiniens. En 2006, Sofka et al. [58] développent une normalisation des MRF générés à des échelles variables, afin de choisir automatiquement l'échelle la plus pertinente. La même année, Cai et Chung [59] exploitent une pyramide de Gaussienne multi-échelles pour rendre leur algorithme de "suivi" de vaisseau plus résistant au bruit autour des vaisseaux les plus fins.

Morphologie mathématique En 2001, Zana et Klein [60] proposent de détecter les formes longilignes des vaisseaux rétiniens non pas par "suivi" ou Hessienne, mais par une transformée "top-hat" avec pour éléments structurants une série de lignes de taille fixe et d'orientation variable (avec un pas de 15°), dont le bruit est corrigé par une ouverture appliquée à la reconstruction géodésique de l'image. Cette approche par opérations de morphologie mathématique (morphomat) sera très populaire dans les années 2010, pour sa simplicité d'implémentation et ses capacités de débruitage. Elles seront souvent utilisées en complément des méthodes énoncées plus haut. Ainsi, en 2011, Rossant et al [61] publient une approche exclusivement morphomat consistant en un prétraitement par fermeture, puis une détection par des filtres "top-hat" et "path-opening" ; leur traitement est multi-échelle. De même, en 2013, Nguyen et al. [62] proposent un algorithme basé sur des détecteurs de lignes multi-échelle (Multi-Scale Line Detector). La même année, Budai et al. [63] reprennent l'approche de la Hessienne évaluée à plusieurs échelles, mais y ajoutent une étape de post-traitement par opérations morphomat.

Bien sûr, la taxonomie des méthodes non supervisées présentées ci-dessus est non exhaustive. Il faudrait y ajouter entre autres les méthodes par seuillage local adaptatif (Hoover et al., 2000 [32]), par modèle de contour actif (Zhao et al. 2015 [64]), ou par transformée de Radon (K. Noronha, 2012 [65]). Elle est aussi contestable : à mesure que l'intérêt pour la segmentation des vaisseaux rétiniens monte, les travaux combinent les approches et il n'est plus possible de les ranger dans telle ou telle catégorie. Mais dans le contexte de cette thèse, cette taxonomie révèle les principaux leviers qui sont à notre disposition – ou qu'il faut mettre à la disposition des algorithmes d'apprentissage machine – pour analyser la vasculature rétinienne. En particulier, elle souligne l'importance d'une analyse multi-échelle et qui s'appuie sur la section des vaisseaux (e.g. par un filtrage convolutif) mais aussi sur leur direction tangente (e.g. "suivi", Hessienne ou morphomat) pour limiter le bruit de segmentation.

En 2004, une évaluation comparative des approches de segmentations place Zana et Klein en tête des méthodes non supervisées [66], avec une exactitude de 93.77%. Mais déjà dans cette étude, la première place est remportée par une méthode supervisée.

Méthodes supervisées

L'étude de Neimeijer et al. (2004 [66]), comparant les méthodes non supervisées existantes et proposant une méthode supervisée, est conduite à l'occasion de la publication de leur base de données DRIVE. Ce n'est pas anodin : les méthodes supervisées ont besoin de telles bases pour entraîner leurs classifieurs statistiques à identifier, parmi une série de caractéristiques décrivant l'image, les combinaisons permettant de distinguer si un pixel appartient à un vaisseau ou non. Les publications de STARE en 2000 [32] puis de DRIVE en 2004 [31] ouvrent donc la voie au développement de ces méthodes, dont on donnera ici quelques exemples.

La méthode supervisée de Neimeijer et al. [66] extrait, pour chaque pixel, un vecteur descripteur à 31 dimensions obtenu par un filtrage convolutif avec des filtres Gaussiens multi-échelles et avec leurs dérivées du premier et second ordre. Trois classifieurs sont ensuite comparés pour reconnaître les vecteurs décrivant des pixels de vaisseaux : un classifieur linéaire, un quadratique et un classifieur des k plus proches voisins (k-NN). Dans toutes leurs expériences, le k-NN est le plus précis et surpasse les méthodes non supervisées. Les auteurs expliquent en partie ce gain par l'existence de biais d'annotation non négligeables dans la vérité terrain : pour faire "mieux" que le second annotateur de DRIVE (94.73% d'exactitude) les algorithmes doivent en réalité se spécialiser sur les biais du premier dont les annotations ont été choisies comme étalon (gold-standard). À ce jeu, les méthodes supervisées ont un avantage considérable sur celles non supervisées. Les performances de la méthode de Neimeijer et al. sont améliorées la même année par Staal et al. (2004 [67]) qui divisent l'image en petites régions définies par leur proximité aux lignes de crêtes des structures ressemblant à des vaisseaux, puis ajoutent au descripteur de chaque pixel les caractéristiques extraites de la région à laquelle il appartient.

En 2006, Soares et al. [68] entraînent un classifieur Bayésien à identifier les pixels appartenant aux vaisseaux rétiniens à partir de caractéristiques extraites par des filtres de Gabor. La même année, Ricci et Perfetti [69] implémentent avec succès une segmentation des vaisseaux rétiniens par machines à vecteurs de support (SVM) qui s'appuie sur un vecteur de caractéristiques généré par des filtres détecteurs de lignes. Les performances des SVMs sont significativement améliorées 4 ans plus tard par You et al. [70] qui proposent deux modifications : **1.** les vaisseaux les plus larges sont traités séparément des autres, **2.** la base d'entraînement est augmentée avec des échantillons annotés automatiquement par un SVM, préalablement entraîné sur des images annotées par un humain. Pour ne pas induire de bruit dans l'ensemble d'apprentissage étendu, seuls sont conservés les échantillons où la certitude du SVM est élevée. Cette approche semi-supervisée permet d'augmenter virtuellement le nombre d'échantillons d'entraînement sans le coût d'une annotation manuelle.

En 2010, Lupascu et al. [71] extraient, pour chaque pixel de l'image, un vecteur de 41-D par concaténation d'une bonne partie des descripteurs présentés plus haut : notamment la convolution avec des MRF et des filtres de Gabor ou Gaussiens multi-échelles, les estimateurs de "vascularité" de Frangi [54] et de Staal [67], ou encore les valeurs et vecteurs propres de la Hessienne. En confiant ce vecteur descripteur à un classifieur Adaboost (Feature Based Adaboost Classifier), ils atteignent 95.97% d'exactitude sur DRIVE. Leur méthode restera à la tête du classement sur DRIVE jusqu'à ce qu'elle soit détrônée par les CNN.

2.2.3 Classification traditionnelle des artérioles et veinules rétiniennes

Avant d'étudier les approches de segmentation des vaisseaux rétiniens par réseaux de neurones, intéressons-nous aux méthodes traditionnelles pour leur classification. On l'a vu à la section 2.1, cette étape est cruciale pour permettre une analyse différenciée des artérioles et veinules de la rétine, par exemple pour mesurer l'AVR. À ma connaissance, il n'existe pas d'approche non supervisée à la classification des vaisseaux, et puisque la première base de données publique annotée A/V n'est publiée qu'en 2013, le développement de ces méthodes est plus tardif que celui des méthodes de segmentation. Ce retard s'explique aussi par une attention moindre portée à ces méthodes par la communauté scientifique : toujours à ce jour, le nombre de papiers publiés sur le sujet n'atteint pas le dixième de leurs équivalents en segmentation (cf. Figure 2.1).

Les méthodes de classification A/V traditionnelles (pré-apprentissage profond) suivent toutes à peu près le même fonctionnement, présenté sur la Figure 2.4. Elles débutent par la squelettisation d'une segmentation vasculaire pré-existante ; puis identifient sur la carte du squelette les points de jonction entre branches et les en effacent, faisant ainsi apparaître des segments continus de vaisseau que l'on désignera par le terme : *branche*. L'algorithme de classification vasculaire en lui-même a alors la charge d'étiquetter chaque branche comme artériole ou veinule. C'est cet algorithme qui diffère d'une méthode à l'autre.

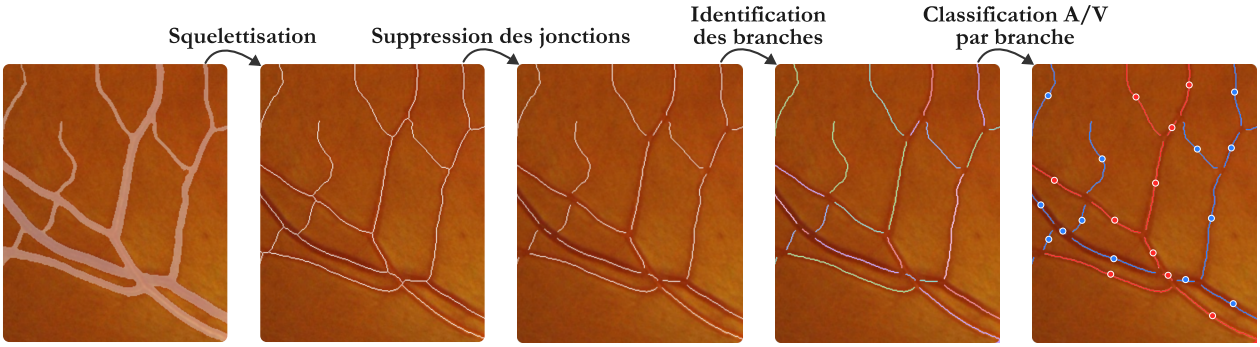


FIGURE 2.4 Vue schématique des étapes précédant la classification des branches vasculaires.

Dans leur review publiée en 2024, Chen et al. [26] recensent 18 articles publiés entre 2003 et 2023 qui approchent la classification des vaisseaux rétiniens par des méthodes de classification traditionnelles pures, c'est-à-dire qui extraient un vecteur de caractéristiques descriptives pour chaque branche puis entraînent un classifieur statistique à distinguer si ce vecteur représente une artériole ou une veinule. Les classifieurs testés sont variés : des méthodes de partitionnement de données (clustering) diffus [72] ou de partitionnement k-means [73, 74], un classifieur bayésien [75], des mixtures de Gaussiennes [76,77], des classifieurs k-NN [78,79]. Après 2013 et la publication de la base annotée AV-DRIVE [37], les classifieurs par analyse discriminante linéaire (LDA) [80–82] et les SVM [83–85] gagnent en popularité.

Concernant les caractéristiques descriptives, les algorithmes proposés utilisent le plus souvent des caractéristiques statistiques (quantiles, moyenne, écart-type) extraites soit directement des intensités le long de la branche [72, 78, 80, 86], soit après avoir converti ces intensités dans un espace de couleur plus judicieux comme LAB ou HSL [73–77, 79, 80, 83, 84]. Quelques articles proposent un choix de caractéristiques plus spécifiques : Neimeijer et al. (2009 [78]) incluent le calibre vasculaire et les angles d'embranchements ; Zamperini et al. (2012 [75]) ajoutent la position du centroïde de la branche relativement au disque optique. D'autres obtiennent aussi de bons résultats avec des descripteurs morphologiques, par exemple en cherchant dans le profil des vaisseaux les reflets lumineux symptomatiques des artérioles [82,83,86] ou par des descripteurs de textures [79]. Pour réduire la complexité algorithmique des classifieurs, certains auteurs diminuent la dimensionnalité du descripteur par PCA (Kondermann, 2007 [83]), ou éliminent ses composantes les moins pertinentes, en les ayant préalablement classées à l'aide d'algorithmes *Random Forest* (Vijayakumar, 2016 [84]), de recherche génétique (Huang, 2018 [81]), ou de *Binary Particule Swarm Optimization* (Irshad, 2021 [85]).

Tous les descripteurs énoncés plus haut sont efficaces pour distinguer le type des plus gros vaisseaux. Néanmoins, le contraste entre artérioles et veinules s'atténuant pour les vaisseaux plus fins, ils sont insuffisants pour classifier les branches secondaires ou tertiaires. D'ailleurs, une majorité des méthodes citées plus haut se contentent de classifier les vaisseaux rétiniens principaux, et souvent uniquement sur leur portion directement mitoyenne au disque optique. Certes, certains articles proposent des stratégies exploitant les connaissances a priori sur le graphe vasculaire rétinien pour améliorer leurs performances : comme Grisan et al. (2003 [72]) qui divisent le disque optique en quadrants pour assurer la détection d'une veine et d'une artère dans chacun, ou Vázquez et al. (2012 [87]) qui ajoutent une étape de "suivi" (tracking) pour regrouper des branches appartenant à un même vaisseau. Mais pour espérer classifier l'ensemble des vaisseaux rétiniens, ces caractéristiques locales ne suffisent pas : il faut tirer profit de la représentation topologique de l'arbre vasculaire.

2.2.4 Modélisation topologique du graphe vasculaire pour la classification A/V

On verra ici comment ont émergé, après l'année 2009, des méthodes de classification A/V s'appuyant sur une modélisation topologique de l'arbre vasculaire rétinien. Mais, en préambule, débutons par un rappel rapide sur la topologie de cet arbre.

Topologie de l'arbre vasculaire rétinien

La topologie de l'arbre vasculaire rétinien est très variable entre individus, mais possède cependant quelques invariants. En règle générale : artérioles et veinules s'alternent sur la surface rétinienne de sorte que les croisements vasculaires impliquent quasi systématiquement des vaisseaux de types différents (l'artériole passera souvent au-dessus de la veinule). Le nombre de vaisseaux sortant du disque optique est variable, mais deux artérioles encadrent toujours la macula (une dans chaque hémisphère supérieur et inférieur).

En théorie, le réseau vasculaire devrait être représenté par un graphe acyclique dirigé, naissant de l'artère centrale et finissant à la veine centrale. La première moitié serait composée des artérioles rétiniennes, le nombre de branches irait donc croissant à mesure qu'elles se ramifient pour laisser place aux capillaires, puis décroîtrait dans la seconde moitié où les capillaires se regroupent en veinules. Mais en pratique, les technologies d'imagerie du fond d'œil n'ont pas aujourd'hui la résolution suffisante pour distinguer les vaisseaux capillaires rétiniens 5 à 10 fois plus fins qu'un cheveu (10 à 15 μm). Par ailleurs, seule la tête de l'artère et de la veine centrales est visible et leurs embranchements dans le disque optique sont particulièrement enchevêtrés. On modélise donc plutôt le graphe vasculaire rétinien par une forêt d'arbres dont les nœuds racines sont placés aux frontières du disque optique et dont les feuilles sont les pointes des vaisseaux à partir desquelles on ne distingue plus leurs ramifications. Dans cette typologie topologique, un *vaisseau* désigne un arbre de cette forêt.

Une telle modélisation topologique en vaisseaux distincts est cependant une représentation avancée du graphe vasculaire rétinien. En pratique, celle obtenue à partir de la segmentation vasculaire pure est bien plus simpliste. Elle consiste en un graphe des *branches* définies à la section précédente (segment continu d'un vaisseau). Dans ce graphe non-orienté, chaque arc représente une branche reliant deux noeuds qui sont de trois types : les *terminaisons* matérialisent les extrémités ou les racines des vaisseaux, elles ne sont connectées qu'à une unique branche ; les *bifurcations* désignent les embranchements où une branche principale donne naissance à deux branches secondaires ; les *croisements* apparaissent lorsque deux vaisseaux s'intersectent, ils relient au moins quatre branches (deux par vaisseaux).

Classification A/V par partition du graphe vasculaire

L'approche topologique pour la classification A/V diffère des méthodes qui consistent à ne pas considérer chaque branche individuellement mais plutôt à propager la classe connue des branches principales vers leurs ramifications difficiles à classifier. Cependant, dans la représentation topologique extraite de la segmentation, artérioles et veinules se croisent, s'entremêlent et se recouvrent mutuellement. Pour propager la classe des branches principales vers les branches secondaires le long de la topologie, il est donc nécessaire de "démêler" les nœuds du graphe afin d'identifier comment sont connectées leurs branches adjacentes. Chaque nœud non terminal doit ainsi être interprété soit comme une bifurcation (si toutes ses branches incidentes appartiennent au même vaisseau), soit comme un croisement (si elles sont issues de vaisseaux différents), soit comme un groupe de terminaisons superposées, soit enfin, comme une combinaison de ces options.

Rothaus et al. (2009 [88]) sont parmi les premiers à proposer une solution à ce problème. Ils interprètent tous les nœuds de rang 3 (avec trois branches incidentes) comme des bifurcations et tous les nœuds de rang 4 comme des croisements diagonaux de vaisseaux. Selon ces hypothèses, ils construisent, pour un graphe vasculaire donné, un système de contraintes binaires où les branches adjacentes à une bifurcation sont considérées d'une même classe et celles adjacentes à un croisement sont de classes opposées. En fournissant alors manuellement les classes A/V de quelques branches, l'algorithme résout le système de contraintes et annote l'ensemble du graphe. Lorsqu'un conflit apparaît entre contraintes, certaines sont relaxées par des heuristiques prédéfinies, et si le conflit persiste, une résolution manuelle est nécessaire.

Cette première approche se limite à séparer les artérioles des veinules, mais les travaux qui suivront s'intéresseront au problème plus général du *partitionnement du graphe vasculaire* (ou *graph tracing*) consistant à isoler les sous-graphes propres à chaque vaisseau (voir Figure 2.5).

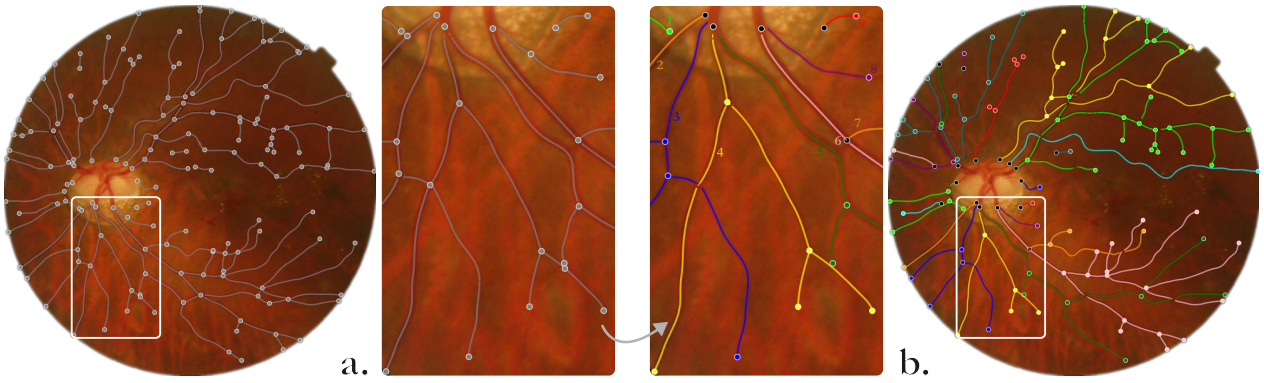


FIGURE 2.5 Partitionnement d'un graphe de branches pour isoler chaque vaisseau.
a. Graphe de branches non-orienté ; **b.** Partition du graphe : chaque vaisseau a sa couleur propre et les nœuds de croisement ont disparu, seuls restent ceux des bifurcations et des terminaisons.

Cette opération se formule comme un problème d'assignation où chaque branche du graphe vasculaire doit être associée à un vaisseau. Suite au partitionnement, les nœuds de croisement sont effacés du graphe de sorte que chaque vaisseau soit représenté par un arbre de branches (voir Figure 2.5b). La classification de l'ensemble du graphe en est alors simplifiée puisqu'elle est réduite à l'attribution d'une classe à chaque vaisseau plutôt qu'à chaque branche (les auteurs proposeront généralement de classer les branches principales de chaque vaisseau par les méthodes supervisées détaillées plus haut). Bienfait supplémentaire, le partitionnement du graphe permet aussi l'analyse de l'arborescence vasculaire et en particulier des bifurcations (seuls nœuds non terminaux restants) dont on peut mesurer l'angle des branches incidentes ou le ratio des calibres. À ce titre, il fait partie de l'arsenal d'algorithmes déployés par les outils semi-automatiques et automatiques qui mesurent les paramètres vasculaires rétiniens.

Partition du graphe par interprétation de la connectivité locale des branches

Une première approche pour partitionner le graphe vasculaire consiste à visiter chaque nœud du graphe pour préciser la connectivité entre ses branches adjacentes : sont-elles toutes connectées entre elles (comme une bifurcation) ou seulement certaines (comme un croisement) et, dans ce dernier cas, lesquelles sont connectées entre elles ? Rothaus et al. (2009 [88]) présupposaient une connexion diagonale des branches adjacentes à un nœud de rang 4. Cependant, ce n'est qu'une possibilité parmi d'autres, comme le montre la Figure 2.6. Pour déterminer la bonne connectivité entre ces branches, il faut s'appuyer sur leurs propriétés géométriques fournies par la segmentation vasculaire.

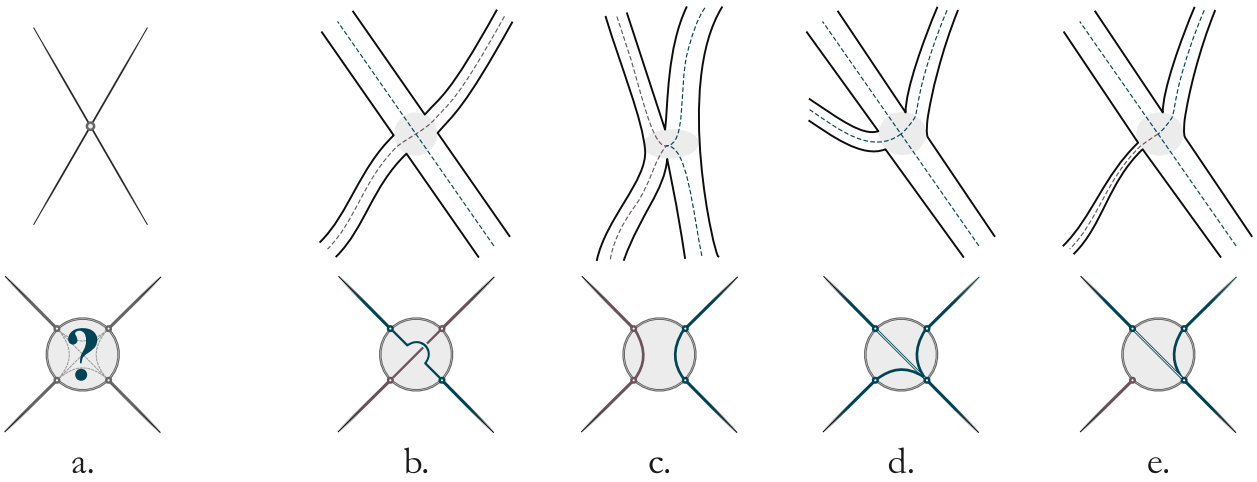


FIGURE 2.6 Résolution de la connectivité locale d'un nœud reliant 4 branches incidentes. **a.** Représentation topologique commune aux quatres configurations suivantes dans lesquelles le nœud représente : **b.** un croisement diagonal de deux vaisseaux ; **c.** un point de contact de deux vaisseaux parallèles ; **d.** deux bifurcations ; **e.** une bifurcation confondue avec une terminaison.

En 2012, Lin et al. [89] estiment la similarité des branches adjacentes à un nœud par un filtre de Kalman étendu. Le filtre se base sur le calibre des branches, sur l'intensité du canal vert, sur la variation de ces deux valeurs à l'abord de la jonction, et sur les angles d'incidence.

En 2014, Dashtbozorg et al. [90] proposent plutôt d'interpréter la connectivité au sein d'un nœud par une série d'heuristiques, chacune spécifique à un type de nœud (i.e. rang 2, 3, 4 ou 5). Par exemple, l'heuristique pour les nœuds de rang 4 détermine s'ils doivent être interprétés comme des bifurcations, des croisements, ou des points de rencontre de terminaison vasculaire et comment relier les branches incidentes en fonction du calibre vasculaire et des angles d'incidence. Cette approche est reprise par Pellegini et al. (2018 [91]) qui remplacent les heuristiques par un algorithme « graph-cut », puis plus récemment par Zhao et al. (2020 [92]) qui les remplacent par une recherche d'ensembles dominants sur le graphe de connectivité des branches.

Partition du graphe par optimisation globale de l'arborescence vasculaire

Parfois, les informations géométriques locales ne suffisent pas à déterminer la connectivité de branches adjacentes à un nœud, et il faut alors exploiter la topologie des nœuds voisins. De et al. (2014 [93], 2016 [94]) proposent de remplacer les heuristiques strictes de Dashtbozorg par des poids continus modélisant la plausibilité d'une connexion entre deux branches. En pratique, ils construisent un *graphe de ligne* : c'est-à-dire un graphe dans lequel une branche vasculaire – un arc dans le graphe standard – est représentée par un nœud dans le graphe de ligne ; et où réciproquement, un nœud du graphe standard est représenté dans le graphe de ligne par plusieurs arcs symbolisant les connexions entre des branches adjacentes (voir Figure 2.7). De et al. pondèrent les arcs du graphe de ligne par un coefficient qui diminue lorsque l'angle entre les deux branches connectées augmente. Puis, ils définissent N vaisseaux (un pour chaque branche racine) et affublent chaque branche d'un vecteur de taille N indiquant la probabilité qu'elle appartienne aux différents vaisseaux. En forçant les vecteurs des

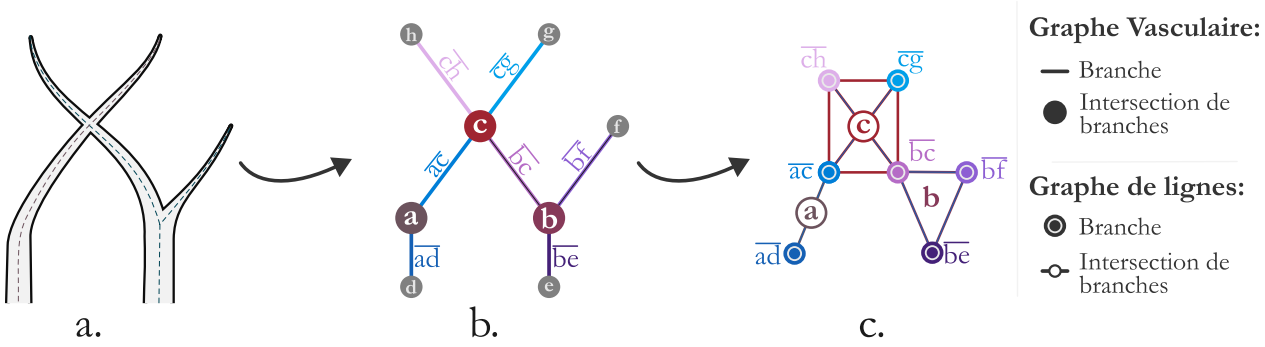


FIGURE 2.7 Construction d'un graphe de ligne à partir de la segmentation vasculaire.

a. Segmentation vasculaire ; **b.** Graphe vasculaire ; **c.** Graphe de ligne.

branches racines à une probabilité de 100% pour leur vaisseau et 0 pour les autres, ils diffusent alors l'influence de chaque vaisseau le long de la topologie du graphe de ligne jusqu'à ce que les vecteurs de probabilité se stabilisent. Les branches sont finalement affectées au vaisseau avec la plus forte probabilité.

En 2015, Estrada et al. [95] proposent une toute autre formulation du problème. Ils postulent l'arborescence vasculaire originale T évoluant en 3D sur la surface rétinienne, dont les nœuds sont uniquement des bifurcations ou des terminaisons (pas de croisement entre vaisseaux) et dont la descendance de chaque branche principale correspond à un unique vaisseau rétinien. Ils formulent alors le graphe vasculaire G tel qu'observé sur l'image de fond d'œil comme une projection P de l'arborescence originale T dans lequel tous les vaisseaux – artérioles comme veinules – se croisent et se confondent. Avec cette formulation, l'objectif de la modélisation topologique consiste à inverser la projection P. Or, comme plusieurs arborescences peuvent avoir la même image par P (cf. Fig. 2.8), Estrada et al. explorent l'ensemble de ces arborescences pour y trouver la plus plausible. Pour ce faire, ils définissent une orientation aux arcs de G afin de produire une arborescence valide, puis ils inversent successivement l'orientation des branches (par ex. Fig. 2.8 b & c) ou altèrent la parenté des branches issues de croisements (par ex. Fig. 2.8 c & d), tout en s'assurant de maintenir la validité globale de l'arbre. Chaque inversion ou altération de parenté génère une nouvelle configuration qu'Es-

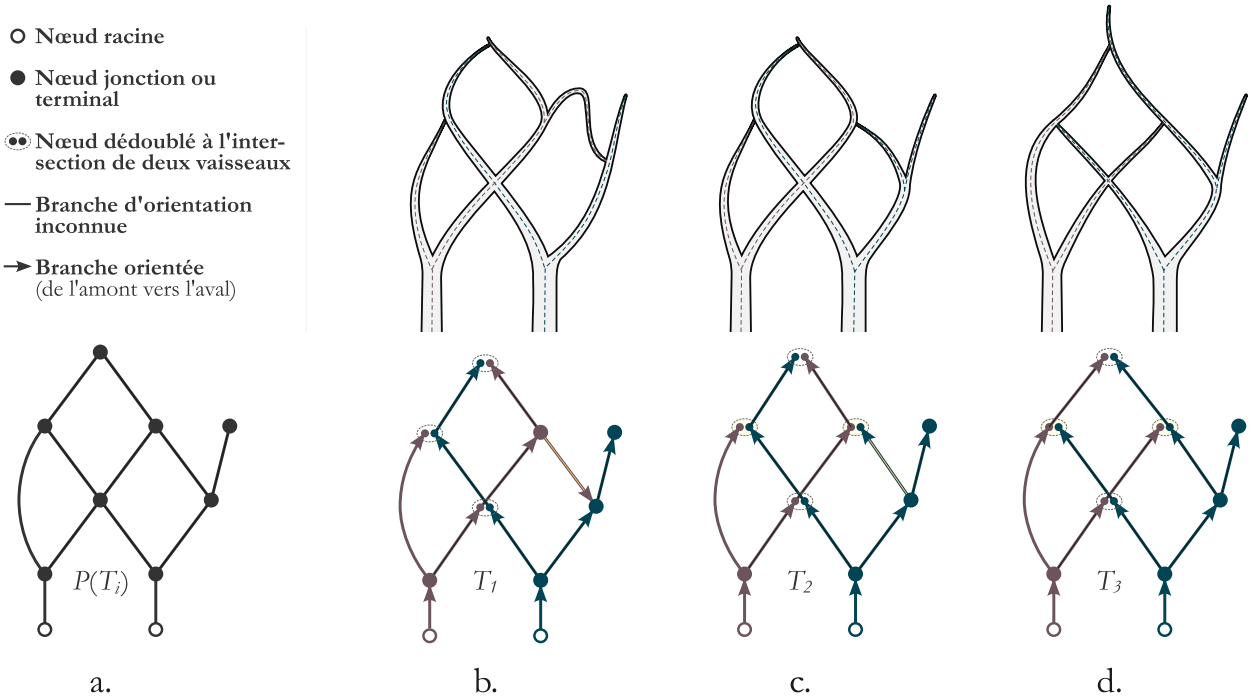


FIGURE 2.8 Variété d'arborescence induisant un même graphe vasculaire non-orienté.
a. Graphe vasculaire non-orienté commun aux trois configurations suivantes. Entre **b.** et **c.** seule l'orientation d'une branche change ; entre **c.** et **d.** seule l'hérédité des deux branches supérieures change.

trada et al. représentent par un nœud dans un méta-graphe d'arborescences. Afin d'identifier la plus plausible, ils calculent un score pour chacune en sommant des mesures de similarité entre branches connectées (selon leurs couleurs, leurs calibres ou leurs orientations). Le méta-graphe est alors exploré par un algorithme de recherche « best-first ».

Ce calcul global de la plausibilité d'une arborescence est un profond changement de paradigme : la connectivité au sein d'une jonction n'est plus simplement interprétée au regard de la similarité de ses branches incidentes, elle intègre aussi la plausibilité des jonctions avales. Dit autrement, puisque la contrainte de validité de l'arborescence lie l'interprétation de certaines jonctions entre elles (le choix d'une configuration pour une jonction pouvant interdire des options pour les autres), l'interprétation évidente de certaines jonctions contribue à l'interprétation de jonctions incertaines. En outre, cette approche permet à Estrada et al. d'inclure des indicateurs généraux de plausibilité comme le taux de croisement incluant à la fois une artériole et une veine qui devrait théoriquement approcher 100 %. Leur méthode se révèle redoutablement efficace, pour la première fois plus de 90 % des pixels du squelette de DRIVE sont correctement classifiés. Elle parvient même à partitionner le graphe vasculaire d'images aquises à grand angle ! Cependant, elle est très sensible aux déconnexions de vaisseaux parfois provoquées par des faux négatifs sur la segmentation. D'ailleurs, ils évaluent leur algorithme sur les segmentations manuelles de DRIVE et sont même contraints d'exclure une image où trop de vaisseaux sont déconnectés du disque optique.

Dans leur méthode, Dashtbozorg et al. prétraitaient le graphe pour y corriger les artefacts de déconnexion (faux négatifs), surconnexion (faux positifs) ou de recouvrement (lorsque les segmentations de deux vaisseaux distincts mais adjacents se superposent et se confondent de sorte que, dans le graphe, un unique arc représente une portion de l'un et l'autre). Cependant, comme pour la résolution de la connectivité des nœuds, leur prétraitement consiste en une série d'heuristiques strictes et locales. En 2013, Lau et al. [96] avaient proposé une méthode d'énumération d'arborescences et d'estimation probabiliste de leur plausibilité qui tenait compte des artefacts de recouvrement vasculaire, mais uniquement de ceux-là. En outre, leur méthode ne traitait que des vaisseaux compris dans la région entre .5 et 2 diamètres du disque optique autour de la cupule. À ce jour, le problème de partitionnement du graphe vasculaire rétinien par une approche d'optimisation globale de l'arborescence mais robuste aux artefacts de segmentation reste irrésolu.

2.2.5 Segmentation sémantique des vaisseaux rétiniens par réseaux de neurones

Nous avons résumé dans les sections précédentes les évolutions de la littérature concernant la segmentation et la classification du réseau vasculaire rétinien par des approches non supervisées ou supervisées traditionnelles. Toutes ces approches s'appuient sur des caractéristiques descriptives choisies manuellement par les auteurs. Ce paradigme change avec l'essor des approches par réseaux de neurones, capables d'apprendre automatiquement à extraire les caractéristiques pertinentes pour résoudre la tâche qui leur est soumise. À partir du début des années 2010, ces modèles ont rapidement conquis l'intérêt des chercheurs travaillant dans le domaine biomédical en général et sur la segmentation vasculaire en particulier.

La révolution de l'*apprentissage profond* et des réseaux de neurones convolutifs

Les *réseaux de neurones entrainés par apprentissage profond* sont des algorithmes conçus pour accomplir une tâche de traitement de signal par une modélisation statistique du lien entre les signaux d'entrée et une consigne souhaitée. Contrairement aux méthodes d'apprentissage traditionnelles, cette modélisation considère le signal d'entrée tel quel (et non des caractéristiques choisies manuellement pour le représenter) et elle s'échelonne en une série de représentations successives de plus en plus abstraites. On ne détaillera pas ici le formalisme théorique qui a déjà été expliqué maintes fois et mieux que je ne pourrais le faire (notamment par Goodfellow, Bengio et Courville, 2016 [97]). On rappellera simplement que chaque représentation est constituée d'un vecteur de caractéristiques dérivé de la représentation précédente par *une couche* de neurones artificiels. Ces neurones sont définis par une fonction mathématique statique alliant une combinaison linéaire suivie d'une transformation non linéaire, ainsi que par des *poids* ajustables qui paramètrent la réponse du neurone. Lors de la phase d'entraînement du modèle, ses poids sont mis à jour itérativement par *descente de gradient* de sorte à minimiser une *fonction de coût* mesurant un taux d'erreur. De tels modèles sont théoriquement capables de modéliser n'importe quelle transformation, sous réserve que le nombre de neurones et d'échantillons d'entraînement soit suffisant.

Cette formulation en réseaux de neurones présente plusieurs avantages comparés aux méthodes d'apprentissage dites traditionnelles. D'abord, sa grande versatilité : régression, classification, segmentation, reconstruction, appliquée à du texte, des images ou des sons, sa diversité d'applications est telle qu'elle semble dotée d'ubiquité ! Et puisque la théorie est commune à toutes, les innovations conçues par exemple pour l'analyse de texte peuvent être transposables à la segmentation d'images. Ensuite, l'extraction des caractéristiques est totalement automatique. De ce fait, le réseau ne considère que les données de l'ensemble d'entraînement et non une connaissance a priori du phénomène étudié, ce qui est généralement

vu comme un atout pour ne pas risquer de biaiser l'interprétation des données [2]. Enfin, les caractéristiques ainsi extraites possèdent généralement de bien meilleures capacités de généralisation. Ces capacités sont telles qu'il est notamment possible de raccourcir grandement le temps d'entraînement d'un modèle en utilisant les poids d'un réseau à l'architecture similaire mais entraîné sur un autre jeu d'entraînement plutôt qu'en les initialisant aléatoirement (par ex. en pré-entraînant un réseau pour ensuite l'affiner sur des images médicales). Pour toutes ces raisons, les réseaux de neurones artificiels entraînés par apprentissage profond sont devenus incontournables dans les conférences et journaux d'interprétation d'images médicales.

Ce formalisme n'est pas récent, dès la fin des années 90, Yoshua Bengio et Yann LeCun (1998 [98]) en avaient proposé une variante pour la reconnaissance d'images (en l'occurrence des chiffres) qu'ils nomment CNN et dont chaque neurone apprend un filtre de convolution. Les couches de neurones convolutifs sont alternées avec des *couches de sous-échantillonnage* qui réduisent la résolution des représentations mais augmentent leur nombre de canaux. Les dernières couches sont *complètement connectées* : elles réduisent par combinaison linéaire la représentation de l'image dotée de la plus faible résolution (et donc du plus grand nombre de canaux) à un vecteur de caractéristiques puis en la prédiction finale. Au début des années 2010, suite aux avancées incrémentales du cadre théorique des réseaux de neurones profonds et à l'explosion des capacités de calcul, la capacité de généralisation de ces modèles éclate au grand jour et révolutionne le domaine de l'apprentissage machine. Ainsi, en 2012, les CNN dépassent largement les performances des autres méthodes d'apprentissage sur la classification des images du challenge ImageNet (Krizhevsky, 2012 [99]).

Deux ans plus tard, Melinscak et al. [100] sont parmi les premiers à porter cette approche à la segmentation des vaisseaux rétiniens. Leur modèle CNN contient 125 k paramètres organisés en 4 couches convolutives et deux couches complètement connectées et atteint les performances prometteuses de 94.7% d'exactitude sur DRIVE. Depuis, la majorité des travaux de recherche en segmentation ou en classification des vaisseaux rétiniens s'intéresse à ces algorithmes : soit en développant des innovations spécifiques à cette tâche, soit en important des avancées de la technologie développées pour des approches plus généralistes. Deux avenues de recherche sont principalement explorées : **1.** le choix d'architecture des réseaux et **2.** leur procédure d'entraînement.

2. On pourrait cependant nuancer cette affirmation : considérer que les données sont «la vérité» stricte, c'est négliger tous leurs biais intrinsèques. De plus, les connaissances a priori sur un phénomène sont souvent de bonnes pistes de recherche pour imaginer des régularisations de la fonction de coût et améliorer la généralisation du modèle.

Évolutions Architecturales

Depuis 2015, les architectures de réseaux de neurones spécialisés dans le traitement d'images et a fortiori dans la segmentation et la classification des vaisseaux rétiniens ont connu de nombreuses évolutions. Il y eut initialement quelques tentatives d'approches non convolutives, comme Li et al. (2016 [101]) qui entraînèrent cinq couches complètement connectées à prédire la carte de segmentation d'un patch de 16 par 16. Mais rapidement, dans la lignée de Melinscak et al., les modèles convolutifs s'imposent.

CNN Les réseaux de neurones convolutifs ont en effet un avantage sur les modèles complètement connectés : dans une couche convolutive, les filtres sont convolués sur toute l'image, là où un modèle complètement connecté associerait un poids à un pixel. Les CNN sont ainsi non seulement plus économiques en paramètres, mais de plus, chaque paramètre bénéficie de l'information de l'ensemble des pixels de l'image. Ou plus exactement l'ensemble des pixels du *patch* : En effet, pour savoir si un pixel est un vaisseau ou non, on présente au modèle un patch de son voisinage dans l'image. Le modèle de Melinscak et al. analyse ainsi des patchs de 64×64 pixels. En 2016, Likowscki et Krawiec [102] améliorent leur travail en appliquant une augmentation de données aléatoire géométrique (rotation, symétrie et redimensionnement) et colorimétrique (teinte et saturation). L'année suivante, plusieurs travaux (Girard et al. 2017 [103] ; Welikala et al., 2017 [104]) proposent d'augmenter le nombre de neurones de la dernière couche d'un à deux. Par cette simple modification, ils convertissent le réseau de segmentation en un modèle de segmentation *sémantique*, c'est-à-dire qui réalise simultanément la segmentation et la classification des artérioles et veinules rétiniennes. Un défaut des modèles CNN demeure cependant : initialement conçus pour la classification d'images, ils nécessitent que chaque pixel soit analysé individuellement pour obtenir la carte de segmentation complète. Cette limitation motive la conception de modèle réseaux de neurones complètement convolutif (FCN).

FCN Proposé pour la première fois par Long et al. (2015 [105]), les FCN troquent les dernières couches complètement connectées des CNN par des couches convolutives avec des filtres 1×1. Par ce subterfuge, la sortie vectorielle des CNN devient matricielle et le modèle prédit dorénavant une image complète. Il est aussi équivariant par translation : une translation de son image d'entrée induit une translation équivalente de sa prédiction, propriété qui est parfaitement adaptée pour les tâches de segmentation. À cause des couches de sous-échantillonnage, la prédiction du modèle est cependant sous-résolue par rapport à l'image d'entrée. Long et al. proposent donc de suréchantillonner (par convolution transposée) à la

résolution de l'image, les cartes produites à différentes étapes du réseau, avant de les combiner pour obtenir la prédiction finale. Cette méthode est adaptée à la segmentation des vaisseaux rétiniens par Fu et al. (2016 [106]) avec leur architecture Deep Vessels dans laquelle les prédictions aux multiples résolutions sont combinées par un champs aléatoire conditionel (CRF). Mais les FCNs devront attendre 2015 pour exprimer leur plein potentiel de segmentation à travers l'architecture U-Net.

U-Net Cette architecture proposée en 2015 par Ronneberger et al. [107] contient deux parties. La première moitié : l'encodeur, est similaire en tout point à un CNN dont on aurait enlevé les couches finales de classification (il ne sera d'ailleurs pas rare de la remplacer par un CNN pré-entraîné sur ImageNet). La seconde : le décodeur, est une réplique par symétrie de l'encodeur ; il réduit progressivement le nombre de caractéristiques par couche tout en augmentant leur résolution jusqu'à atteindre celle de l'image. Encodeur et décodeur sont reliés à chaque résolution par une "skip-connection" qui concatène les prédictions du premier à celles fraîchement suréchantillonnées du second. On le verra dans l'étude exhaustive de la littérature (chapitre 7), cette architecture sera la plus plébiscitée par les publications traitant de segmentation sémantique des vaisseaux rétiniens. En 2019, Hemelings et al. [39] l'étudient en détail sur cette tâche : ils réduisent le nombre de caractéristiques prédites par chaque couche pour éviter le surapprentissage dans ce contexte de volume réduit de données d'entraînement. Pour la même raison, ils montrent un effet bénéfique de l'utilisation de dropout entre les couches les plus profondes. Surtout, ils mettent en évidence l'importance de la taille du champ d'activation [3] dans la classification vasculaire. Ils proposent de l'augmenter en choisissant des tailles de filtres convolutifs 5×5 et des convolutions diluées. Les bénéfices des convolutions diluées furent par la suite approfondis par Biswas et al. (2020 [108]) ou Hussain et al. (2022 [109]). D'autres auteurs proposent de substituer chaque couche convolutive par un bloc qui en comprend trois avec des filtres de taille 1×1, 3×3 et 5×5 (D. Yang et al., 2020 [110] ; X. Yang et al. 2022 [111]). Des modèles plus denses sont aussi proposés, renforçant le nombre de connexions entre les couches convolutives qui travaillent à différentes échelles. C'est le cas du FR-UNet de Lio et al. (2022 [112]) dont on trouve une implémentation publiée avec ses paramètres entraînés dans le paquet python Fundus-Image-Toolbox (Gervelmeyer, 2025 [113]). Toutes ces publications tentent d'améliorer la cohérence des prédictions entre les gros et les petits vaisseaux en amplifiant les capacités d'analyse multi-résolution déjà bien présentes dans les U-Net. De ce point de vue, elles s'inscrivent dans la lignée des approches multi-échelles imaginées 10 à 20 ans plus tôt pour segmenter ces mêmes vaisseaux avec des

3. Le champ d'activation est défini comme la taille de la fenêtre dans l'image de départ qui peut théoriquement contribuer au calcul d'un pixel dans celle prédite, ou dit autrement, c'est la taille cumulée obtenue par composition de tous les filtres de convolution.

algorithmes non supervisés. J'ai moi-même contribué à ce mouvement pendant ma maîtrise en 2018 à l'occasion d'une publication (Lepetit-Aimon et al., 2018 [114]) dans laquelle j'adjoignais au U-Net classique, une branche travaillant sur l'image sous-résolue.

Bloc d'attention Malgré le recours à de multiples résolutions de travail, la segmentation sémantique des vaisseaux les plus fins reste un défi pour les U-Net. Une interprétation est alors formulée : pour segmenter les vaisseaux indépendamment de leur taille, la multiplicité de résolutions ne suffit pas, encore faut-il les « aiguiller » de sorte que le modèle s'appuie sur les résolutions basses pour les vaisseaux larges, et celles élevées pour les vaisseaux fins. Zhang et al. (2019 [115]) proposent d'implémenter cet « aiguillage » en insérant des *modules d'attention* dans un U-Net, là où les caractéristiques de l'encodeur sont concaténées à celles, plus basse résolution, du décodeur. L'idée d'attention avait été développée pour les CNN par Woo et al. (2018 [116]) : elle consiste à « désactiver » certains neurones d'une couche convolutive en multipliant leur réponse par un masque. Ce masque est calculé par une branche parallèle qui analyse les entrées de la couche pour déterminer quel neurone doit être désactivé. Une fonction sigmoïde à la fin de cette branche assure que sa valeur soit comprise entre 0 et 1 (tel un masque). Les neurones ainsi shuntés peuvent, réciproquement, se spécialiser à l'analyse de certains objets, par exemple les plus petits vaisseaux. Woo et al. avaient proposé de combiner deux formes d'attention dans les CNN : une attention par canal et une attention spatiale. En 2019, Mou et al. [117] exploitent précisément cette idée d'*attention duale* pour la segmentation des vaisseaux rétiniens. Puis en 2021, Li et al. [118] montrent qu'en intégrant des modules d'attention au décodeur du U-Net, il est possible d'atteindre les performances de l'état de l'art avec un modèle économe en paramètres et ne travaillant qu'à 3 résolutions (comparés aux 5 des U-Net standard).

Transformeurs Comme on l'a vu à la section précédente, la classification des vaisseaux éloignés du disque optique est particulièrement délicate sans s'appuyer sur une approche topologique. Or le champ d'attention limité des FCN leur interdit l'apprentissage d'une telle approche : lorsqu'ils classifient les pixels en périphérie de l'image, le disque optique n'est tout simplement plus visible pour ces modèles. Un problème similaire affectait les réseaux de neurones analysant le langage naturel, et trouva sa résolution avec l'apparition des modèles dits "transformeurs". Ces modèles appréhendent la tâche bien différemment de leurs prédécesseurs : plutôt que d'apprendre à reconnaître des séquences de mots, ils modélisent l'interaction de chaque mot avec ses voisins plus ou moins proches : au sein d'une phrase, d'un texte, ou aujourd'hui d'un roman (GPT-4 peut analyser un contexte d'environ 96 000 mots). Je n'expliquerai pas ici leur fonctionnement, hormis qu'ils reposent sur un mécanisme

d'attention croisée (*self attention*) leur permettant de modéliser l'interaction de mots séparés d'une distance arbitrairement longue et uniquement limitée par l'espace mémoire à disposition. Cette méthode fut proposée pour l'analyse du langage en 2017 par Vaswani et al. dans leur article : « Attention is all you need » [119], puis transposée à l'analyse d'images 4 ans plus tard par Dosovitskiy et al. : « An image is worth 16x16 words » [120] qui introduit les *vision transformers* (ViT). La capacité de cette architecture à modéliser des interactions longues distances semble évidemment un atout pour la segmentation sémantique des vaisseaux rétiniens, et rapidement des auteurs les intègrent à leurs modèles : soit par des couches de *self-attention* placées sur les *skip-connection* du U-Net (Zhang et al., 2022 [121] ; Yu et al., 2022 [122]) ou entre ses couches convolutives (Yi et al., 2023 [123]) ; soit en remplaçant toutes les couches de convolutions par des modules transformeurs *Swin* (Li et al., 2024 [124]).

On pourrait ajouter à cette liste des principaux « courants architecturaux » : les approches *squeeze and excitation* (Shen et al., 2022 [125]) ou les réseaux de neurones par graphe (GNN) (Mishra et al., 2021 [126]). Mais la frise chronologique présentée dans cette section permet déjà de distinguer les tendances de la littérature. Les réseaux de neurones convolutifs semblent être particulièrement adaptés à la segmentation et à la classification des vaisseaux, probablement grâce aux aptitudes naturelles des approches convolutives déjà mises en évidence lors du développement d'algorithmes non supervisés. Parmi les modèles convolutifs, une majorité a en commun l'architecture U-Net, dont les multiples résolutions semblent particulièrement bénéfiques pour analyser tous les vaisseaux, quel que soit leur calibre. Les approches récentes par transformeurs améliorent marginalement les performances de segmentation et de classification, mais au prix d'une importante hausse du nombre de paramètres. Ces architectures, particulièrement gourmandes en données d'entraînement, semblent largement bridées par le faible volume de jeux de données annotées disponibles.

Notons que ces «courants architecturaux» n'évoluent pas dans des couloirs isolés. Au contraire, la littérature fourmille d'articles piochant dans l'un ou dans l'autre, combinant et recombinant leurs idées pour former des architectures inédites. Les baptiser devient alors un exercice quasi poétique, voire ésotérique pour peu que tous les courants aient été réunis dans une même architecture, comme pour ce papier sobrement intitulé : *Global Transformer and Dual Local Attention Network via Deep-Shallow Hierarchical Feature Fusion* [127]. Plus sérieusement, je ne peux m'empêcher de constater une certaine hubris quant à la complexité des approches proposées qui semblent parfois moins être motivées par une intuition scientifique que par la composition de mots-clés. En 2022, Galdran et al. [128] avaient déjà noté la forte expansion paramétrique et computationnelle des modèles proposés, sans nécessairement induire une hausse équivalente des performances. Ils avaient alors montré qu'un modèle simple

composé de deux U-Net placés l'un derrière l'autre (baptisé W-Net) égalait ces architectures en segmentation comme en classification, sous réserve qu'il soit entraîné correctement.

Ce foisonnement de la littérature pose aussi un problème de validation. Les propositions de modèles énoncées plus haut sont souvent motivées par l'intuition d'une adéquation entre une formulation de réseau et la tâche visée, mais cette intuition est rarement interrogée puis confirmée ou infirmée. Certes, les papiers contiennent régulièrement une étude ablative pour démontrer l'intérêt de l'intégration successive des idées qu'ils proposent. Mais les réseaux de neurones, dont l'entraînement peut produire des résultats hautement variables, présentent un vrai défi de répétabilité et de comparaison entre papiers. En effet, comment savoir si les fluctuations de performance entre deux architectures s'expliquent bel et bien par les propriétés du modèle proposé, par une synergie nouvelle engendrée par une composition judicieuse de modules, ou si elles sont simplement dues à la hausse du nombre de paramètres, à un meilleur choix d'hyperparamètres, de jeu de données ou de procédure d'entraînement ? Gardons cette question en tête, on y reviendra...

Évolution de la procédure d'entraînement

L'amélioration des procédures d'entraînement est abordée par la littérature en parallèle des recherches d'architectures. Et si elle est éclipsée par ces dernières en termes de nombre de papiers, elle propose tout de même quelques idées qui méritent d'être mentionnées ici.

Approches Génératives On l'a vu, le manque de données annotées pour l'entraînement pose une difficulté majeure pour le développement de réseaux de neurones fiables. Pour atténuer cette difficulté, des travaux exploitent des modèles génératifs pour créer de fausses images de fond d'œil à partir d'une segmentation vasculaire connue. Cette approche est d'abord mise en œuvre avec un réseau adversarial génératif (GAN) (Costa et al., 2018 [129] ; Lahiri et al. 2020 [130] ; Andreini et al., 2022 [131]), puis plus récemment avec un modèle de diffusion (Go et al., 2024 [132]). Mais ces méthodes présentent un risque d'introduire des hallucinations dans l'ensemble d'entraînement. Une autre catégorie d'approches exploite l'entraînement adversarial différemment : le générateur est entraîné à prédire la segmentation de vaisseaux sur des images non annotées par un jeu adversarial avec un modèle discriminateur qui tente d'identifier dans deux couples images de fond d'œil/segmentation vasculaire lequel provient du jeu d'entraînement annoté et lequel a été produit par le générateur. (Wu et al., 2019 [133] ; Kamran et al., 2021 [134]) . Dans les deux cas, ces méthodes permettent de tirer profit des nombreuses images de fond d'œil publiques non annotées, mais offrent des gains en performances généralement limités.

Métriques de validation Depuis les premiers algorithmes de segmentation des vaisseaux rétiniens, la métrique de validation a toujours été l'exactitude [4]. Mais alors que les algorithmes s'améliorent, cette métrique est de moins en moins pertinente pour les évaluer. **1.** Elle donne trop d'importance aux pixels du fond (qui constituent 80 à 90 % de la zone d'intérêt) de sorte qu'un algorithme ne détectant aucun vaisseau atteint une exactitude moyenne avoisinant les 85%... Pour cette raison, l'indice Dice qui ignore les pixels vrai-négatifs [5] gagne en popularité depuis une dizaine d'années pour évaluer les méthodes de segmentation vasculaire : à ma connaissance, il est utilisé pour la première fois comme tel en 2015 (Mukherjee et al., 2015 [135]). **2.** L'exactitude donne relativement moins d'importance à la détection des petits vaisseaux qu'à ceux plus larges (qui sont composés de plus de pixels) et elle est très sensible aux variations de contour des vaisseaux qui ne sont pourtant pas toujours bien définis. Afin de mieux évaluer la capacité d'un modèle à détecter des vaisseaux quels que soient leur taille ou contour, Shit et al. proposent en 2021 [136] une variation du dice qui ne tient compte que des pixels appartenant au squelette vasculaire : le *center-line dice* ou CL Dice. On trouve une métrique analogue mais sous un autre nom dans un article publié 9 ans plus tôt par Gegundez et al. : *A Function for Quality Evaluation of Retinal Vessel Segmentations* (2012 [137]). Dans ce papier, ils proposent aussi deux autres métriques : l'une similaire au score Dice mais intégrant une tolérance à la variation des contours vasculaires, et l'autre mesurant la similarité de connectivité par décompte des composantes connectées de la segmentation. **3.** Cette évaluation de la connectivité ou plus largement de la similarité topologique est l'angle mort ultime de l'exactitude et des métriques mesurant des taux de pixels. En effet, quelques pixels absents d'une segmentation n'auront que très peu d'impact sur la valeur de ces métriques, quand bien même ces pixels déconnecteraient une branche du reste du graphe vasculaire en transformant complètement sa topologie. En 2021, Araujo et al. [138] proposent de mesurer la similarité topologique en évaluant combien de paires de pixels connectés [6] dans la vérité terrain le sont aussi dans la prédiction. Cependant, leur article n'est pas publié dans une revue et sera peu repris. Tout récemment, Berger et al. (2024 [139]) ont commis : *Pitfalls of topology-aware image segmentation*. Dans cette prépublication ArXiv, ils rapprochent trois applications de segmentation sensibles à la topologie : la segmentation des vaisseaux sur image de fond d'œil, des routes sur image satellite et des neurones sur images du cerveau captées par microscope électronique. Ils identifient des lacunes communes aux validations

4. L'éxactitude (*accuracy* en anglais) est définie par : $\text{acc} = \dfrac{TP + TN}{TP + TN + FP + FN}$ où TP et TN sont les vrais positifs et négatifs et où FP et FN sont les faux positifs et négatifs.

5. L'indice Dice-Sørensen, analogue au score de classification F1, est défini par : $\text{dice} = \dfrac{2TP}{2TP + FP + FN}$.

6. Araujo et al. considèrent une paire de pixels connectés s'il existe un chemin continu de pixels appartenant à la segmentation pour les relier.

topologiques de ces méthodes : notamment que leurs jeux de données, conçus pour la segmentation, contiennent des erreurs impactant fortement les métriques topologiques (par ex. quelques pixels manquants à la segmentation de vaisseaux dans DRIVE). Ils montrent aussi un manque de maturité des métriques topologiques (nombre de Betti, erreurs d'appariement de Betti [140]) qui indiquent parfois des niveaux de performances contradictoires.

Fonctions de coût Les problèmes des métriques de validation identifiés plus haut se posent de manière analogue aux fonctions de coût guidant l'apprentissage des réseaux de neurones ; avec une contrainte supplémentaire cependant : ces dernières doivent être différentiables. Pour la segmentation sémantique, la fonction de coût généralement privilégiée est l'*entropie croisée* qui donne une bonne approximation différentiable de l'exactitude, mais souffre des mêmes lacunes. Plusieurs recherches explorent donc d'autres options. En 2016, Milletary et al. [141] proposent une formulation différentiable du score Dice : le *soft-Dice*, dans un contexte de segmentation volumétrique sur image IRM. Cette fonction de coût sera reprise deux ans plus tard pour la segmentation des vaisseaux rétiniens (Soomro et al., 2018 [142]). D'autres auteurs proposent de conserver l'entropie croisée mais en la pondérant en fonction de la difficulté des régions (Wang et al., 2020 [143]). Dans leur article proposant le CL-Dice, Shit et al. (2021 [136]) formulent une fonction de coût correspondante en s'appuyant sur une version différentiable de la squelettisation. Leur article se limite à la segmentation vasculaire, mais sera étendu à la classification par Jian et al. (2024 [144]). Enfin, en 2023, Mao et al. [145] appliquent à la segmentation vasculaire la *topo-loss* : une fonction de coût proposée quatre ans plus tôt par Hu et al. (2019 [146]) qui exploite des diagrammes de persistance (*persistance homology diagram*) pour pénaliser les variations du nombre de composantes connectées ou d'anses topologiques lorsqu'on fait évoluer le seuil de segmentation appliqué à la prédiction. Malheureusement, en l'absence d'une validation topologique standardisée, il est difficile de distinguer les bienfaits de ces méthodes comparées au soft-Dice ou à l'entropie croisée.

Limitations des réseaux de neurones

Malgré les centaines d'articles sur la segmentation et la classification des vaisseaux rétiniens par réseaux de neurones, certaines lacunes de ces modèles semblent persister inexorablement.

Incohérence topologique L'exactitude en segmentation des U-Net et de leurs successeurs est particulièrement élevée, les erreurs qui subsistent concernent donc un petit nombre de pixels. Dans leur récente revue des algorithmes de segmentation des vaisseaux rétiniens, Li et al. (2025 [25]) identifient deux types de lacunes communes aux réseaux de neurones : l'omission des vaisseaux les plus fins et la discontinuité dans la segmentation des vaisseaux, en

particulier en cas de faible contraste, d'occlusion par une lésion ou proche des croisements. Concernant la classification, le taux d'erreur est plus important. La revue de Chen et al. (2024 [26]) l'attribue à un déclin important des performances pour les petits vaisseaux et aux changements erronés mais récurrents de la classe prédite au sein d'une branche. Toutes ces erreurs de segmentation et de classification semblent pointer vers une cause commune : la méconnaissance de la structure topologique sous-jacente au graphe vasculaire rétinien. Au fond, les réseaux de neurones convolutifs sont comparables aux méthodes qui les ont précédés et qui s'appuyaient sur une modélisation morphologique locale des vaisseaux. Certes, ils en démultiplient le potentiel en accroissant considérablement la complexité et le nombre de paramètres, mais ils en conservent aussi les lacunes topologiques fondamentales. Chen et al. (2024 [26]) notent d'ailleurs que les rares approches alliant modèles convolutifs et analyse du graphe obtenaient généralement les meilleurs résultats en classification. Ils recommandent donc le développement d'algorithmes de post-traitement des cartes de segmentation sémantique prédites par réseaux de neurones pour en assurer la cohérence topologique.

Lacune de généralisation Les réseaux de neurones sont généralement reconnus pour leur bonne capacité de généralisation, à la condition qu'une variété suffisante de données d'entraînement leur soit disponible. Or, en analyse d'images médicales – particulièrement lorsqu'il est question de segmentation sémantique – les données annotées sont rares. Dans ce domaine, on connaît donc plutôt les réseaux de neurones pour leur propension au surapprentissage (*over-fitting*) et à l'apprentissage des biais de données. Les modèles de segmentation et de classification des vaisseaux rétiniens y sont particulièrement vulnérables étant donné le faible nombre de jeux de données spécialisés pour ces tâches. Li et al. (2025 [25]) soulignent d'ailleurs dans leur revue que ces données sont insuffisantes pour capturer la variabilité et la complexité inhérentes aux situations cliniques réelles. Pour en avoir entraîné pendant plusieurs années, je peux moi-même attester que ces modèles sont plutôt capricieux : un simple changement de caméra peut faire effondrer les performances d'un modèle même si les nouvelles images sont objectivement de meilleure qualité ! Cette spécialisation au jeu d'entraînement est particulièrement vraie pour la classification vasculaire : les performances mesurées sur l'ensemble de test de la base choisie pour entraîner le modèle sont très différentes de celles obtenues sur une autre base. Les modèles sont ainsi généralement entraînés et testés sur chaque jeu individuellement. Et malheureusement, la montée en complexité des architectures, l'accroissement de leur nombre de paramètres, et la recherche perpétuelle de nouveaux records de performance sur des jeux de validation très limités aggravent cette tendance au surapprentissage, au détriment des capacités de généralisation des modèles.

Stagnation des performances Plusieurs méta-analyses publiées récemment compilent les performances des algorithmes de segmentation (Khandouzi, 2022 [22] ; Kumar, 2023 [23] ; Qin, 2024 [24] ; Liu, 2025 [25]) et de classification (Chen, 2024 [26]) des vaisseaux rétiniens, et les tendances qu'ils observent sont intrigantes. À leur arrivée en 2015, les CNN grimpent rapidement dans les classements de segmentation vasculaire sans toutefois égaler la méthode de Lupascu et al. (2010 [71]) qui combinait de nombreux descripteurs conçus manuellement. Ce palier sera franchi à partir de 2018 et de l'introduction des U-Net pour la segmentation sémantique des réseaux rétiniens : ainsi Xiao et al. (2018 [147]) atteignent 96.55% d'exactitude sur DRIVE avec un Res-U-Net. Mais depuis, les performances des méthodes proposées oscillent entre 96% et 97% sur cette métrique, et ce, en dépit des nombreux efforts de recherche pour améliorer les architectures des modèles et leur procédure d'entraînement. Concernant la classification, Chen et al. (2024 [26]) concluent que, bien qu'elles constituent la majorité des travaux récents, les approches par réseaux de neurones obtiennent des performances comparables en moyenne aux approches topologiques pures. Il semble donc que ces modèles aient atteint, depuis quelques années, un plafond de performances difficile à dépasser.

2.3 Paramétrisation du réseau vasculaire rétinien à partir de segmentation sémantique automatique

Les techniques de segmentation et de classification des vaisseaux énumérées à la section précédente ont beaucoup évolué depuis 2002 et le développement du premier logiciel de mesure semi-automatique des paramètres vasculaires rétiniens. On verra ici comment ces nouvelles techniques furent incorporées à ces logiciels, permettant une paramétrisation semi-automatique puis complètement automatique du réseau vasculaire des images de fond d'œil. On commencera par un rappel des paramètres vasculaires exploités par les études cliniques.

2.3.1 Paramètres vasculaires rétiniens

Calibre vasculaire La mesure de diamètre des vaisseaux est centrale pour identifier le rétrécissement symptomatique des artérioles. Sa mesure automatique repose initialement sur une détection de contours sur le profil d'intensité du canal vert, dans une direction orthogonale au vaisseau (Sherry, 2002 [13]). Cette mesure peut être affinée en ajustant un modèle gaussien à ce profil (Li, 2005 [148]). Cette méthode est particulièrement adaptée pour déterminer le calibre des artères réfléchissant la lumière en leur centre. Aujourd'hui, on considère généralement que les cartes de segmentations automatiques sont suffisamment fiables pour une mesure directe du calibre, par exemple par le calcul d'une carte de distance (*Distance Transform*) [149].

Tortuosité Une élévation de la tortuosité vasculaire est associée très tôt à l'hypertension et à l'athérosclérose (Scheie et al., 1953 [150]). Mais l'appréciation de la tortuosité est subjective, et définir un indice qui correspond à cette appréciation tout en étant calculable automatiquement à partir du squelette vasculaire n'est pas trivial. Une première formule est proposée par Lotmar et al. (1979 [151]) qui l'a définie pour chaque branche comme le ratio de la longueur de l'arc (le nombre de pixels du squelette) sur la longueur de la corde (la distance euclidienne entre le premier et le dernier pixel de la branche). En 1999, Hart et al. [152] définissent une série d'indices de tortuosité en fonction de la courbure locale κ. Ils concluent que la formule correspondant le mieux au sentiment clinique de la tortuosité est l'intégrale de κ^2 le long de la courbe. Grisan et al. (2008 [153]) proposent de calculer la tortuosité par une moyenne des ratios arc sur corde calculée entre chaque point d'inflexion de la courbe. Enfin, Trucco et al. (2010 [154]) généralisent la formule de Hart et montrent que la valeur de la tortuosité est sensible au calibre des vaisseaux. Quelles que soient la définition choisie, la tortuosité est toujours calculée à partir du squelette des artérioles ou des veinules sur lequel chaque branche a été identifiée.

Dimension fractale Les multiples ramifications du réseau vasculaire rétinien sont des structures délicates à modéliser pour les algorithmes proposés avant les années 2000. Il est donc proposé d'estimer le taux de ramification par une approche globale estimant la dimension fractale. Mainster (1990 [155]) propose de compter le nombre de pixels du squelette à l'intérieur de cercles centrés sur le disque optique et de rayon variable. Il trace alors la courbe du logarithme du nombre de pixels en fonction du logarithme des rayons et en mesure la pente pour obtenir la dimension fractale du graphe. Une implémentation plus rapide dite de *Box-Counting* est proposée par Liebovitch et Toth (1989 [156]). Elle consiste à couvrir la région d'intérêt de grilles de carrés dont on fait varier la longueur c du côté. La dimension fractale est alors donnée par la pente de la courbe du nombre de carrés contenant une partie du squelette en fonction de c. C'est cette méthode qui est retenue aujourd'hui. Il a été montré que des changements de la dimension fractale sont associés à l'hypertension, la DR, ou encore la mortalité liée aux coronaropathies ou aux AVC (Zu, 2014 [157]).

Paramètres des bifurcations Les modèles physiques d'écoulement du sang dans les vaisseaux permettent de prédire les caractéristiques géométriques optimales aux bifurcations (Murray, 1926 [158]). Les écarts entre les valeurs théoriques de ces caractéristiques et leurs valeurs mesurées sont symptomatiques d'une altération des propriétés physiques des vaisseaux, par exemple à cause du diabète (Luo et al., 2017 [159]). Pour quantifier ces écarts, Martinez-Perez et al. (2000 [160]) ont proposé une série d'indicateurs géométriques des bifurcations. Soit une branche parente (de diamètre d_0) donnant naissance à deux branches filles (de diamètre d_1 et d_2, avec $d1 > d2$), tel qu'illustré sur la Figure 2.9. On note θ_1 et θ_2 les angles de variation de direction entre la branche mère et chaque branche fille.

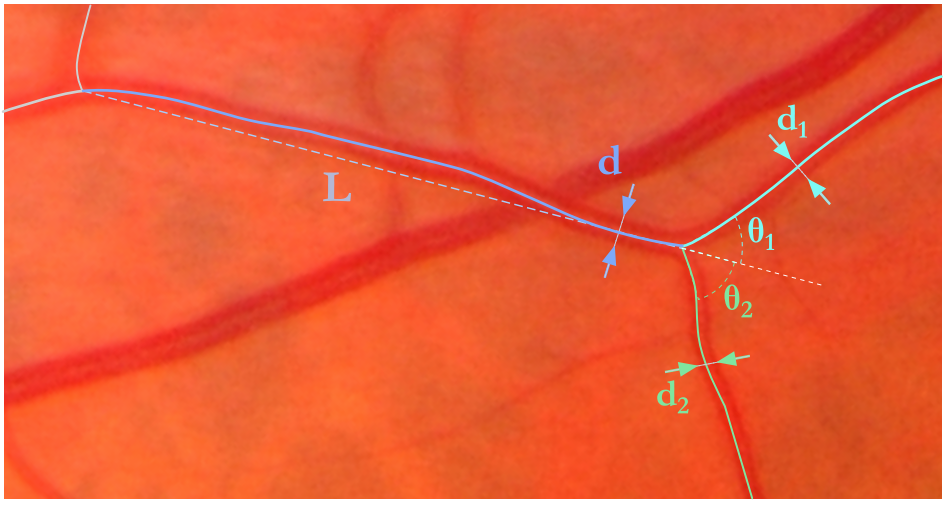


FIGURE 2.9 Typologie des grandeurs géométriques à une bifurcation.

Martinez-Perez et al. définissent :

- Le ratio longueur-diamètre : $\frac{L}{d}$;
- Le coefficient d'embranchement (ou facteur d'expansion) : $\frac{d_1^2 + d_2^2}{d_0^2}$;
- Le facteur d'asymétrie : $\frac{d_1}{d_2}$;
- L'angle d'embranchement : $\theta_1 + \theta_2$; et l'asymétrie angulaire : $|\theta_1 - \theta_2|$;
- L'exposant de jonction : e tel que $d_0^e = d_1^e + d_2^e$
  (Murray avait suggéré que la valeur optimale pour les artères était $e = 3$).

Ces paramètres sont les plus complexes à évaluer automatiquement car ils nécessitent de distinguer les bifurcations parmi toutes les jonctions du graphe vasculaire puis d'identifier quelle branche est la branche principale. En d'autres termes, leur mesure automatique nécessite de résoudre le problème de partition du graphe vasculaire (cf. Section 2.2.4).

2.3.2 Logiciel Semi-Automatique et Automatique de mesure des paramètres vasculaires rétiniens

Depuis le développement d'**IVAN** en 2004, plusieurs logiciels de mesure des paramètres vasculaires rétiniens virent le jour. Les huit principaux sont recensés dans le tableau 2.2.

Le premier est **SIVA** (Cheung et al., 2010 [20]) développé par une équipe de recherche à Singapour. La liste des paramètres vasculaires qu'il extrait est bien plus complète que celle d'IVAN (tous les paramètres mentionnés plus haut sont mesurés), et il étend aussi la zone de mesure jusqu'à 2 diamètres du disque optique autour de la papille. Pour réaliser cet exploit, SIVA confie la segmentation des vaisseaux à un algorithme non supervisé qui analyse le gradient de l'image pour identifier le squelette vasculaire (Garg et al., 2007 [161]), la partition du graphe est réalisée par une recherche globale de l'arborescence optimale (Lau et al. 2013 [96]) puis chaque sous-graphe est classifié comme artérioles ou veinules par une approche non supervisée exploitant le contraste de couleur entre artères et veines au sein d'une même image. Ces étapes sont menées sous la supervision d'un utilisateur qui a la charge de les valider et, si besoin, de les corriger.

En 2011, une équipe écossaise publie **VAMPIRE** (Perez-Rovira et al., 2011 [162]). Ce logiciel vise un domaine d'application similaire à SIVA. Il segmente les vaisseaux par un classifieur bayésien et des ondelettes multi-échelles, mais n'embarque ni méthode pour identifier automatiquement les artérioles et veinules, ni algorithme de partition du graphe : les nœuds à 3 branches incidentes sont simplement interprétés comme des bifurcations. Cette approche est moins fiable que celle de SIVA, mais contrairement à ce dernier, VAMPIRE analyse l'ensemble de l'image de fond d'œil. Le logiciel s'exécute sous la supervision d'un utilisateur.

TABLEAU 2.2 Comparaison des logiciels automatiques et semi-automatiques de mesure des paramètres vasculaires rétiniens

seg. sem. : Segmentation Sémantique (segmentation et classification simultanée) ; MSLD : Multi-Scale Line Detector ;

	Semi-Automatique			Complètement Automatique				
	IVAN	**SIVA**	**VAMPIRE**	**QUARTZ**	**SIVA-DLS**	**AutoMorph**	**TVBM**	**RMHAS**
Publication	2004	2010 [20]	2011 [162]	2015 [163]	2020	2022 [164]	2022 [165]	2022 [149]
ROI	Standard	Standard & Étendue	Complète	Complète	Standard & Étendue	Standard & Complète	Standard & Complète	Standard & Complète
Temps par image	20 min	25 min	-	54 s	quelques secondes	< 2 s	< 2 s	< 2 s
Code Public	-	-	Sur demande	-	-	Code et Modèle	Code seul	-
Algorithme de modélisation du graphe vasculaire								
Segmentation	-	Non-Supervisée (Garg 2007 [161])	Supervisée ondelettes	Supervisée MSLD	FCN (Xu 2019 [166])	U-Net	LUNet (Fhima 2024 [167])	U-Net
Classification	-	Clustering par contraste	-	Arbre de décision par couleur	*seg. sem.*	*seg. sem.*	*seg. sem.*	U-Net dédiés
Partition du graphe	-	Globale (Lau 2013 [96])	-	Locale	-	-	-	Séparation A/V
Paramètres Vasculaires								
Calibre & AVR	√	√	√	√	√	√	√	√
Tortuosité		√	√	√		√	√	√
Dimension Fractale		√	√			√	√	√
Bifurcations		√	√				√	√
Modélisation Topologique		√		(√)				√

La même équipe publiera en 2020 une version complètement automatique nommée **SIVA-DLS** qui tirera profit d'un FCN pour la segmentation et la classification des vaisseaux. Cependant, cette version ne permettra la mesure que des calibres vasculaires et de l'AVR.

En 2015, Fraz et al. [163] réduisent le temps d'annotation par image de 25 minutes à 54 secondes en étant les premiers à proposer une solution complètement automatique : **QUARTZ**. Les vaisseaux y sont segmentés par une approche supervisée exploitant des détecteurs de lignes multi-échelles (MSLD), puis identifiés dans le graphe par résolution locale de la connectivité des branches et enfin catégorisés A/V par un arbre de décision et des caractéristiques colorimétriques. Le logiciel mesure uniquement le calibre, l'AVR et la tortuosité des vaisseaux.

Enfin en 2022 sont publiés trois outils par trois équipes différentes (anglaise, israélo-belge et sino-australienne) : **AutoMorph** (Zhou et al., 2022 [164]), **TVBM** (Fhima et al., 2022 [165]) et **RMHAS** (Shi et al., 2022 [149]). Ces trois méthodes ont en commun l'utilisation de modèles U-Net pour la segmentation et la classification des vaisseaux. Mais seul RMHAS effectue une modélisation topologique du graphe vasculaire. Pour identifier chaque vaisseau, ils extraient simplement le graphe séparément des cartes de segmentation des artérioles et de celles des veinules. Ils n'ont ainsi pas besoin de partitionner le graphe, mais sont particulièrement sensibles aux erreurs de classification.

2.3.3 Limites des logiciels existants

Incohérence entre standards L'objectif motivant le développement de ces logiciels semi et complètement automatiques était de faire disparaître la subjectivité et la variabilité inter-observateur en remplaçant les annotateurs humains par un protocole de mesure déterministe et standardisé. Cet objectif n'est qu'à moitié rempli. Certes, ces logiciels sont déterministes, mais leur variété crée une duplicité des standards. Chacun propose en effet une sélection différente de paramètres vasculaires, et les paramètres qu'ils ont en commun ne sont pas forcément calculés sur la même région ou avec la même formule. Surtout, chacun a son propre algorithme de segmentation et de classification, exploitant des méthodes radicalement différentes et accroissant encore les différences de mesure d'un logiciel à l'autre. En 2022, Mautuit et al. [168] avaient comparé les valeurs d'AVR, de CRAE et de CRVE mesurées par IVAN, SIVA et VAMPIRE. Alors que ces paramètres ont une définition unique et théoriquement standardisée, Mautuit et al. observent une différence statistiquement significative du calibre moyen mesuré par chaque logiciel, et montrent que, même après avoir corrigé ce biais, les mesures de VAMPIRE sont faiblement corrélées avec celles des deux autres logiciels. Ils concluent cependant que tous obtiennent des corrélations similaires avec les principaux paramètres vasculaires (hypertension, pression systolique et diastolique, etc.).

Accessibilité restreinte Ces incohérences entre logiciels sont aggravées par l'opacité qui les entoure : sur les huit, seuls deux ont vu leur code être rendu public (les deux plus récents : AutoMorph et PVBM). Il est donc impossible de comparer leur implémentation pour identifier les sources de dissensus, définir des bonnes pratiques, et ainsi faire évoluer les protocoles de mesure. L'inaccessibilité de ces logiciels est aussi problématique pour les chercheurs en vision par ordinateur qui sont privés d'outils pour évaluer la capacité de leurs algorithmes de segmentation et de classification à produire des cartes pertinentes pour la paramétrisation du réseau vasculaire rétinien.

Concurrence des approches end-to-end Danielescu et al. ont récemment recensé les usages des analyses automatiques des vaisseaux rétiniens comme indicateur de pathologies non ophtalmiques (2024 [169]). Ils concluent qu'un large consensus existe sur la corrélation entre les paramètres vasculaires (en particulier les mesures de calibres) et les facteurs de risque cardiovasculaires, et que des recherches prometteuses les relient aussi à des pathologies cérébrales, rénales ou respiratoires. Ils indiquent cependant que ces corrélations sont souvent modérées et trop limitées pour être exploitées à des fins prédictives. Sur ce point, les paramètres vasculaires sont concurrencés par des indicateurs end-to-end prédits par des réseaux de neurones directement à partir des images de fond d'œil. C'est par exemple le cas du *retinal age gap* : la différence entre l'âge réel d'un patient et l'âge prédit à partir d'une image de sa rétine par un CNN (Zhu et al., 2023 [170]). On pourrait ainsi être tenté de délaisser les approches par modélisation du graphe vasculaire pour celles end-to-end, mais, alors que les algorithmes de segmentation sémantique vasculaire sont enfin matures, il me semble que la paramétrisation vasculaire peut révéler son plein potentiel. Surtout en considérant le gain d'interprétabilité qu'elle représente comparé à ces indicateurs « boîte noire ».

Modélisation topologique insuffisante La maturité des modèles de segmentation et la classification des vaisseaux rétiniens ouvrent de nouveaux horizons aux analyses paramétriques sur de larges cohortes. Elles rendent possible la mesure du calibre de l'ensemble des vaisseaux et non plus simplement des principaux ; l'analyse différenciée des branches en fonction du nombre de bifurcations qui les séparent des vaisseaux centraux ; ou encore une analyse fine de l'évolution morphologique et géométrique des vaisseaux à travers des données longitudinales. Mais pour permettre ces études, encore faut-il que la modélisation du graphe vasculaire soit suffisamment fiable et précise. C'est un dernier reproche qu'on peut formuler à l'encontre de ces logiciels : seulement trois extraient une représentation topologique du graphe vasculaire. Et parmi les logiciels récents exploitant des CNN, RMHAS est le seul à produire cette modélisation. Son approche pour le faire est d'ailleurs simpliste et très sensible aux erreurs de classification.

CHAPITRE 3 MÉTHODOLOGIE GÉNÉRALE

Ce chapitre synthétise la littérature de la modélisation automatique du graphe vasculaire rétinien et pointe des lacunes. Il expose ensuite les axes de recherche que j'ai choisis pour y répondre durant mon doctorat, et les raisons motivant ces choix.

3.1 Synthèse de la littérature

À l'origine de la littérature décrite dans le chapitre précédent se trouve une ambition formulée en défi : l'ambition de comprendre et prévenir les pathologies vasculaires par l'observation de la vasculature du fond d'œil ; le défi d'automatiser la paramétrisation des artérioles et veinules rétiniennes. De cette ambition et de ce défi, naît une effervescence académique autour des questions de segmentation, puis de classification, et finalement de modélisation topologique du graphe vasculaire rétinien. On essaie plusieurs approches pour détecter ces vaisseaux : d'abord non supervisées, en exploitant leur profil reconnaissable par des filtres convolutifs ou leur structure tubulaire identifiable par l'analyse hessienne ou par des opérateurs morphomat. Puis, lorsque des jeux de données annotés sont publiés, les approches supervisées s'imposent. Les CNN, et plus précisément l'architecture U-Net, surpassent toutes les méthodes antérieures pour la segmentation, et se hissent au niveau des méthodes topologiques pour la classification. Trente-six ans et près de 2000 articles ont succédé à la première publication en 1989 d'une méthode de segmentation des vaisseaux rétiniens. Les problèmes d'alors – l'amplitude de variation des diamètres de vaisseaux ou les artefacts d'illumination et de contraste – ont largement été résolus, et de nouveaux sont apparus. Concernant les réseaux de neurones, deux limitations persistent : **1.** ces modèles sont particulièrement sujets au surapprentissage, surtout compte tenu du faible volume de données annotées pour leur entraînement et leur validation ; **2.** ils semblent méconnaître la structure topologique sous-jacente au graphe vasculaire rétinien : les cartes qu'ils prédisent contiennent des artefacts de segmentation et de classification impliquant peu de pixels mais à fort impact topologique.

Les méthodes de modélisation topologique du réseau vasculaire rétinien – c'est-à-dire de partition du graphe afin de représenter chaque vaisseau comme un arbre dont le nœud racine est connu – sont évoquées dans les revues de littérature de classification comme des post-traitements prometteurs pour corriger a posteriori les artefacts topologiques des CNN. Initialement développées pour permettre la classification des artérioles et veinules avant l'arrivée des réseaux de neurones, ces méthodes semblent peu à peu retomber dans l'oubli. Elles peinent notamment à trouver leur place dans les logiciels de paramétrisation des vaisseaux rétiniens qui leur préfèrent des approches plus simplistes. Pourtant, elles permettraient une identification plus fiable des bifurcations par ces logiciels.

Au terme de cette revue de littérature, je formule quatre constats associés aux quatre étapes successives de la modélisation du graphe vasculaire rétinien : sa segmentation, sa classification, l'extraction de sa topologie, et sa paramétrisation.

Constat 1 : Les méthodes de **segmentation** ont fait l'objet de très nombreuses publications ces dernières années et voient aujourd'hui leurs performances stagner. À mon sens, cette stagnation s'explique par deux facteurs. D'une part, les données d'entraînement et de validation ne sont plus suffisantes pour tirer profit des architectures de réseaux de neurones toujours plus complexes ; d'autre part, les métriques évaluant les modèles ont atteint leur limite d'applicabilité : au niveau de performances actuel, elles sont trop sensibles aux biais d'annotation et pas assez aux artefacts topologiques.

Constat 2 : La **classification** des artérioles et veinules est dix fois moins étudiée que leur segmentation. Pourtant, la classification des petits vaisseaux éloignés du disque optique reste incertaine pour les modèles convolutifs ; et ces algorithmes sont par ailleurs sujets à des incohérences topologiques : inversant régulièrement la classification d'un vaisseau au milieu d'une branche. En outre, les jeux de données annotés pour la classification A/V sont encore plus rares que ceux annotés pour la segmentation, ce qui présente un défi supplémentaire à la capacité de généralisation des modèles de classification. Ce point est d'autant plus inquiétant que les images de fond d'œil issues des cohortes de patients dont on souhaite modéliser la vasculature rétinienne n'ont pas nécessairement les mêmes caractéristiques d'éthnie ou d'acquisition que celles des jeux ayant entraîné les modèles.

Constat 3 : L'**extraction de la topologie** de l'arbre vasculaire rétinien, maillon essentiel pour sa modélisation, était au cœur de l'attention entre 2010 et 2015 lorsqu'elle assistait la classification A/V en réalisant la partition du graphe vasculaire. Néanmoins, elle semble aujourd'hui délaissée par la communauté de chercheurs. Remise au goût du jour et combinée aux réseaux de neurones, elle pourrait pourtant corriger certains de leurs artefacts topologiques.

Constat 4 : Quatre outils de **paramétrisation** des vaisseaux rétiniens basés sur des CNN ont été publiés depuis 2020, mais un seul (RMHAS) modélise topologiquement le réseau vasculaire comme une forêt d'arbres, et la méthode choisie pour ce faire est très sensible aux erreurs de classification A/V. Le seul autre logiciel proposant une telle modélisation (SIVA) est semi-automatique et limite son analyse à 2 diamètres du disque optique autour de la papille. Ces outils sont donc en retard sur la littérature de modélisation topologique. J'ajoute que, pour 3/4 d'entre eux, le code n'est pas accessible librement. Cette situation n'est pas anecdotique : elle prive les chercheurs en segmentation et en classification des outils pour évaluer leurs algorithmes dans la perspective d'une paramétrisation des vaisseaux rétiniens.

3.2 Problématique de recherche

Les réseaux de neurones et plus spécifiquement les CNN ont reçu beaucoup d'attention de la communauté de chercheurs travaillant sur l'identification des artérioles et veinules rétiniennes. Si ces modèles culminent en haut des classements de l'état de l'art, ils restent limités par leur tendance au surapprentissage et par leur incapacité à apprendre la structure topologique sous-jacente à la vasculature de la rétine. Les logiciels récents de paramétrisation automatique des vaisseaux rétiniens qui exploitent ces modèles ont soit abandonné la modélisation topologique de l'arbre vasculaire, soit ont recours à une approche simpliste qui ne tient pas compte des artefacts topologiques de segmentation ou de classification.

Je formule donc pour cette thèse la problématique suivante : *Comment mettre à profit les recherches récentes appliquant les CNN à la segmentation et la classification des vaisseaux sur image de fond d'œil, pour extraire une modélisation topologique robuste de l'arbre vasculaire rétinien en vue d'en extraire des paramètres cliniques fiables ?*

3.3 Objectifs spécifiques

Parmi les multiples questions que soulèvent cette problématique, j'ai choisi d'orienter mon doctorat sur les objectifs spécifiques suivants.

Objectif I : **Construire un nouveau jeu de données annotées pour l'entraînement et la validation de modèles de segmentation et de classification vasculaire sur image de fond d'oeil.** Peu de jeux de données publics d'images de fond d'œil sont annotés pour la segmentation des vaisseaux rétiniens, et encore moins distinguent les artérioles et les veinules. Que ce soit pour mieux représenter la variété des conditions cliniques à l'entraînement des modèles, ou pour renforcer la confiance dans l'évaluation de leur qualité, de nouvelles annotations doivent être collectées.

Objectif II : **Améliorer les propriétés intrinsèques de généralisation des CNN appliqués à la classification A/V.** Les efforts conduits pour l'objectif I ne suffiront jamais à faire disparaître complètement le risque de surapprentissage qui plane sur les CNN, et en particulier ceux qui classifient les artérioles et les veinules. Je propose donc d'aborder ce problème par un second angle : modifier le formalisme des neurones convolutifs pour améliorer leurs propriétés intrinsèques de généralisation.

Objectif III : **Modéliser la topologie et la géométrie de l'arbre vasculaire rétinien à partir de cartes imparfaites de segmentation A/V.** Puisque les réseaux convolutifs peinent à apprendre la cohérence topologique du réseau vasculaire, il est nécessaire d'adapter les méthodes de modélisation de l'arbre vasculaire pour qu'elles tiennent compte des imperfections présentes dans les cartes de segmentation et de classification. Pour ce faire, on peut s'appuyer sur la littérature qui traite du partitionnement optimal du graphe vasculaire développée entre 2010 et 2015 pour la classification A/V.

3.4 Approche générale

Ces objectifs spécifiques sont traités dans les trois prochains chapitres de la thèse, dont voici un aperçu de leurs motivations et approches respectives.

3.4.1 MAPLES-DR : une campagne d'annotation des structures anatomiques et pathologiques de la rétine

À l'heure où les réseaux de neurones fleurissent dans tous les secteurs du traitement du signal, la pression sur la disponibilité des données annotées augmente. Dans le domaine de la segmentation d'image médicale, l'annotation des images est fastidieuse et coûteuse. Il n'est donc pas étonnant que les jeux de données publics dédiés à la segmentation et la classification des vaisseaux sur les images de fond d'œil soient insuffisants pour représenter la variété des situations cliniques. Dans les dernières années, cette situation tend néanmoins à se résorber avec la publication en 2022 des jeux comme FIVES [36], RETA [40], ou celui de Leuven-Haifa [41] en 2024. J'ai moi-même posé ma pierre à l'édifice avec la publication en 2024 de MAPLES-DR (*MESSIDOR Anatomical and Pathological Labels for Explainable Screening of Diabetic Retinopathy).*

Ce projet fut certainement le plus long de mon doctorat. Il trouve son origine en 2018 lors d'une collaboration entre le LIV4D, le MILA et une équipe de rétinologues canadiens qui souhaitaient améliorer le programme de télédépistage de la rétinopathie diabétique au Canada en l'assistant d'outils de diagnostic automatique. Étant donné l'expertise du LIV4D dans la segmentation automatique des structures rétiniennes et l'importance donnée à l'interprétabilité en IA médicale, il fut décidé que l'algorithme impliquerait des étapes de segmentation de structures anatomiques et pathologiques pour établir son diagnostic (contrairement à une approche end-to-end). Mais, à l'époque, les bases de données publiques d'images de fond d'œil pour entraîner des modèles de segmentation de lésions sont tout aussi limitées que celles pour les vaisseaux, et la nécessité d'en annoter une nouvelle nous apparaît rapidement.

Dans les années qui suivirent, sept cliniciens se sont donc relayés pour annoter quatre types de structures anatomiques (les vaisseaux, le disque optique, la papille et la fovéa) et six types de structures pathologiques symptomatiques de la DR sur 200 images de fond d'œil issues du jeu de données public MESSIDOR [171].

Cette exhaustivité d'annotation est inédite. J'ai donc dû concevoir un protocole d'annotation dédié s'appuyant : d'une part sur des présegmentations prédites par des algorithmes puis corrigées par les rétinologues ; et d'autre part sur une plateforme d'annotation web implémentée pour l'occasion. Par la suite, j'ai pu réutiliser ces outils et les segmentations des vaisseaux pour annoter les cartes des artérioles et les veinules sur 100 images de MAPLES-DR. Ces cartes ont été mises à profit dans les deux chapitres suivants.

3.4.2 STEERED CNN : l'équivariance par rotation au service de la classification des artères et veines rétiniennes

Pour approcher le second objectif, je prends la littérature à revers : plutôt que d'accroître la complexité architecturale et paramétrique des modèles de classification des vaisseaux rétiniens, je vais plutôt chercher à réduire leur nombre de paramètres en exploitant une propriété de leur tâche : la classification ne dépend pas de l'orientation des vaisseaux. Les CNN qui réalisent cette classification doivent donc être *équivariant par rotation* : c'est-à-dire qu'une rotation de l'image présentée en entrée doit théoriquement induire une rotation analogue sur la carte de prédiction en sortie. Les CNN possèdent naturellement une propriété similaire, ils sont équivariants par *translation* : par définition de l'opérateur de convolution, une translation de leur entrée produit une translation équivalente de leur sortie ; mais l'équivariance par *rotation* n'est pas inscrite dans leur formalisme. Le modèle doit donc l'apprendre à partir des données d'entraînement. Et lorsque peu de données sont disponibles – comme c'est le cas pour la classification A/V rétinienne – cet apprentissage est approximatif, et peut être source d'erreur.

Cette lacune m'a soufflé une idée assez peu étudiée dans la littérature : modifier le formalisme des neurones convolutifs – au cœur de tous les CNN et des U-Nets – afin d'orienter leurs filtres selon la direction locale de chaque vaisseau de l'image avant d'appliquer la convolution. Ainsi, chaque filtre du modèle pourrait participer à la détection de tous les vaisseaux d'une image, quel que soit leur orientation, au contraire des neurones convolutifs standards qui perçoivent différemment les vaisseaux horizontaux et verticaux. En mutualisant les filtres de cette façon, les *steered* CNN devraient nécessiter moins de paramètres pour réaliser la même tâche, ils seraient ainsi moins sujets au surapprentissage et plus aptes à généraliser.

3.4.3 FUNDUS VESSELS TOOLKIT : modélisation géométrique et topologique de l'arbre vasculaire rétinien

Le troisième objectif de cette thèse traite des dernières étapes de la modélisation du graphe vasculaire rétinien : l'extraction de sa topologie et de sa géométrie pour mesurer des paramètres vasculaires. Il porte plus spécifiquement une ambition théorique : montrer comment les techniques de partitionnement optimal de l'arbre vasculaire, développées initialement comme un prétraitement topologique en vue de classer chaque branche comme artériole ou veinule, peuvent à nouveau retrouver leur place aujourd'hui, comme post-traitement des cartes de segmentation sémantique prédites par CNN pour en corriger les incohérences topologiques.

Après être parvenu à la conclusion qu'aucun logiciel de paramétrisation n'offrait le cadre dont j'avais besoin, je dus me résoudre à développer moi-même un nouvel outil : la librairie Python baptisée *Fundus Vessels Toolkit* (FVT). En me lançant dans ce développement, je n'imaginais pas la nuée d'anicroches dont recèle la modélisation topologique du réseau vasculaire rétinien. La variabilité naturelle des vaisseaux produit en effet tant d'exceptions, tant de cas étranges, que l'implémentation de FVT se révéla être une épreuve de minutie algorithmique.

L'ultime chapitre de méthode de cette thèse est donc traversé par deux mouvements. D'une part, le développement théorique d'une nouvelle méthode de correction topologique des imperfections contenues dans les cartes de segmentation prédites par CNN. Et d'autre part, l'ajustement minutieux de chacune des étapes qui jalonnent le processus d'extraction de la géométrie et de la topologie de l'arbre vasculaire rétinien.

CHAPITRE 4 MAPLES-DR : UNE CAMPAGNE D'ANNOTATION DES STRUCTURES ANATOMIQUES ET PATHOLOGIQUES DE LA RÉTINE

Le présent chapitre est consacré à *MAPLES-DR* (MESSIDOR Anatomical and Pathological Labels for Explainable Screening of Diabetic Retinopathy), un nouveau jeu de segmentation des structures anatomiques et pathologiques de la rétine pour 200 images de fond d'œil issues de la base publique MESSIDOR [171]. Il détaille le choix du protocole d'annotation (résumé sur la Figure 4.1), en présente les résultats et se conclut par une étude inédite de la variabilité inter-observateur lors de la segmentation des lésions rétiniennes.

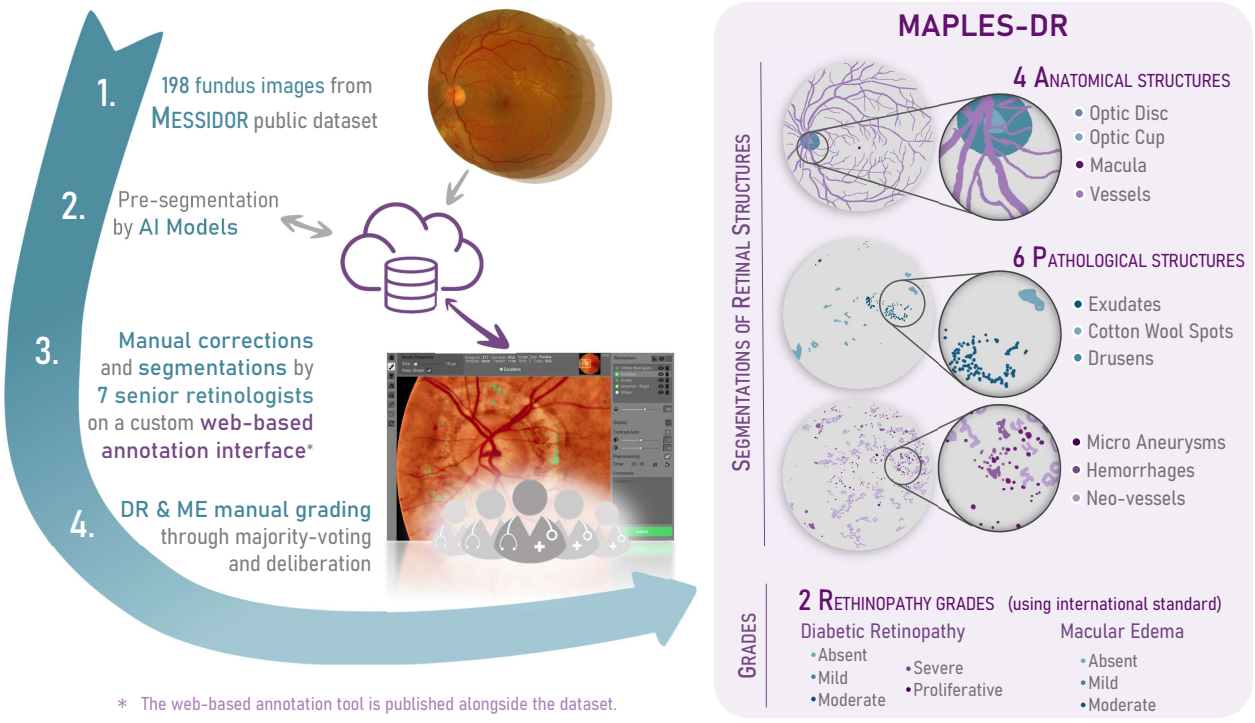


FIGURE 4.1 Résumé graphique du protocole d'annotation de MAPLES-DR.

4.1 Motivations et travaux connexes

Enjeux du dépistage automatique de la DR

La rétinopathie diabétique (DR) est une complication du diabète qui cause des lésions sur la micro-vasculature rétinienne et conduit dans plus d'un quart des cas à une détérioration de la vue. Elle concerne 22% des patients diabétiques, soit 103,12 millions d'adultes en 2020 [172], et est la première cause de cécité dans le monde pour les adultes entre 25 et 75 ans [173]. Sa progression lente se caractérise cliniquement par des stades définis par l'apparition de lésions sur la rétine qui doivent être surveillées chez les populations à risque pour assurer une prise en charge précoce et espérer préserver leur vision [174]. En Amérique du Nord, 40% des patients diabétiques ne sont pas dépistés pour la DR malgré de larges programmes de dépistage [175–177]. Cependant, l'expérience fournie par ces programmes montre une amélioration du taux de patients surveillés grâce au dépistage par télé-ophtalmologie [178], et suggère que des algorithmes de diagnostic automatique de la DR pourraient encore démultiplier leur efficacité, en réduisant le coût des examens de dépistage pour augmenter leur fréquence, et en assurant une prise en charge immédiate des patients référés [174].

Durant la dernière décennie, les CNN puis les Vision Transformers furent appliqués avec succès aux diagnostics automatiques de DR à partir d'images de fond d'œil, en tous cas dans le cadre de recherches académiques. Au cœur du développement de ces algorithmes supervisés sont les bases de données publiques d'images annotées, sur lesquelles sont entraînés puis validés ces modèles [179]. Parmi elles, les bases de données Eyepacs [180] et MESSIDOR-2 [171] ont permis aux algorithmes de dépistage automatique d'atteindre des performances supérieures aux exigences de la FDA (sensibilité supérieure à 85% et spécificité supérieure à 82,5%) [179], ou à celles d'experts humains [181].

Pourtant, il persiste un manque de confiance du personnel médical envers ces technologies qui n'atteignent pas les standards d'une *"explainable AI"*. D'une part, ces modèles ne sont pas interprétables : les règles implicitement utilisées pour produire les diagnostics nous sont inintelligibles et rien n'assure qu'elles soient similaires à celles des rétinologues [179,182,183]. Cette opacité est d'autant plus inquiétante aux regards des imperfections des vérités terrains utilisées pour évaluer ces modèles : les bases de données publiques de DR sont spécifiques à une population et leurs protocoles d'annotations ne sont pas standardisés. De ce fait, les modèles entraînés sur l'une sont spécifiques à ses biais et sont incompatibles avec les autres [179,183]. D'autre part, ces modèles se prêtent mal à une collaboration entre AIs et ophtalmologistes humains car les diagnostics produits par ces modèles sont étayés de trop peu d'explications.

Segmenter les structures rétiniennes pour dépister la DR

Pour dépasser ces limitations, il nous est apparu qu'une partie de la réponse résidait dans l'amélioration des labels d'entraînement. En effet, l'incertitude concernant l'évaluation des algorithmes et les problèmes de généralisation appellent à plus de diversité dans les bases de données de validation et à une meilleure documentation de leurs biais et de leur processus d'annotation. Par ailleurs, la faible explicabilité des algorithmes appelle à plus d'exhaustivité et de pertinence clinique des labels qui guident leurs entraînements, bien au-delà des simples diagnostics. Dans leur rapport *"Four principles of Explainable AI"* [184], le NIST (National Institute of Standards and Technology) rappelle que pour être satisfaisante, la justification d'une prédiction doit être «*meaningfull*» pour son public cible. C'est-à-dire, dans notre cas, qu'elle soit formulée dans un vocabulaire dont la sémantique est familière aux ophtalmologistes. Or les grades de DR ne constituent qu'une petite partie de ce vocabulaire : les *justifications cliniques* s'appuient plutôt sur les structures anatomiques ou pathologiques de l'œil (vaisseaux, macula, lésions rouges ou claires, etc.) [185–187]. Ainsi, que ce soit pour faire converger *a priori* les modèles vers des représentations compatibles avec les connaissances cliniques ou pour interpréter *a posteriori* ces représentations en les comparant à des biomarqueurs cliniquement reconnus, les bases de données annotées au niveau du pixel ont un rôle crucial dans le développement de modèles de dépistage explicable.

Bases de données publiques pour la segmentation des structures rétiniennes

Annoter au niveau du pixel les structures anatomiques ou pathologiques de la rétine requiert considérablement plus de temps et d'effort qu'apposer un simple diagnostic de DR. Les bases de données proposant de telles annotations sont donc naturellement bien moins nombreuses. La plupart ne proposent des annotations que pour un seul type de structure : par ex. vaisseaux [31, 36, 40, 45], disque optique [188–190], exsudats [191], ou micro-anévrismes [192].

Concernant les structures pathologiques, seulement quatre bases de données proposent des annotations pour plusieurs lésions symptomatiques de la DR (i.e. micro-anévrisme, hémorragies, exsudats et CWS). *FGADR* [193] et *Retinal Lesions* [194] contiennent de très nombreuses images mais en segmentant grossièrement les zones pathologiques. Au contraire, *IDRiD* [195] et *DDR* [196] fournissent des segmentations précises mais pour moins d'images (respectivement 81 et 757 images). Aucune ne contient d'annotation de structures anatomiques. Par ailleurs, les articles qui accompagnent ces bases de données se concentrent sur la hausse des performances de segmentation qu'elles permettent, mais ne décrivent que très brièvement le protocole ayant produit leurs annotations et aucun ne mesure leur variabilité inter-observateur. Il est donc très difficile de juger de leur fiabilité.

4.2 Méthode d'annotation de MAPLES-DR

4.2.1 Sélection des images

Lors de la conception du projet, il fut décidé que la campagne d'annotation (initialement conçue comme une campagne pilote) se limiterait à 200 images d'une base de données publique d'images de fond d'œil. Le choix de la base se porta sur MESSIDOR parce qu'elle servait de référence pour valider les algorithmes de dépistage de DR et qu'elle ne disposait d'aucune annotation de structures rétiniennes. La sélection des 200 images (parmi les 1200 de MESSIDOR) fut randomisée de sorte que 30 soient saines, 59 soient de stade R1, 55 soient R2 et 56 soient R3 (selon les diagnostics de DR fournis par MESSIDOR). Cette distribution n'est pas représentative de la prévalence de la rétinopathie ni dans le jeu de données MESSIDOR, ni dans la population réelle (c.f. Figure 4.2). Néanmoins, elle assure un nombre suffisant d'images pour chaque stade de la maladie tout en se concentrant sur ceux critiques pour le dépistage (R1 et R2).

À la fin de la campagne d'annotation, il nous est apparu que 2 images étaient des duplicatas [1]. Bien qu'ayant déjà été annotées par nos experts, ces images furent exclues de MAPLES-DR, réduisant le nombre total d'images à 198. Leurs cartes de segmentations et diagnostics furent tout de même rendues publiques à des fins de transparence et d'étude de variabilité.

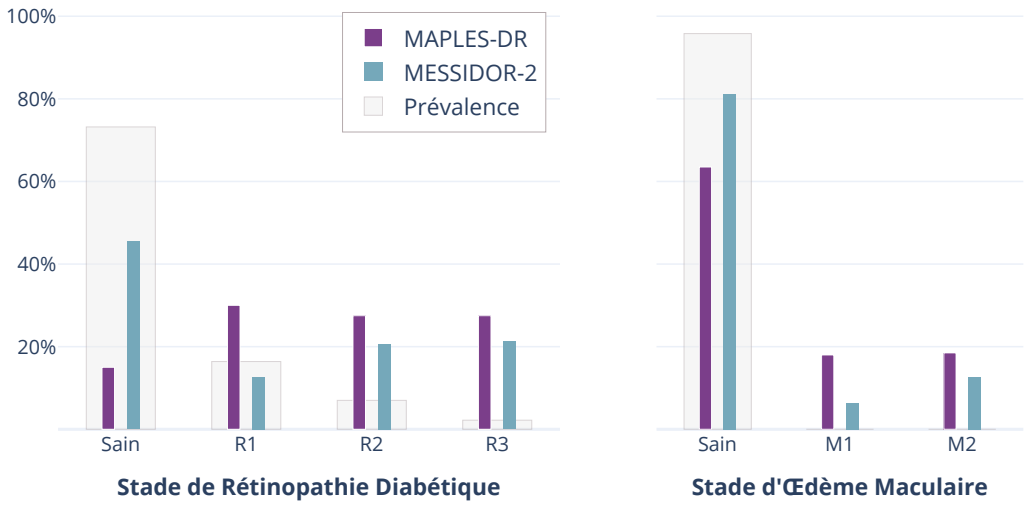


FIGURE 4.2 Différences de prévalence de DR et ME entre MAPLES-DR (violet), MESSIDOR au complet (turquoise) et une population dépistée par téléophtalmologie [2] à Toronto (gris).

1. La base de données originale MESSIDOR contenait plusieurs images accidentellement dupliquées portant des noms différents. Cette erreur fut corrigée ultérieurement dans MESSIDOR-2 [197], mais malheureusement notre travail s'appuyait sur la première version.

4.2.2 Sélection des structures anatomiques et pathologiques à annoter

Le choix d'inclure ou non les structures rétiniennes dans la liste des annotations de MAPLES-DR fut concerté avec les rétinologues en s'appuyant sur le rôle de chacune dans l'histopathologie de la DR et dans son dépistage.

Structures anatomiques Bien que les structures anatomiques rétiniennes soient présentes dans toutes les images, y compris les images saines, leur aspect et leur proximité avec les lésions fournissent de précieuses informations pour le diagnostic. La morphologie des *vaisseaux rétiniens* est révélatrice des stades de la DR : une augmentation de la tortuosité artériolaire est associée aux stades léger et modéré [198], tandis que le «venous beading» et la dilatation des veinules sont plutôt des symptômes des stades prolifératifs sévères. De plus, la gravité d'une lésion dépend souvent de sa position par rapport au *disque optique* ou à la *macula*. Ainsi, l'œdème maculaire est évalué en comptant le nombre de lésions situées à un ou deux diamètres du disque optique par rapport à la macula. La Figure 4.3 présente un exemple d'annotations de structures anatomiques de MAPLES-DR.

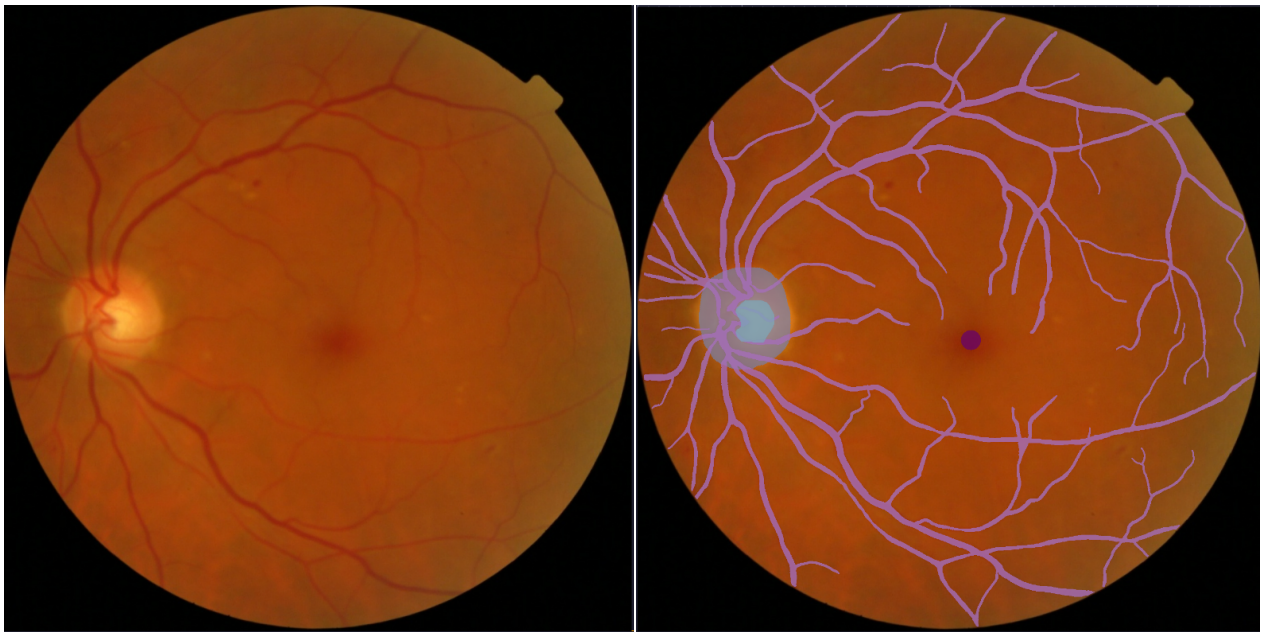


FIGURE 4.3 Structures anatomiques annotées dans MAPLES-DR (Bleu sombre : disque optique, bleu clair : papille, violet sombre : macula, violet clair : vaisseaux).

Lésions rouges Le diabète altère les parois des vaisseaux provoquant parfois des dysfonctionnements de la microvasculature rétinienne et l'apparition de microanévrismes, d'hémorragies, d'anomalies microvasculaires intrarétiniennes (IRMA) ou de néovaisseaux. Ces structures pathologiques sont appelées « *lésions rouges* ». Les *microanévrismes* – petites dilatations circulaires de capillaires – sont des signes précoces de dysfonctionnement microvasculaire et sont associés à une DR légère. Aux stades plus avancés de la pathologie apparaissent les *hémorragies intrarétiniennes* puis les *IRMA* dont la prolifération coïncide avec le stade sévère de la maladie, ce qui indique un risque de 50 % de développer une néovascularisation en l'espace d'un an. L'émergence de *néovaisseaux* (NV) marque la transition vers le stade prolifératif, le stade le plus sévère du dépistage de la DR, qui nécessite une consultation immédiate chez un ophtalmologiste. En effet, s'ils ne sont pas traités, ces néovaisseaux risquent de produire des hémorragies prérétiniennes ou vitréennes, entraînant une perte visuelle majeure. La distinction entre les NV et les IRMA est difficile à établir à partir des seules images du fond d'œil et nécessite normalement une angiographie à la fluorescéine. En l'absence de cette modalité, les deux furent annotés comme NV dans MAPLES-DR.

Lésions claires Dans les stades graves de la DR, la rétine s'épaissit et des *exsudats durs* – des dépôts clairs autour de capillaires endommagés – peuvent apparaître. Des ischémies plus sévères provoquent parfois un blocage du transport axonal dans la couche de fibres du nerf optique, qui peut à son tour entraîner l'apparition de lésions connues sous le nom de *nodules cotonneux*. Ces lésions se caractérisent par leur aspect blanc et leurs contours flous. Si leur étiologie principale est la rétinopathie diabétique, les nodules cotonneux (CWS) peuvent également être observés dans d'autres maladies vasculaires comme l'hypertension artérielle. Enfin, MAPLES-DR fournit aussi des annotations de *Drusens*. Bien que ces lésions claires soient plus symptomatiques de la dégénérescence maculaire liée à l'âge (DMLA) que de la DR, leur aspect est similaire à celui des exsudats et peut être confondu avec eux. Ils sont donc inclus dans MAPLES-DR pour marquer la distinction entre eux.

4.2.3 Recrutement des annotateurs

Les cartes de segmentation de MAPLES-DR ont été annotées par sept rétinologues canadiens affiliés à cinq hôpitaux différents : l'Hôpital Maisonneuve-Rosemont, le University Health Network (à Toronto), le CHUM (Centre Hospitalier Universitaire de Montréal), le Centre Hospitalier Universitaire Saint-Justine, et l'Université de Sherbrooke. Tous les annotateurs étaient des rétinologues seniors qui ont été recrutés suite à leur implication dans des programmes de téléophtalmologie pour la détection de la DR au Québec et en Ontario.

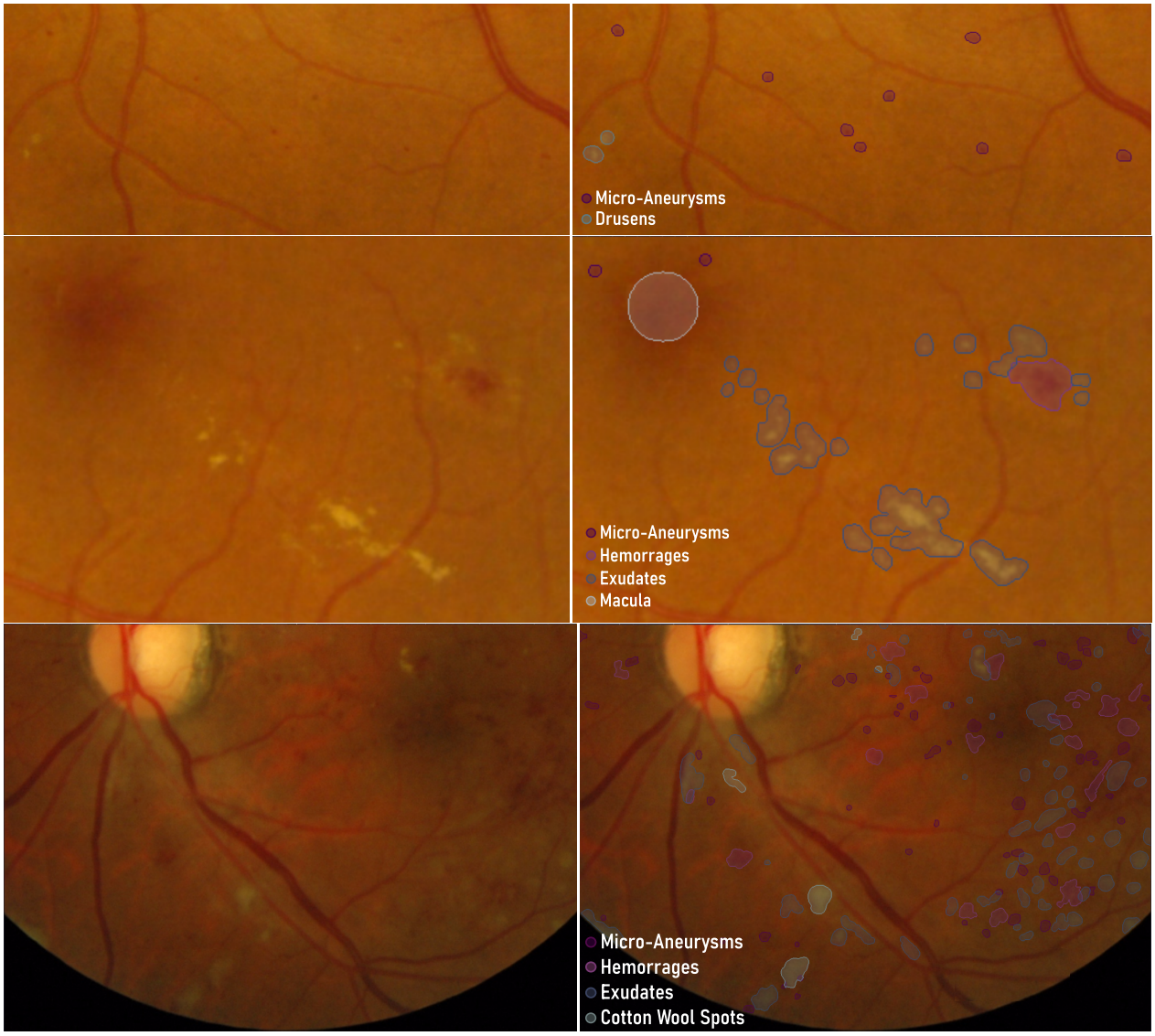


FIGURE 4.4 Exemples d'annotations de structures pathologiques

4.2.4 Conception du protocole d'annotation

MAPLES-DR se limite à l'annotation de 200 images, mais elle fut conçue comme une campagne pilote préparant le terrain à l'annotation de grandes bases de données (i.e. les milliers d'images de télémédecine canadienne) par la collaboration de multiples annotateurs répartis dans les différentes provinces canadiennes. Aussi, une attention particulière fut portée à maximiser l'efficacité des annotateurs lors de la conception du protocole d'annotation. Les caractéristiques de l'outil d'annotation détaillées ci-dessous témoignent de cette ambition.

Prétraitement des images de fond d'oeil

Les dimensions des images de fond d'œil dans MESSIDOR varient de 1440×960 à 2304×1536 pixels. Pour uniformiser la résolution des cartes de segmentation dans MAPLES-DR, toutes les images furent recadrées et redimensionnées à une résolution de 1500×1500 pixels. Ensuite, deux algorithmes d'amélioration d'image furent appliqués pour aider les annotateurs dans leur lecture. Le premier visait à corriger les variations d'illumination tout en préservant les teintes de l'image, par une soustraction de la médiane et par une égalisation d'histogramme CLAHE sur le canal de luminosité. Le second maximisait les contrastes en normalisant indépendamment les intensités de chaque canal couleur de manière à étendre au maximum leur plage de valeurs (voir Figure 4.5). Ces deux algorithmes de prétraitement pouvaient être activés individuellement ou en combinaison par le clinicien lors de l'annotation. Ils furent particulièrement utilisés lors de l'annotation des vaisseaux et des lésions rouges.



(a) Sans pré-traitement (b) CLAHE (c) Maximisation des constrastes (d) Combinaison des deux pré-traitements

FIGURE 4.5 Prétraitements disponible sur la plateforme d'annotation pour améliorer la lecture des image de fond d'oeil

Préannotation automatique

La segmentation médicale est une tâche intrinsèquement fastidieuse. Aussi, plutôt que d'annoter de zéro, les rétinologues furent chargés d'examiner des cartes de segmentation générées par IA (que nous appellerons *cartes préannotées*) et d'en corriger les éventuelles erreurs. Cette méthode, bien moins laborieuse, fut utilisée pour annoter les exsudats, les microanévrismes, les hémorragies et les vaisseaux sanguins. Pour les autres structures pathologiques (néovaisseaux, drusen et CWS), les bases de données publiées à l'époque n'étaient malheureusement pas suffisantes pour entraîner des modèles de segmentation performants. L'annotation des autres structures anatomiques (macula, disque optique et papille) étant plus simple, elle ne nécessitait pas de préannotations.

En pratique, la génération des cartes préannotées fut confiée à deux modèles du niveau de l'état de l'art en 2018 : l'un chargé de segmenter les lésions [199], l'autre de segmenter les vaisseaux [114]. Les valeurs de seuillage des cartes de probabilités prédites par ces deux modèles furent cependant légèrement réduites de façon à favoriser les faux positifs par rapport aux faux négatifs. Ce choix reposait sur l'hypothèse qu'examiner les structures préannotées et supprimer celles incorrectes serait une tâche plus aisée que repérer et annoter celles manquantes. Inversement, toutes les lésions préannotées sur les images saines (selon les étiquettes DR de MESSIDOR) furent automatiquement effacées, afin d'éviter à l'annotateur de devoir les supprimer manuellement. Enfin, lorsque la première moitié des segmentations de vaisseaux fut corrigée, on réentraîna le modèle de préannotation vasculaire et on mit à jour les préannotations de la seconde moitié. Le temps moyen nécessaire à l'annotation des vaisseaux fut ainsi réduit de 10% entre la première et la seconde moitié.

Plateforme web d'annotation

Ce protocole de collecte de segmentations par correction de préannotation n'était pas aussi répandu en 2018 qu'il ne l'est aujourd'hui. À l'époque, aucun outil libre ne trouvait satisfaction à nos yeux. Nous avons donc conçu les maquettes d'une plateforme d'annotation dont le développement fut confié à une équipe d'étudiants en génie logiciel. J'ai supervisé leur travail puis ai complété, maintenu, et déployé la plateforme. Son code a récemment été rendu public sur Github, accompagné d'une CLI Python et de documentation.

La collaboration entre des experts éloignés géographiquement fut le premier défi à relever lors de la conception. Elle fut assurée par une solution double. **1.** la plateforme d'annotation prit la forme d'une application web accessible en tout temps sur le serveur du laboratoire. Pour des raisons de sécurité évidentes, l'accès fut protégé par un portail d'authentification et on assigna à chaque rétinologue un compte utilisateur propre. **2.** La plateforme fut dotée d'un système d'affectation de tâches : chacune étant liée à une image et un annotateur et y associant une liste des structures rétiniennes à réviser et compléter. Cette liste de tâches à compléter est présentée aux rétinologues immédiatement après leur authentification (cf. Figure 4.6). Chacun dispose de sa propre version des annotations de sorte que plusieurs rétinologues puissent annoter la même image. Une CLI en Python me permettait de répartir préalablement les tâches parmi l'équipe d'annotateurs puis de modifier ces affectations, d'en surveiller la complétion et de télécharger les annotations corrigées.

Après avoir sélectionné une tâche, l'utilisateur est transporté vers le cœur de la plateforme : l'éditeur d'annotation. Celui-ci est avant tout un outil de visualisation et de révision des multiples structures rétiniennes, chacune possédant son calque et sa couleur propre (cf. Fi-

gure 4.7). L'utilisateur peut cacher chaque calque individuellement et ajuster leur opacité afin de les rendre semi-transparents. L'éditeur permet aussi de cacher ou révéler toutes les annotations d'un appui sur une touche, et ainsi d'examiner par clignotement les pixels de l'image qu'elles recouvrent. Dans la même optique, le mode "bordure" remplace les cartes de segmentation par une carte de leurs contours, facilitant la visualisation des structures détourées. Enfin, l'annotateur peut à son gré zoomer et activer individuellement ou en combinaison les prétraitements de l'image fundus (cf. Section 4.2.4).

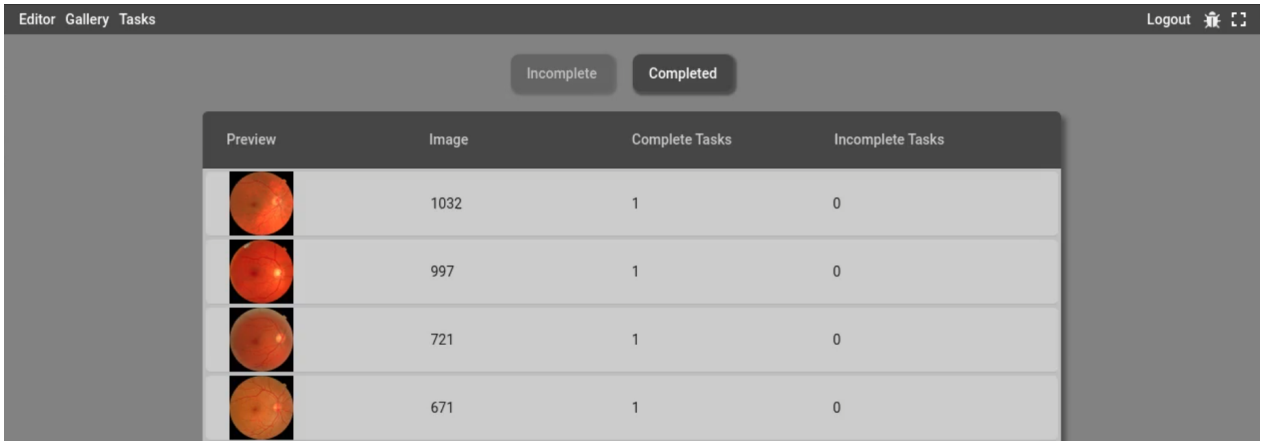


FIGURE 4.6 Interface de sélection d'une tâche parmi celles assignées.

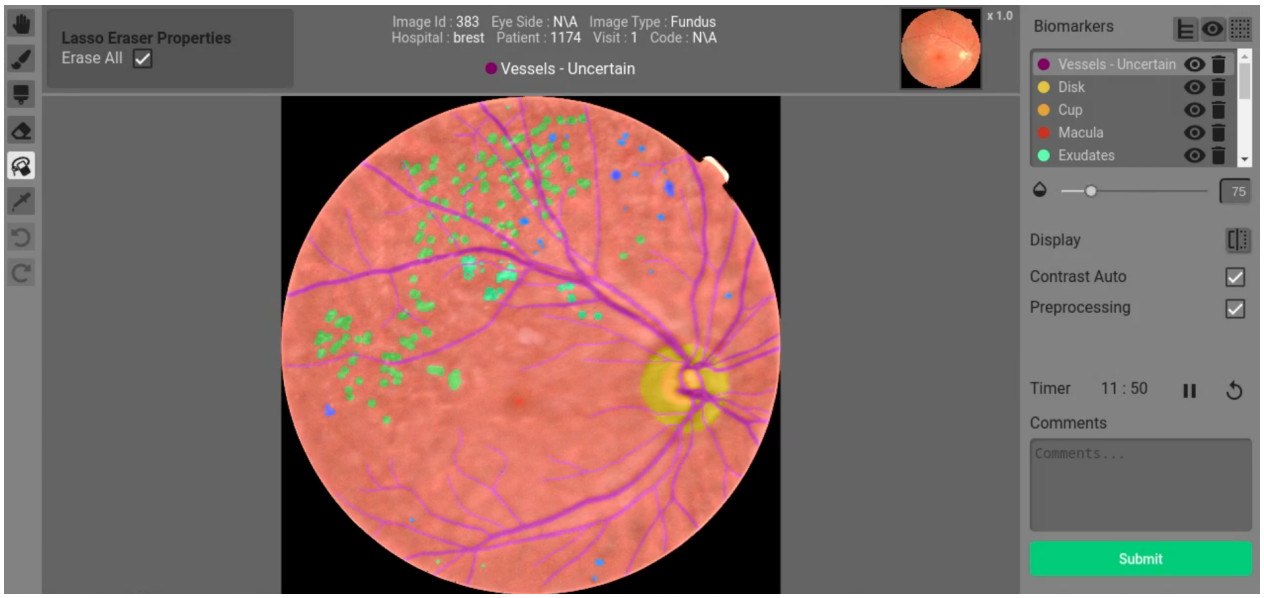


FIGURE 4.7 Interface d'annotation des structures rétiniennes.

Les outils d'édition en eux-mêmes sont similaires à ceux de n'importe quel logiciel de dessin. Un pinceau et une gomme circulaires de taille réglable ajoutent ou enlèvent des pixels du calque d'annotation actif. Les deux sont doublés d'outils de détourage permettant de délimiter des zones en traçant leurs contours. Une option du pinceau implémentée dans la plateforme la distingue des logiciels de dessin traditionnels : elle transforme son comportement pour qu'il n'ajoute pas de nouveaux pixels, mais plutôt qu'il remplace le type d'annotations existantes sans en modifier la segmentation. Cette option est particulièrement utile pour requalifier une préannotation d'exsudats en drusen, ou pour spécifier le type d'un vaisseau en artère ou veine. Après avoir corrigé l'annotation, le rétinologue peut laisser un commentaire puis, d'un clic, la sauvegarder et charger la suivante. Le temps d'annotation est enregistré automatiquement.

Instructions fournies aux retinologues

En préparation de la campagne d'annotation, les 10 structures rétiniennes à segmenter furent regroupées en 4 catégories : **1.** lésions rouges (micro-anévrismes, hémorragies et néovaisseaux), **2.** lésions claires (exsudats, drusen, CWS), **3.** vaisseaux, **4.** structures anatomiques hors vaisseaux (disque optique, papille, macula). Puis on attribua à chaque rétinologue l'une des quatre catégories de sorte qu'il n'eût à annoter que les structures associées. Chacune des 200 images fut ainsi vue par plusieurs rétinologues, mais chaque type de biomarqueur n'a été examiné que par un seul expert : l'un segmentant les vaisseaux, l'autre les lésions rouges, etc. Cette division des tâches d'annotation visait à spécialiser chaque clinicien dans l'étiquetage d'un type de structure rétinienne pour simplifier la courbe d'apprentissage des outils d'annotation et accélérer l'ensemble du processus. Cette approche, combinée à la pré-segmentation des lésions et vaisseaux, permit de réduire le temps pour annoter une image à 22 minutes cumulées en moyenne.

Hormis la liste des structures à segmenter et des recommandations sur l'utilisation des outils d'annotation, aucune instruction explicite ne fût fournie aux cliniciens : ni sur une définition clinique des structures (nous faisions totalement confiance dans leur expérience médicale), ni sur le niveau de détail attendu. Radsch et al. [200] a récemment montré que pour améliorer la qualité des annotations biomédicales, il était préférable de fournir des images d'exemples plutôt que des instructions textuelles même détaillées. Les cartes de préannotations remplissent parfaitement cette fonction. En présentant à l'annotateur des lésions petites (leur diamètre médian était de 13 pixels) et nombreuses ; elles incitent à segmenter chaque lésion individuellement (au lieu d'encercler la zone générale) et à prêter attention à toutes les lésions, même les plus petites.

4.2.5 Bilan du travail d'annotation

Les structures rétiniennes de MAPLES-DR furent en réalité majoritairement revues et corrigées par seulement trois rétinologues (cf. Figure 4.8) : Dr. Boucher (87% des lésions claires), Dr. Brent (62% des lésions rouges) et Dr. Duval (100% des vaisseaux). Au total, 69 heures ont été nécessaires pour réviser et annoter les 200 images.

La macula, le disque optique et la papille furent les plus rapides à annoter, avec une moyenne de 2 minutes par image ; suivis par la segmentation des vaisseaux, avec une moyenne de 6 minutes par image. Pour cette dernière, la pré-segmentation semble avoir été d'une aide précieuse : 77,5 % des pixels annotés comme vaisseaux dans MAPLES-DR sont issus des préannotations, et moins de 8,3 % des pixels préannotés ont été manuellement identifiés comme faux positifs et effacés (cf. Figure 4.9). Pour autant, la segmentation des vaisseaux a nécessité d'importantes corrections : en moyenne 56k pixels par image ont été segmentés manuellement, majoritairement des petits vaisseaux (voir Figure 4.10a).

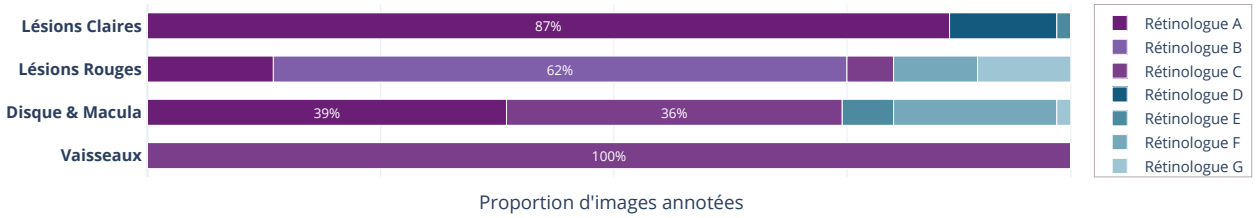


FIGURE 4.8 Proportion d'images annotées par chaque rétinologue.

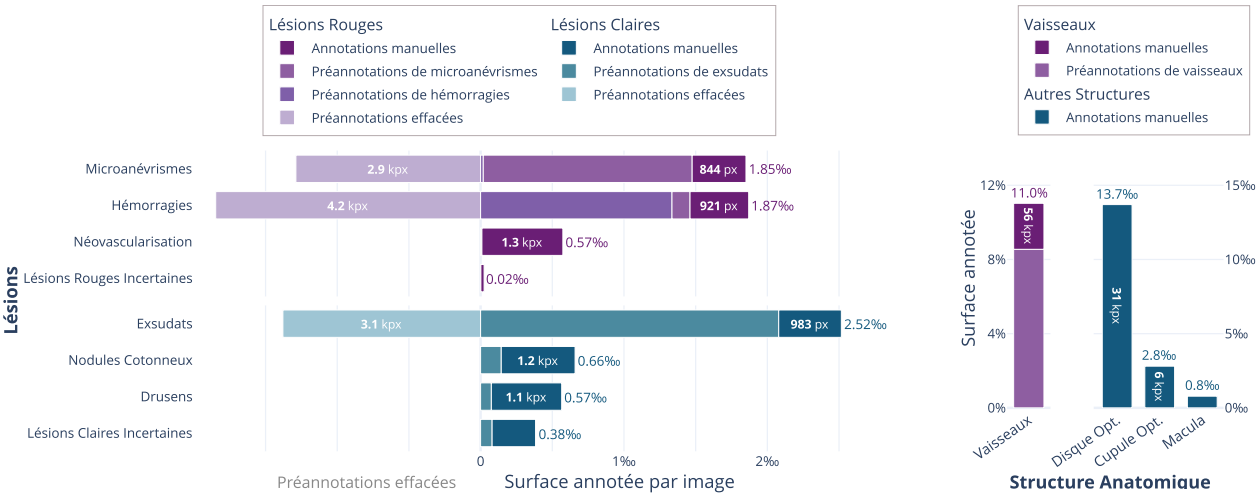
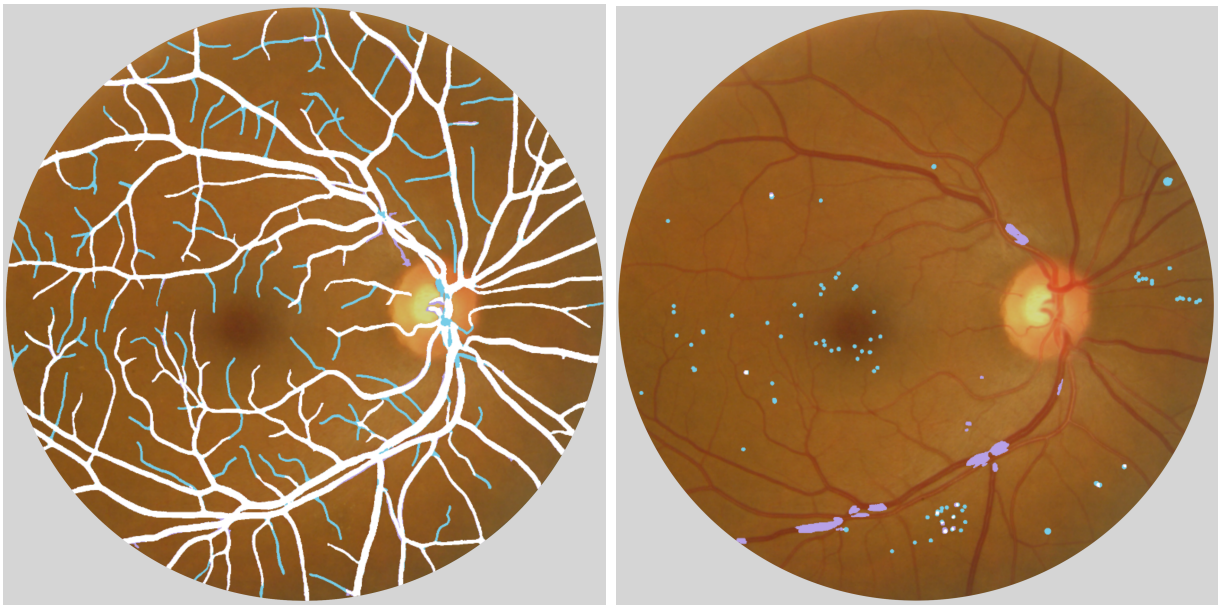


FIGURE 4.9 Vue quantitative du travail d'annotation pour chaque catégorie de structures rétiniennes. Les barres sombres correspondent aux segmentations ajoutées manuellement, celles claires correspondent aux préannotations effacées. Chacune est annotée du nombre de pixels corrigés par image. Le pourcentage indique l'aire occupée par chaque structure rapportée à l'aire des images.



(a) Vaisseaux (b) Toutes lésions confondues

FIGURE 4.10 Préannotations corrigées par les rétinologues. (*Blanc* : préannotation inchangée ; *Turquoise* : segmentation ajoutée manuellement, *Violet* : préannotation effacée manuellement.)

Les structures pathologiques ne concernent qu'une petite fraction des annotations de MAPLES-DR : elles couvrent moins de 0.5% des images. Pourtant, c'est bien elles qui ont nécessité le plus grand effort d'annotation. Le temps passé sur chaque image est très variable : de quelques dizaines de secondes pour les images saines à une heure pour les plus pathologiques. Comme pour les vaisseaux, l'annotation manuelle des structures pathologiques s'appuie largement sur les pré-segmentations. Pour les lésions qui en disposaient (microanévrismes, hémorragies et exsudats), seul un quart des pixels a été ajouté manuellement, les 3 quarts restant proviennent directement des cartes préannotées. Mais contrairement aux vaisseaux, les cartes de lésions préannotées contenaient plus de faux positifs : plus de 58 % des hémorragies préannotées furent effacées, de même que 47 % des microanévrismes et 40 % des exsudats (cf. Figure 4.9). La plupart des hémorragies et des microanévrismes effacés étaient en réalité des vaisseaux, comme le montre la Figure 4.10b. En moyenne, l'annotation des structures pathologiques prit 6 minutes par image pour les lésions claires et 10 minutes pour les lésions rouges.

4.2.6 Préparation des données pour la publication

La campagne d'annotation s'est conclue par un travail de formatage des cartes de segmentation corrigées en vue de leur publication. Tout d'abord, un rapide parcours des annotations a permis d'éliminer quelques faux positifs évidents (par ex. une hémorragie sur un vaisseau) introduits par les préannotations mais ayant échappé à la vigilance des rétinologues. Cette brève inspection révéla une segmentation manquante pour la macula (l'image étant centrée

sur le disque optique, la macula est hors-champ) ; et six pour le disque optique à cause de contours trop flous pour être précisément délimités. Le jeu de données fut ensuite divisé en un ensemble d'entraînement et un ensemble de test. Pour s'assurer de la représentativité de ce dernier, on construisit un vecteur indiquant la présence ou l'absence de chaque type de lésion, puis on appliqua la stratification «*iterative multilabel-shuffle*» proposée par Sechidis et al. [201] de sorte que l'ensemble de test soit constitué de 60 images laissant le reste (138 images) pour l'ensemble d'entraînement. Cette séparation est clairement apparente dans l'arborescence de l'archive publique de MAPLES-DR afin que toutes les publications futures qui évalueraient leurs méthodes avec MAPLES-DR se réfèrent bien au même ensemble de test.

Il faut noter que, conformément aux instructions de partage de MESSIDOR, nous n'avons pas publié les images de fond d'œil ayant servi de base à nos annotations. Elles doivent être téléchargées manuellement depuis le site du consortium. Cependant, afin d'éviter à l'utilisateur l'effort d'extraction des régions d'intérêt depuis les images de MESSIDOR et de leur redimensionnement au format de MAPLES-DR , j'ai implémenté une librairie python dédiée à l'automatisation de cette tâche, sous réserve que l'utilisateur lui indique un chemin local vers l'archive de MESSIDOR. Cette librairie télécharge aussi automatiquement les annotations de MAPLES-DR, permet de les exporter localement dans le format souhaité, inclut les algorithmes de prétraitement d'images de fond d'œil et expose même les informations complémentaires sur l'annotation (préannotations, commentaires, temps, etc.). Elle est publiée sous la forme d'un paquet python disponible sur Github et pip et est documentée ici en anglais et en français !



FIGURE 4.11 Page d'accueil de la documentation du paquet `maples-dr`.

4.3 Évaluation de la fiabilité des annotations

Annoter à partir de cartes pré-segmentées par intelligence artificielle (IA) n'est pas un choix anodin lorsqu'on construit une base de données qui a justement pour vocation l'entraînement et la validation d'algorithmes supervisés. Ces cartes introduisent fatalement des biais dans les annotations qui risquent de se propager aux nouveaux modèles, compromettant leur acuité. Par ailleurs, quand bien même les rétinologues avaient segmenté sans préannotation, la variabilité inter-observateur naturelle entre eux est aussi une source d'incertitude qui doit être mesurée pour assurer la fiabilité des annotations de MAPLES-DR. Cette section évalue l'ampleur de ces biais et de cette variabilité. D'abord par l'étude des deux paires d'images dupliquées de MAPLES-DR, puis par la réannotation des structures pathologiques d'une dizaine d'images par 3 annotateurs, avec et sans préannotations.

4.3.1 Estimation de la variabilité inter-observateur sur les duplicatas

Chaque structure rétinienne annotée dans MAPLES-DR ne le fut que par un unique clinicien, excluant de facto toute étude de variabilité inter-observateur. Ou plutôt, cette étude aurait été exclue si 2 duplicatas ne s'étaient pas glissés par inadvertance dans la liste des images annotées. Cet accident permet une première estimation de la variabilité inter-observateur présente dans MAPLES-DR. Cependant, étant donné la taille très réduite de l'échantillon (2 images), nous restreindrons notre étude à une approche qualitative.

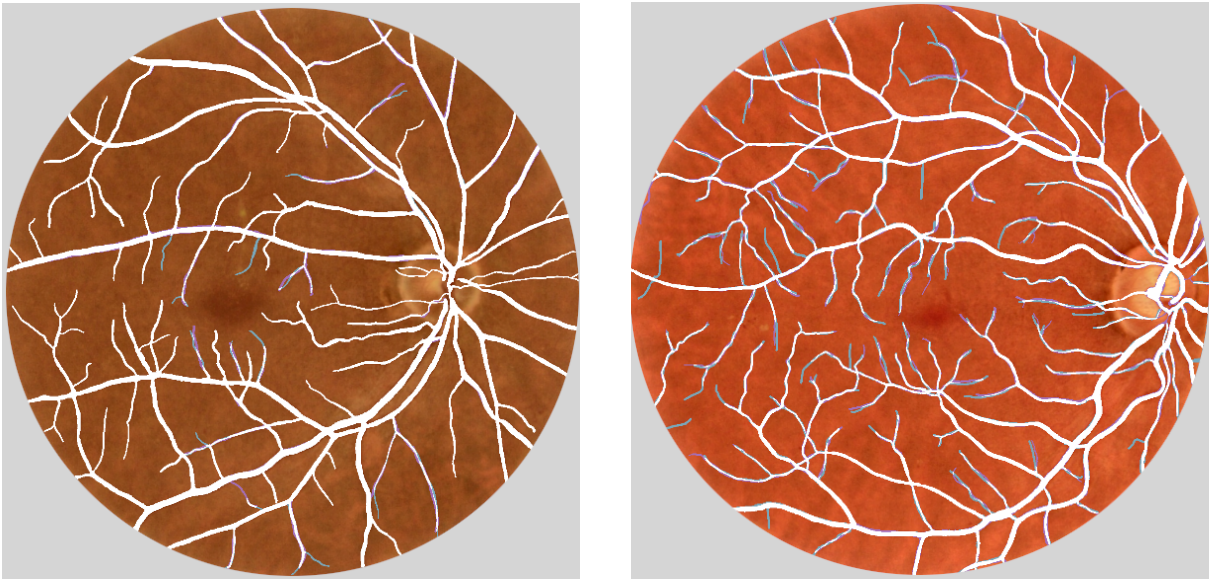


FIGURE 4.12 Visualisation comparative des annotations vasculaires pour les images dupliquées de MAPLES-DR. (*Blanc* : Pixels communs aux deux annotations, *Violet* : pixels inclus uniquement dans le jeu d'entraînement, *Bleu* : pixels inclus comme doublons)

Commençons avec les annotations vasculaires, puisque ce sont celles qui nous intéressent le plus pour la thèse. Les vaisseaux les plus larges font l'objet d'un consensus parfait entre les deux cartes de segmentation, probablement parce qu'ils proviennent directement des préannotations (cf. Figure 4.12). La variabilité intra-observateur [2] se manifeste plutôt sur les petits vaisseaux : souvent, ils sont annotés sur les deux cartes mais avec un tracé différent. Ces observations corroborent celles de Kai et al. [36] qui ont mesuré un dice intra-annotateur de 0.9679 en moyenne sur 40 images annotées deux fois par 5 annotateurs.

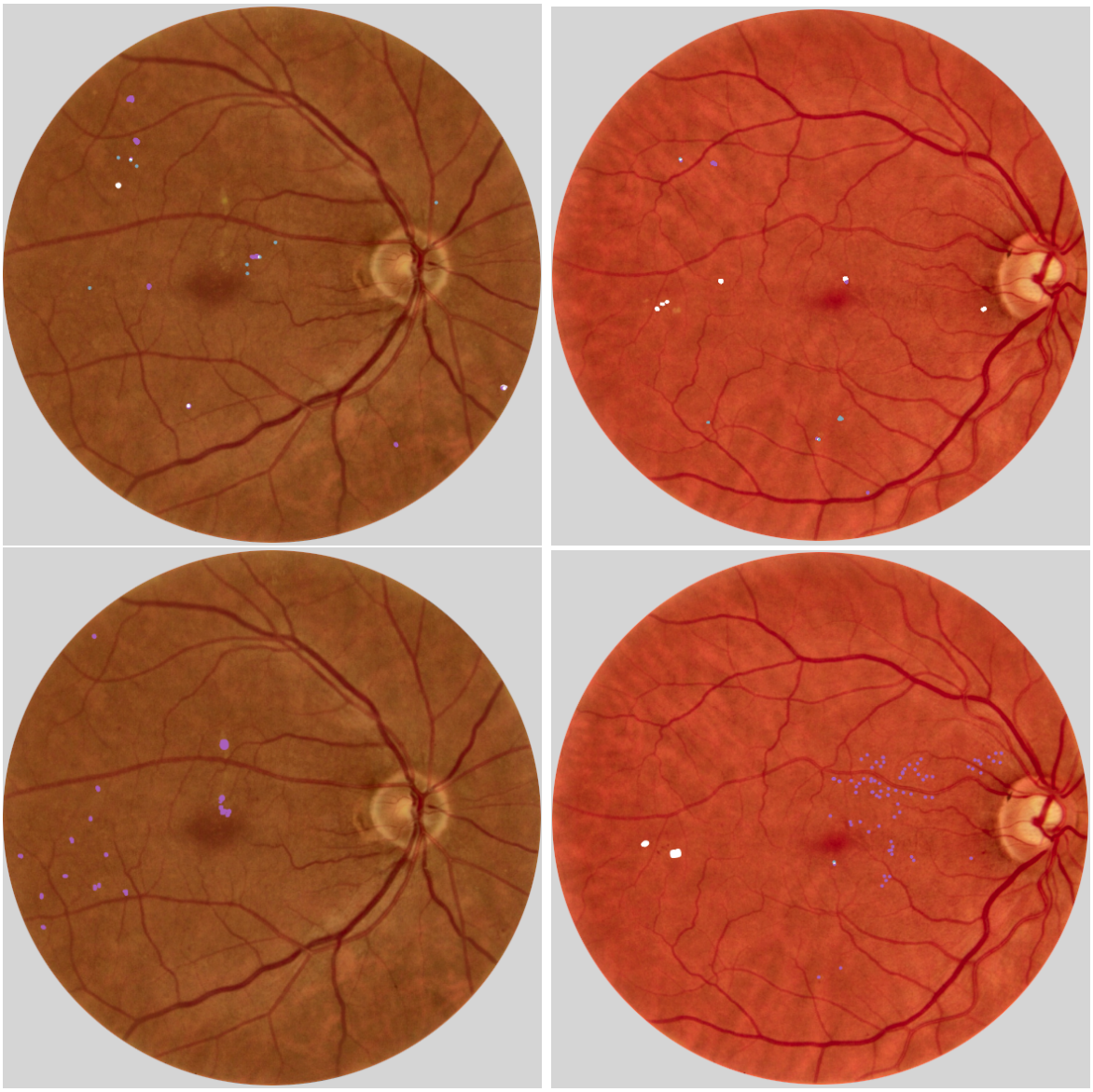


FIGURE 4.13 Visualisation comparative des lésions rouges (en haut) et lésions claires (en bas) pour les images dupliquées de MAPLES-DR. (Même légende que la Figure 4.12.)

2. Tous les vaisseaux ont été revus et corrigés par Dr. Duval. L'étude des doublons est donc dans ce cas une étude *intra*-observateur et non *inter*-observateur.

Pour les structures pathologiques, les observations sont radicalement différentes. Sur les 28 lésions rouges annotées par les rétinologues, seulement 13 sont communes aux deux. Pour les lésions claires, ce nombre tombe à 3 parmi 83 lésions annotées ! Au-delà de cette variabilité de détection des lésions, la visualisation comparative des annotations pathologiques (Figure 4.13) révèle des désaccords sur la segmentation des lésions communes, en particulier lorsque ces lésions ne sont pas présentes dans la carte de pré-segmentation. Ces résultats interpellent : ils appellent une étude approfondie de la variabilité d'annotation des structures pathologiques rétiniennes, avec et sans préannotation. Surtout que, à notre connaissance, aucun article n'a été publié sur le sujet à ce jour. C'est donc sur ces structures que nous avons concentré notre étude de variabilité inter-observateur.

4.3.2 Variabilité inter-observateur sans préannotation de lésions

Dans un premier temps, on s'intéressa à la segmentation manuelle en l'absence des cartes de préannotations. L'objectif était double : d'une part évaluer la variabilité inter-observateur sans elles, et d'autre part établir une segmentation de référence à laquelle on puisse comparer les annotations de MAPLES-DR obtenues avec elles. Pour cette première phase, on demanda donc à trois rétinologues d'annoter, chacun de leur côté, toutes les structures pathologiques de 51 images de MAPLES-DR. Nous aurions aimé conduire une étude analogue sur les vaisseaux mais, sachant que le temps nécessaire pour annoter les vaisseaux d'une image sans préannotations avoisine les deux heures [31], l'effort d'annotation fut jugé irréaliste.

La variabilité inter-observateur est de deux ordres. La première est liée aux désaccords de détection : lorsqu'une structure est considérée comme pathologique et annotée comme telle par un rétinologue, mais qu'elle est absente de la carte de segmentation produite par un autre. La seconde implique des désaccords de segmentation : une même structure est annotée par deux rétinologues, mais ses contours diffèrent. Ces deux types de variabilité doivent être traités individuellement. La variabilité de *détection* entre deux observateurs fut quantifiée par un score F1 défini sur le nombre de lésions plutôt que le nombre de pixels [3]. La variabilité de *segmentation* fut quant à elle évaluée par un calcul de l'intersection sur union (IoU) en ne considérant que les lésions conjointement annotées par les deux observateurs. La Figure 4.14 présente ces scores pour chaque type de lésion et chaque paire de rétinologues.

3. Le nombre de lésions annotées par un observateur O1 est tout d'abord calculé par un algorithme de composantes connectées. Ensuite, une lésion est considérée comme commune si au moins un de ses pixels a aussi été annoté par le second observateur O2. Puisque le nombre de lésions varie grandement d'un observateur à l'autre, le nombre de lésions annotées par O1 commune à O2 peut être très différent du nombre de lésions annotées par O2 commune à O1. Le score F1 est donc obtenu par moyenne harmonique des ratios – calculés indépendamment pour O1 et O2 – du nombre de lésions communes sur le nombre de lésions annotées. (En choisissant un observateur comme vérité terrain, ces ratios seraient équivalent à la precision et au recall.)

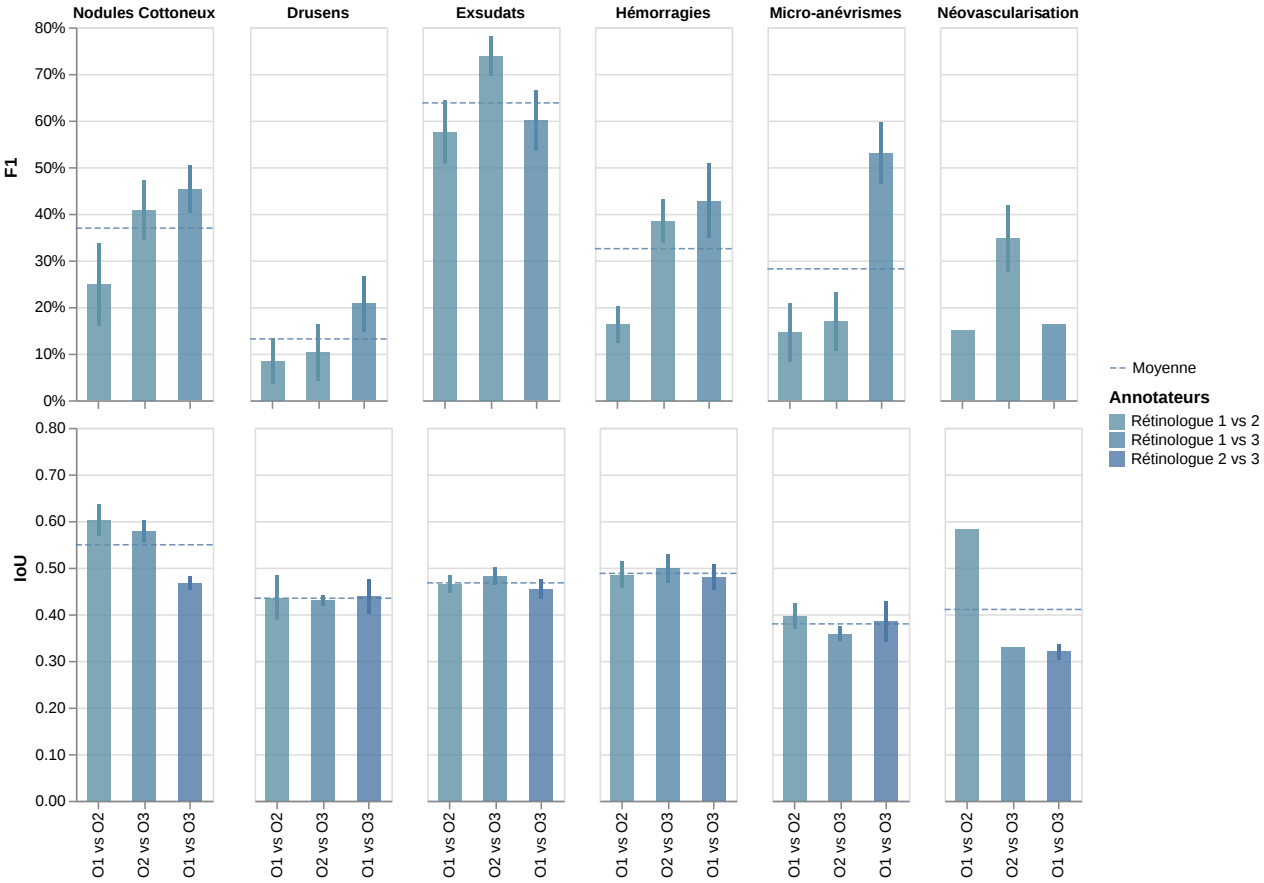


FIGURE 4.14 Variabilité inter-observateur lors de segmentation de lésions sans préannotation. (En haut : variabilité de détection ; en bas : variabilité de segmentation. Chaque paire d'annotateurs est évaluée individuellement. La ligne pointillée correspond à la moyenne des paires.)

Les rétinologues s'accordent le plus sur la détection des exsudats, mais le score F1 n'atteint que 61% en moyenne. Pour les micro-anévrismes, les hémorragies et les CWS, on mesure un accord moyen entre 28% et 33% et il tombe à 13,3% pour les drusens [4]. Une piste d'explication pour comprendre cet accord supérieur pour les exsudats et inférieur pour les drusens est à chercher dans leur fréquence respective (voir Figure 4.10). Les exsudats – symptômes courants de la DR – sont particulièrement présents dans les fonds d'œil de MAPLES-DR : 22 par image en moyenne ; alors que les drusens – plutôt associés à la DMLA – y sont rares : en moyenne 3 par image et la majorité n'en contiennent pas. Il est probable que la vigilance des annotateurs soit plus faible pour ces lésions rares, conduisant à une annotation partielle.

Contrairement à celle de détection, la variabilité de segmentation est homogène parmi les types de lésions (si l'on exclut celles qui n'ont pas été simultanément identifiées par les deux observateurs). Quelle que soit la paire de rétinologues considérée, l'IoU varie entre 0.4 et 0.6.

4. Nous n'analyserons pas les néovaisseaux qui ont été identifié sur trop peu d'images pour obtenir des résultats significatifs.

Résumons par des ordres de grandeur mesurés sur toutes les lésions confondues : lorsqu'elles sont annotées sans préannotation, environ un tiers des lésions identifiées par chaque rétinologue ne l'est par aucun des deux autres (en excluant les drusens pour lesquels ce rapport est plutôt de deux tiers), et un quart des lésions annotées par au moins deux rétinologues n'est pas annoté par le troisième. Pour une même lésion annotée par deux observateurs, en moyenne un pixel sur deux n'est pas commun. Pour plus de détails, un billet de blog analyse dans le détail et par de multiples graphes interactifs les résultats de cette étude.

Ces niveaux de variabilité étant évidemment inattendus, nous les avons présentés aux rétinologues pour tenter de comprendre leurs causes. Au-delà de la variabilité naturelle, trois facteurs potentiels ont été identifiés :

1. Les prétraitements d'amélioration du fond d'oeil n'ont pas été utilisés par tous les annotateurs. Ceux qui les ont activés ont eu tendance à annoter plus de lésions.
2. La distinction entre micro-anévrismes et les petites hémorragies est difficile à partir du fond d'oeil simple et nécessite normalement l'angiographie. Une partie de la variabilité de détection de ces lésion est donc due à des désaccords sur leur type. En considérant indifféremment les deux lésions rouges comme une classe commune, le score F1 moyen pour la détection de ces lésions passe de 28,2% et 32,5% (respectivement pour les micro-anévrismes et les hémorragies) à 58,4%.
3. L'absence de consignes spécifiques a conduit à des styles d'annotation différents d'un annotateur à l'autre : certains annotant chaque lésion précisément alors que d'autres détouraient globalement les amas.

Il est en tous cas certain que, compte tenu de cette variabilité inter-observateur, toute tentative pour extraire une vérité terrain unique à partir des triples cartes de segmentation ne produirait pas un résultat fiable.

4.3.3 Variabilité inter-observateur avec préannotation de lésions

La précédente section conclut à une variabilité inter-observateur très élevée lors de l'annotation de lésions à partir de zéro. Qu'en est-il lors d'annotation par correction de cartes pré-segmentées ? Pour le savoir, nous avons reproduit l'expérience en fournissant des préannotations aux trois rétinologues. Un modèle de segmentation plus récent fut utilisé de sorte qu'on ajouta les nodules cotonneux à la liste des lésions préannotées de MAPLES-DR : i.e. les micro-anévrismes, les hémorragies et les exsudats. Les drusens et les néovaisseaux furent écartés de l'étude faute de modèle pour les segmenter. Les métriques de variabilité de détection et de segmentation de ces lésions sont présentées sur la Figure 4.15. Elles sont spectaculairement plus élevées que celles obtenues sans préannotation.

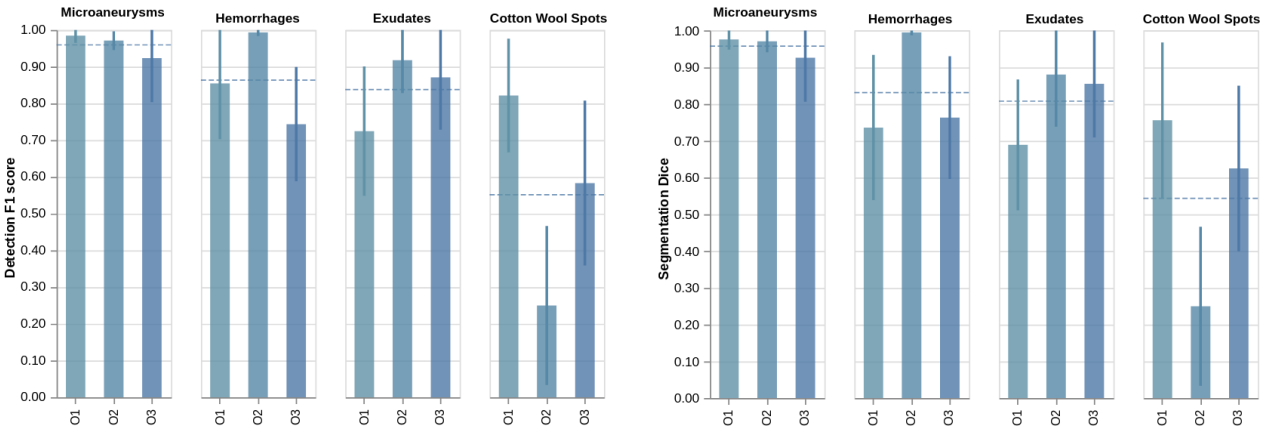


FIGURE 4.15 Variabilité Inter-observateur lors de segmentation de lésions **avec** préannotation. (À gauche : variabilité de détection ; à droite : variabilité de segmentation.)

Deux phénomènes concourent à ce résultat : **1.** les lésions correctement détectées par l'IA ont exactement la même segmentation pour tous les rétinologues qui les ont conservées. **2.** Bien qu'aucune instruction n'ait été donnée aux rétinologues sur le niveau de détail attendu lors de la segmentation manuelle d'une lésion, les lésions préannotées sont autant d'exemples de segmentations souhaitées, agissant comme une consigne implicite et harmonisant le « style » des annotations ajoutées par les différents annotateurs.

Seuls les nodules cotonneux semblent conserver une variabilité inter-observateur élevée malgré les préannotations. Cette différence s'explique par la qualité et le faible nombre de préannotations pour cette lésion. Sur les 16 images, seuls 14 CWS furent préannotés par le modèle et un seul fut considéré valide par les trois rétinologues. Les autres furent retirés manuellement, mais pas par tous les rétinologues simultanément, entraînant un taux élevé de désaccord.

En bref, annoter les structures pathologiques du fond d'œil en corrigeant des cartes pré-segmentées diminue significativement la variabilité inter-observateur en introduisant une référence commune pour tous les annotateurs. Toutefois, l'ampleur de cette réduction repose fortement sur la bonne qualité du modèle de pré-segmentation. Les différences importantes observées sur les deux duplicatas de MAPLES-DR résultent ainsi certainement de la moins bonne qualité du modèle initialement utilisé en 2018.

4.3.4 Biais introduit par la préannotation des lésions

Le recours aux préannotations est un moyen efficace pour réduire la variabilité inter-observateur lors de l'annotation des lésions rétiniennes. Mais si la moitié des hémorragies ou micro-anévrismes de MAPLES-DR sont directement issus de ces préannotations, quels biais introduisent-elles dans les annotations finales ? Idéalement, les annotations de zéros auraient permis de construire une segmentation de référence pour répondre à cette question, mais elles diffèrent

trop d'un rétinologue à l'autre pour définir une vérité terrain fiable (voir section 4.3.2). Nous nous rabattons donc sur une analyse qualitative du style d'annotation qui compare les distributions de taille des lésions et de leur nombre par image pour cinq annotateurs : *O1*, *O2*, et *O3* sont les rétinologues seuls, *CP* est un rétinologue avec préannotation (i.e. la version de MAPLES-DR), et *IA* sont les préannotations avant correction.

Ces distributions sont présentées par la Figure 4.16. Elles confirment la disparité de style présentée à la section 4.3.2 entre les rétinologues : les lésions dessinées par O2 sont en moyenne deux à quatre fois plus grandes que celles de O1 et O3, et O1 a annoté en moyenne deux fois plus de lésions que O2 ou O3. Concernant les biais liés aux pré-segmentations : la filiation entre les préannotations (IA) et les corrections (CP) est évidente tant par le nombre de lésions que par leurs tailles : elles sont plus nombreuses et de tailles plus diverses que celles annotées uniquement par les rétinologues. Mais les annotations CP portent aussi la trace des corrections : leurs tailles s'alignent sur O1 ou O3, et leur nombre sur O1. En résumé, les préannotations biaisent les annotations vers une sensibilité accrue. En l'absence d'une vérité terrain fiable, difficile de dire si elle est synonyme d'exhaustivité ou de faux positifs...

Enfin, un pic de lésions de 78 pixels est visible pour O1, O3 et CP. Cette taille correspond en réalité à l'outil d'annotation par défaut (un cercle de diamètre 10 pixels) et semble correspondre à la taille des plus petits détails annotés par les cliniciens. Bien qu'anodin, ce détail souligne l'importance sous-estimée du choix des paramètres par défaut.

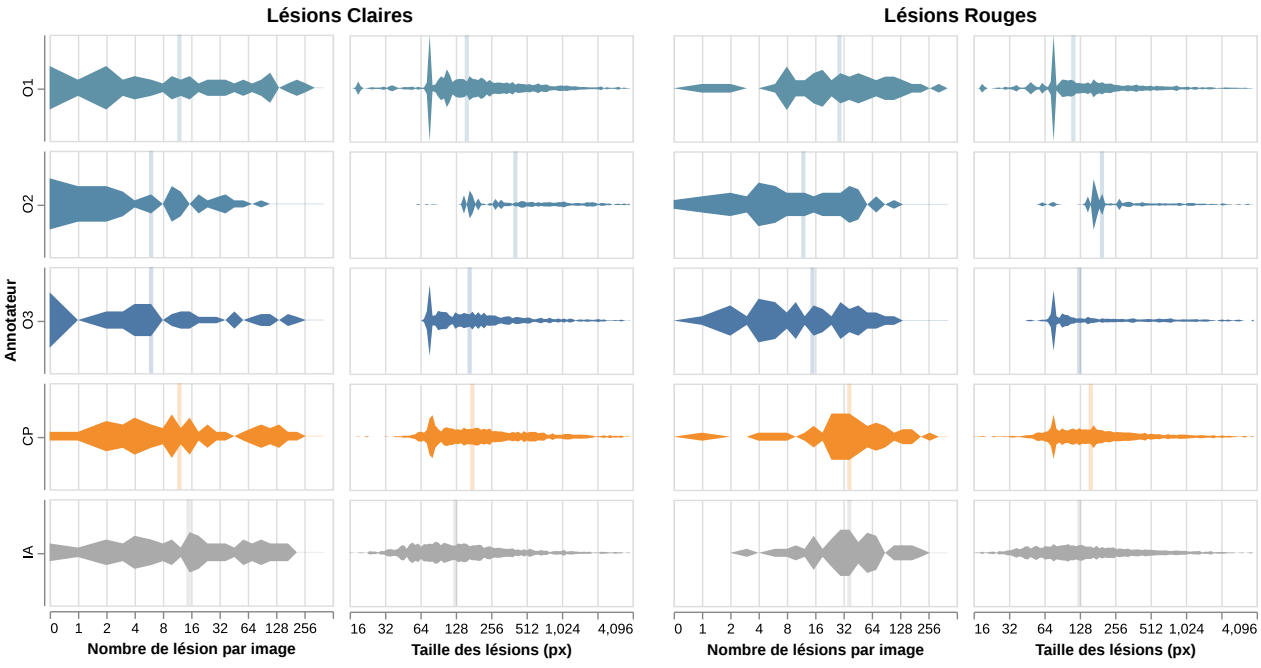


FIGURE 4.16 Distribution de la taille et du nombre de lésions par images avec et sans préannotation. (*O1*, *O2*, *O3* : annotations manuelles sans préannotations ; *CP* : annotation par correction de préannotations ; *IA* : préannotations avant corrections. La barre verticale indique la valeur médiane.)

4.3.5 Entraînement de modèles de segmentations avec MAPLES-DR

Afin d'évaluer la pertinence de MAPLES-DR pour l'entraînement de modèles de segmentations, nous avons entraîné cinq modèles à segmenter les lésions rétiniennes (i.e. micro-anévrismes, hémorragies, exsudats et CWS) sur des jeux de données publics différents : MAPLES-DR, IDRiD [195], DDR [196], FGADR [193], et Retinal Lesions [194]. Tous furent construits à partir du même protocole rudimentaire : un modèle U-Net simple est entraîné à segmenter conjointement les quatre lésions rétiniennes par minimisation d'une fonction de coût Dice en utilisant la traditionnelle descente de gradient stochastique (SGD) comme optimiseur. Pour accélérer l'apprentissage, les poids initiaux de l'encodeur sont pré-entraînés sur ImageNet. Tous les modèles furent ensuite évalués sur les ensembles de tests individuels associés à chacun de ces jeux de données. Les mIoU mesurés pour chaque combinaison d'ensembles de formation et de test sont compilés dans le tableau 4.1.

Pour chaque ensemble de test, le modèle qui obtient les meilleures performances est celui qui a été entraîné sur l'ensemble d'entraînement correspondant. Ce résultat est attendu étant donné l'écart entre les jeux de données : tant en termes des caractéristiques d'images de fond d'œil (ethnicité, qualité et résolution de l'image, colorimétrie, etc.) que du style d'annotation (segmentations grossières pour Retinal Lesions, précises pour MAPLES-DR, IDRiD et DDR ou mélange des deux pour FGADR). Malgré ces divergences, un modèle entraîné sur un jeu de données peut, dans une certaine mesure, se généraliser aux autres. Nous avons mesuré cette capacité en classant les modèles par mIoU sur chaque jeu de données, puis en calculant la moyenne des classements. Le modèle entraîné sur MAPLES-DR obtient le deuxième meilleur classement moyen de 2.8, derrière IDRiD, dont le classement moyen est de 2.2. Ces deux jeux de données sont pourtant ceux contenant le plus petit nombre d'images d'entraînement, mais leurs images de fond d'œil de bonne qualité et leurs annotations précises de chaque lésion semblent particulièrement profiter aux capacités de généralisation des modèles.

Tableau 4.1 Performance en généralisation de modèles de segmentation de lésions par ensemble d'entraînement. Les jeux de données d'entraînement sont triés en fonction de leur classement moyen sur tous les ensembles de test.

Jeu d'entraînement	Performance par ensemble de test (mIoU)					Rang
(nb. image)	IDRiD	MAPLES-DR	DDR	FGADR	Ret. Les.	Moyen
IDRiD (54)	**57.6%**	37.5%	35.6%	28.4%	26.0%	2.2
MAPLES-DR (118)	39.7%	**46.8%**	31.0%	26.9%	27.8%	2.8
DDR (450)	48.6%	35.4%	**43.0%**	26.4%	24.8%	3.2
FGADR (1096)	33.0%	28.8%	31.7%	**47.4%**	25.5%	3.4
Retinal Lesions (948)	28.6%	27.3%	28.6%	28.4%	**48.2%**	3.4

4.4 Discussion et Contributions

Le choix de corriger des cartes pré-segmentées plutôt que d'annoter de zéro, bien qu'exigé par l'ambition de la tâche de segmentation, comportait le risque d'introduire un biais dans les annotations collectées. Mais les différentes expériences d'évaluation de MAPLES-DR semblent valider ce choix : si un biais a été introduit, il a accru la sensibilité d'annotation. Elles montrent surtout que l'annotation par correction de préannotations a permis de diminuer significativement la variabilité inter-observateur, en particulier lors de l'annotation des lésions. Dans l'ensemble, le dyptique préannotation et outils d'annotation web s'est révélé efficace : réduisant le temps nécessaire à la segmentation des vaisseaux de 2 heures [31] à 6 minutes en moyenne !

Au début du projet, nous imaginions que les seuls écueils à anticiper seraient liés à la conception des outils d'annotation, mais si le projet fut si long, c'est que des obstacles apparurent ensuite là où nous ne les attendions pas. Les valeurs de variabilité inter-observateur en détection et en segmentation de lésions nous ont notamment pris par surprise. À notre connaissance, il n'en existe aucune mention dans la littérature. Nous-mêmes ne l'avons pas inclus dans l'article de MAPLES-DR [5], préférant y consacrer un article dédié ultérieurement. À mon sens, elles mettent au jour un sérieux problème d'alignement entre les objectifs théoriques de la segmentation des structures rétiniennes (la détection des lésions pour le diagnostic de pathologie, l'identification des vaisseaux pour en extraire la géométrie ou la topologie, etc.) et la fonction objectif optimisée en pratique par les modèles de segmentation pendant leur entraînement. Cette dernière prête en effet beaucoup d'importance aux pixels en pourtour des lésions et des vaisseaux. Or, si l'on en croit les résultats de l'étude de variabilité, les contours de ces structures ne semblent pas centraux dans la pratique médicale, ou en tous cas ne font pas l'objet d'un consensus sur leur positionnement exact. À l'inverse, la fonction objectif accorde trop peu de considérations à des aspects pourtant cruciaux comme la continuité de la topologie vasculaire.

Il reste néanmoins que la qualité des annotations de MAPLES-DR offre aux modèles qu'elles entraînent de bonnes capacités de généralisation et constitue une base solide de pré-entraînement [202]. Et au-delà des modèles de segmentation, l'exhaustivité inédite des types de structures annotées est précieuse pour ouvrir la voie à des approches qui combinent ces structures pour un dépistage interprétable de la DR (Legault et al., 2025 [203]).

5. Nous signalons tout de même dans la section **Usage Note** du papier que les CWS, les drusens et les néovaisseaux sont insufissament représentés dans MAPLES-DR pour constituer un jeu d'entraînement fiable, et que ces annotations seraient plus propices à servir de base de pré-entraînement.

Concernant cette thèse, le projet MAPLES-DR apporte deux contributions. D'une part, les 198 nouvelles annotations des structures anatomiques rétiniennes rejoignent et diversifient celles des autres jeux de données publics, contribuant ainsi à l'amélioration des modèles de segmentation des vaisseaux, du disque et de la macula. D'autre part, j'ai pu tirer profit de la plateforme d'annotation pour distinguer les artérioles des veinules sur 100 cartes de segmentation vasculaire de MAPLES-DR. Le prochain chapitre exploite ces cartes pour entraîner une nouvelle formulation de neurones convolutifs : les *steered* CNN.



CHAPITRE 5 STEERED CNN : L'ÉQUIVARIANCE PAR ROTATION AU SERVICE DE LA CLASSIFICATION DES ARTÈRES ET VEINES RÉTINIENNES

Le formalisme des *SCN* (ou *Steered* CNN) présenté dans ce chapitre modifie la définition des neurones convolutifs des CNN pour y inscrire l'équivariance à la rotation. Cette nouvelle définition vise à réduire le nombre de paramètres nécessaires à ces modèles et espère ainsi améliorer leur capacité de généralisation.

Ce chapitre est probablement le plus mathématique de la thèse, il suit le formalisme énoncé au début de ce document.

5.1 Motivations et travaux connexes

5.1.1 Motivations à l'équivariance par rotation

Dans le chapitre sur les réseaux convolutifs [204] issu de leur manuel « Deep Learning » publié en 2016, Ian Goodfellow, Yoshua Bengio et Aaron Courville motivent l'efficacité des CNN pour le traitement d'images par une triple propriété que leur confère l'opérateur convolutif : l'équivariance par translation, la faible densité de connectivité (*sparse interaction* et *connectivity*), et la mutualisation des paramètres (*parameter sharing*). En effet, parce que la taille des filtres de convolution est très inférieure à celle de l'image analysée, chaque pixel prédit par le modèle ne dépend que d'une portion de l'image centrée sur ce pixel (*sparse interaction*), réduisant considérablement le temps de calcul nécessaire à l'inférence ou l'entraînement d'un CNN comparativement à un modèle complètement connecté. En outre, puisque les mêmes filtres de convolution sont appliqués à toutes les régions de l'image d'entrée (*parameter sharing*), le nombre de paramètres est non seulement restreint mais aussi indépendant de la résolution de l'image, à contrario des modèles complètement connectés qui attribuent un jeu de paramètres à chaque position du vecteur d'entrée. L'empreinte mémoire lors de l'utilisation et du stockage des modèles convolutifs est donc faible. En outre, le modèle est équivariant par translation : la translation de son image d'entrée induit une translation analogue de sa prédiction ; ou dit autrement : le modèle n'est pas sensible aux coordonnées absolues des objets qu'il segmente. Cette indifférence à la position est particulièrement pertinente pour la segmentation de vaisseaux qui se déploient sur la totalité de l'image. Mais le principal avantage de la réutilisation intensive des filtres de convolution au sein d'un CNN est ailleurs : elle impose à chaque paramètre de participer aux traitements de l'ensemble des pixels de

l'image. Cette mutualisation des fonctions d'un paramètre agit à la fois comme régularisation et comme dopant de son efficacité statistique : lors d'un entraînement, il voit en effet sa valeur ajustée non pas une fois par image mais une fois par pixel de celle-ci !

Cependant, tous ces bénéfices des CNN se limitent aux transformations de translation et ne s'appliquent pas aux rotations. Rien n'assure que la segmentation ou la classification d'un objet (par exemple, un vaisseau) ne varie avec l'orientation sous laquelle il est présenté. La Figure 5.1 montre d'ailleurs des variations dans la classification A/V de certains petits vaisseaux lorsqu'on fait subir des rotations à l'image de fond d'œil présentée au CNN. Pire, la littérature de classification vasculaire pré-apprentissage profond obtient de très bons résultats en analysant le profil des vaisseaux [78, 82], c'est-à-dire en analysant des caractéristiques particulièrement anisotropiques ! Or, pour être équivariant par rotation, un CNN dont la prédiction reposerait sur de telles caractéristiques devrait nécessairement consacrer plusieurs filtres de convolution à la détection d'une seule de ces caractéristiques, mais dans toutes ses orientations. Plus exactement, on peut imaginer que, lors de la phase d'apprentissage, les échantillons d'entraînement présentant des vaisseaux horizontaux induisent une mise à jour des paramètres très différentes de ceux présentant des vaisseaux analogues mais verticaux. Ces différences peuvent permettre la spécialisation des filtres de convolution vers chaque orientation de vaisseaux, mais au prix de plus nombreuses itérations et échantillons d'entraînement avant que le modèle ne converge. Inscrire l'équivariance par rotation dans le formalisme des

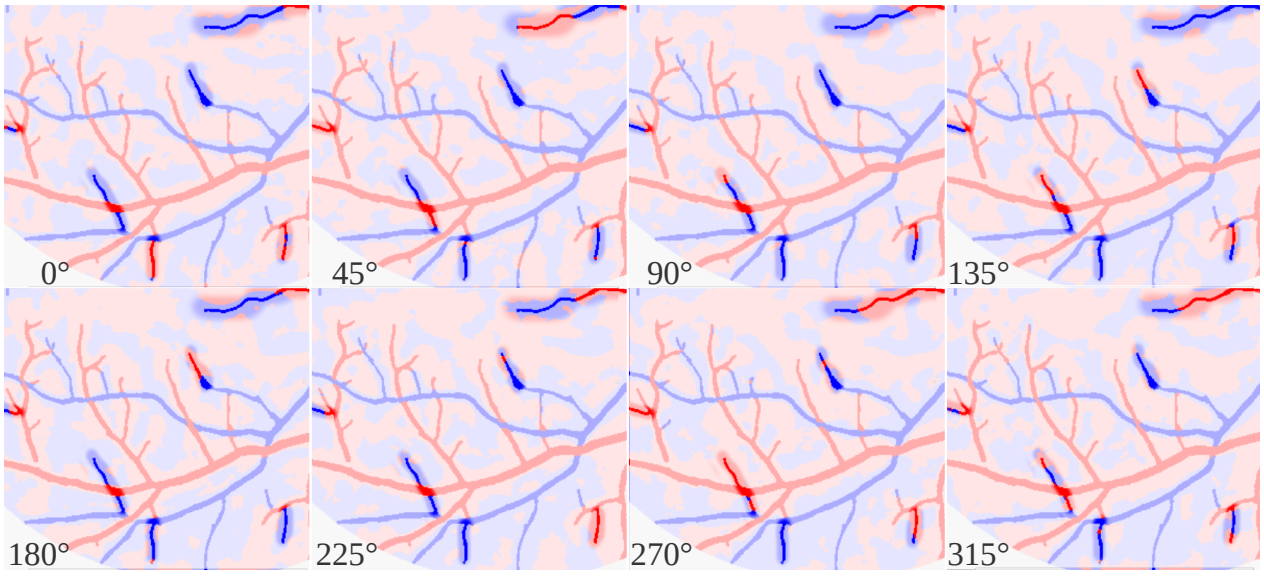


FIGURE 5.1 Variations de la classification A/V prédite par un même CNN, pour différentes rotations initiales de l'image d'entrée

(Pour simplifier leur comparaison, les vaisseaux variants sont en surbrillance et les cartes présentées ont subi un rotation inverse à celle infligée à l'image fond d'oeil.)

CNN devrait permettre l'économie de cette déduplication des filtres de convolution et ainsi la réduction de la taille du modèle et du temps de calcul nécessaire à son entraînement. L'intuition communément admise du rasoir d'Ockham laisse même espérer que cette diminution du nombre de paramètres augmente d'autant les capacités de généralisation du modèle.

5.1.2 CNN et équivariance par rotation

De nombreuses tâches de segmentation biomédicale (dont la segmentation des vaisseaux ou leur classification en artères ou veines) sont par essence équivariantes par rotation. Puisque rien dans l'implémentation standard des CNN n'assure cette propriété, elle doit nécessairement être apprise par le modèle lors de l'entraînement. Pour favoriser cet apprentissage, une première solution consiste à intégrer une augmentation géométrique à la boucle d'entraînement : chaque image de fond d'œil subit une rotation d'un angle aléatoire et la même transformation est appliquée à la vérité terrain. Durant l'entraînement, le modèle apprendra ainsi à classifier les vaisseaux, quelle que soit leur orientation. Cette solution naïve est utilisée par la quasi-totalité des travaux appliquant les réseaux de neurones à la segmentation sémantique du réseau rétinien, bien qu'elle soit sujette aux limitations énoncées à la section précédente. Il existe pourtant des travaux qui visent à améliorer l'équivariance par rotation des CNN, qu'on peut regrouper en deux familles.

La première famille s'inspire de l'augmentation géométrique par rotation des images en entrée du modèle. Cheng et al. [205] proposent d'inciter le réseau à conserver constante sa représentation d'une image donnée, alors qu'elle lui est présentée sous des orientations variables. En pratique, cette incitation prend la forme d'un terme de régularisation ajouté à la fonction de coût qui fait converger les caractéristiques latentes calculées à partir de différentes versions d'une même image. Une autre approche est proposée par Laptev et al. [206] : elle repose aussi sur la démultiplication par rotation des images présentées au modèle, mais étend la procédure à la phase d'inférence. Pour obtenir une prédiction invariante à l'orientation, le modèle analyse successivement l'image selon une série d'orientations, puis, parmi toutes les prédictions obtenues, il ne conserve que celle avec la probabilité la plus forte. Ces deux approches affichent des gains de performance pour la reconnaissance d'objets et ne nécessitent presque aucune modification des architectures existantes. Cependant, si leur invariance par rotation s'avère utile pour la classification d'images, elle n'est pas applicable dans le contexte de la segmentation sémantique où chaque objet a sa propre orientation locale.

La seconde famille abandonne la rotation des échantillons d'entraînement et propose de l'appliquer plutôt aux noyaux de convolution eux-mêmes. Chaque noyau est pivoté dans n orientations avant d'être convolué avec sa carte d'entrée, produisant ainsi n cartes de sortie. Ces

multiples versions de chaque carte sont soit fusionnées immédiatement après chaque couche convolutive par un max-pooling [207], soit maintenues tout au long du réseau – chaque version n'étant convoluée qu'avec des noyaux orientés dans une unique direction à l'aide de *group convolution* – et regroupées à la dernière couche [208, 209]. Lors de l'entraînement, les mises à jour obtenues pour chaque orientation d'un noyau sont moyennées avant de lui être appliquées. De cette façon, bien que chacune ait son orientation propre, les différentes versions d'un même noyau restent toujours cohérentes entre elles. Autrement dit, les méthodes de cette seconde famille implémentent l'équivariance à la rotation par la mutualisation des paramètres sur les différentes orientations ! Et comme on pouvait s'y attendre, elles démontrent une amélioration impressionnante de l'efficacité des paramètres pour les tâches de classification et de segmentation sémantique. Cependant, les travaux cités plus haut réalisent la rotation des noyaux de convolution par une interpolation bilinéaire : cette méthode est simple à implémenter mais est peu efficace en temps de calcul et produit des artefacts pour tout angle de rotation qui n'est pas un multiple de 90°. Une solution à ce problème se trouve avec les *steerable filters*.

5.1.3 Steerable Filters : des filtres de convolution orientables

En 1991, Freeman et al. [210] proposent une méthode pour convoluer efficacement une image avec de multiples versions d'un même filtre, chacune orientée différemment. Pour ce faire, ils introduisent une nouvelle famille de filtres convolutifs : les *Steerable Filters* qui sont définis par une combinaison linéaire d'une série de *filtres élémentaires* et qui peuvent être orientés dans une direction arbitraire en modifiant la pondération de la combinaison linéaire tout en maintenant les filtres élémentaires constants.

Plus formellement, un filtre polaire $\Psi(\rho, \theta)$ est dit *steerable* si sa rotation par un angle α (notée $\mathrm{Rot}_\alpha\Big[\Psi(\rho, \theta)\Big] = \Psi(\rho, \theta - \alpha)$) peut s'écrire comme la combinaison linéaire de H filtres élémentaires (notés $\mathbf{\Psi}_h$) par des coefficients linéaires qui sont fonctions de α (notés $a_h(\alpha)$) :

$$\Psi(\rho, \theta - \alpha) = \sum_{h=1}^{H} a_h(\alpha)\ \mathbf{\Psi}_h(\rho, \theta) \tag{5.1}$$

Grâce à la reformulation de la rotation de ces filtres, la convolution d'une image $\mathbf{X}$ avec un *steerable filter* orienté dans une direction arbitraire α se factorise comme :

$$\mathbf{X} * \mathrm{Rot}_\alpha\Big[\mathbf{\Psi}\Big] = \sum_{h=1}^{H} \Big(a_h(\alpha) \ \cdot \ \mathbf{X} * \mathbf{\Psi}_h\Big) \tag{5.2}$$

Autrement dit, après avoir convolué en amont l'image avec le jeu prédéfini de filtres élémentaires $\mathbf{\Psi}_h$, on peut, par combinaison linéaire des résultats avec les coefficients appropriés, simuler la rotation du filtre dans autant de directions que désirées.

Considérons un exemple trivial : le filtre $\mathbf{W}(\theta) = \cos(\theta)$, dont on omettra temporairement la partie radiale par simplicité. Le développement trigonométrique de sa rotation donne :

$$\begin{aligned}\mathbf{W}(\theta - \alpha) &= \cos(\theta - \alpha) \\ &= \cos(\alpha)\cos(\theta) + \sin(\alpha)\sin(\theta) \\ &= a_1(\alpha) \cdot \mathbf{\Psi}_1(\theta) + a_2(\alpha) \cdot \mathbf{\Psi}_2(\theta)\end{aligned}$$

Ce filtre admet donc une décomposition en deux filtres élémentaires $\mathbf{\Psi}_1(\theta) = \cos(\theta)$ et $\mathbf{\Psi}_2(\theta) = \sin(\theta)$, dont la combinaison linéaire par les coefficients $a_1(\alpha) = \cos(\alpha)$ et $a_2(\alpha) = \sin(\alpha)$ décrit bien la rotation de $\mathbf{W}$. En précalculant les convolutions avec ces filtres élémentaires ($\mathbf{Y}_1 = \mathbf{X} * \mathbf{\Psi}_1$ et $\mathbf{Y}_2 = \mathbf{X} * \mathbf{\Psi}_2$) on peut déduire la convolution de l'image $\mathbf{X}$ par le filtre $\mathbf{W}$ après rotation d'un angle α :

$$\mathbf{X} * \mathrm{Rot}_\alpha\Big[\mathbf{W}\Big] = \cos(\alpha) \cdot \mathbf{Y}_1 + \sin(\alpha) \cdot \mathbf{Y}_2$$

La notation en nombres complexes est particulièrement adaptée aux *steerable filters*. En effet, dans ce domaine, la rotation d'un filtre – c'est-à-dire le déphasage de sa composante polaire – se traduit par un facteur du type $\mathrm{e}^{-\mathrm{i}\alpha}$. On peut par exemple montrer que tout filtre de la forme $\Psi(\rho, \theta) = \mathrm{Re}\Big[\ G(\rho)\mathrm{e}^{\mathrm{i}k\theta}\ \Big]$ est *steerable* et que sa rotation est décrite par :

$$\Psi_k(\rho, \theta - \alpha) = \cos(k\alpha)\ \mathrm{Re}\Big[\ G(\rho)\mathrm{e}^{\mathrm{i}k\theta}\ \Big] \ + \ \sin(k\alpha)\ \mathrm{Im}\Big[\ G(\rho)\mathrm{e}^{\mathrm{i}k\theta}\ \Big] \tag{5.3}$$

5.1.4 Steerable CNN

En 2016, Cohen et al. [211] étendent leur propre travail [208] et proposent l'utilisation de *steerable filters* pour implémenter la rotation des noyaux convolutifs. Pour ce faire, ils reparamétrisent les noyaux de convolution de leur *Group Equivariant Convolutional Layer* par une combinaison linéaire de *steerable filters* et inventent les *Steerable CNN*. Ce formalisme a depuis été modifié et reformulé à plusieurs reprises [212, 213], et a même été étendu pour implémenter l'équivariance à des transformations de symétries axiales ou de redimensionnement [214–216]. L'architecture de base reste cependant toujours la même : chaque carte de caractéristiques est calculée selon plusieurs orientations jusqu'à la couche finale, où une orientation pour chaque pixel et chaque caractéristique est sélectionnée par max-pooling. Ces architectures atteignent les performances de l'état de l'art pour la classification d'objet alors qu'elles utilisent considérablement moins de paramètres.

Cependant, leur implémentation de l'équivariance est partielle car les caractéristiques ne sont calculées que pour n orientations. En outre, sélectionner l'orientation la plus pertinente selon les probabilités de classification rend cette décision sujette aux fluctuations de prédiction du modèle. Dans le cas de la classification des vaisseaux en particulier, leur orientation peut être directement dérivée du squelette vasculaire. Les *steered* CNN proposés dans ce chapitre partagent avec les *Steerable CNN* le reparamétrage des filtres de convolution par des « steerable filters », mais contrairement à ces derniers qui appliquent leurs filtres de convolution dans une série d'orientations prédéfinies, les *steered* CNN alignent dynamiquement leurs noyaux de convolution sur les orientations locales des vaisseaux.

5.2 Étude préliminaire : isotropismes des CNN classifieurs de vaisseaux

On l'a vu, la méthode la plus populaire pour réaliser l'équivariance par rotation des CNN standard reste l'augmentation de données géométriques : en faisant subir une même rotation aléatoire à l'image d'entrée et à la vérité terrain. On peut cependant s'interroger sur l'efficacité de cette méthode et ses implications sur les filtres de convolution appris.

Par ailleurs, les méthodes qui implémentent l'équivariance par rotation dans les CNN sont généralement motivées par l'existence d'un gaspillage de paramètres pour adapter le modèle aux différentes orientations. Mais la véracité et l'ampleur de ce postulat n'ont, à ma connaissance, jamais été vérifiées. Avant d'exposer le fonctionnement des *steered* CNN, il me paraît important d'évaluer combien de paramètres un modèle de classification d'artères-veines consacre-t-il à apprendre les mêmes caractéristiques dans plusieurs orientations ? On pourrait être surpris de la réponse...

Mesure du tropisme des neurones

La question n'est pas triviale : les fonctions modélisées par une couche convolutive ne sont pas soigneusement rangées dans des noyaux individuels. Elles émergent plutôt de l'interaction de plusieurs filtres entre eux, parfois combinés avec des filtres de la couche précédente ou de la couche suivante. Comment identifier alors, parmi les poids du modèle, une fonction spécifique et ses répliques orientées différemment ? Une réponse indirecte peut être trouvée par une question plus simple : dans un modèle globalement équivariant par rotation (par exemple grâce à une augmentation de données géométriques), combien de neurones ne sont pas isotropiques ? En effet, pour que la prédiction du modèle reste constante lors d'un changement d'orientation de l'image, de tels neurones doivent nécessairement compenser mutuellement les variations de leur prédiction respective. D'une manière ou d'une autre, ces neurones doivent se relayer pour assurer collectivement des fonctions qu'ils ne pourraient assurer seuls sans briser l'équivariance du modèle. Ainsi, en mesurant la différence entre l'équivariance globale du modèle et l'isotropisme local de ses neurones, on peut déduire la proportion de paramètres présentant une redondance liée à l'orientation. Cette section décrit une version polaire de la densité spectrale de puissance comme mesure du tropisme (ou de l'isotropisme) des neurones convolutifs d'un CNN.

Soit $\mathrm{I}(i,j)$ une image présentée à l'entrée d'un réseau, et $\phi : \mathbf{I} \to \mathbf{y}$ l'application qui calcule la prédiction $\mathrm{y}(i,j)$ d'un neurone du réseau. On définit $\mathbf{\Phi}(\theta) = \mathrm{Rot}_{-\theta}\Big[\ \phi\Big(\mathrm{Rot}_{\theta}\big[\mathbf{I}\big]\Big)\ \Big]$, où la carte $\mathbf{\Phi}(\theta)$ est obtenue en présentant au réseau une rotation de l'image $\mathrm{I}(i,j)$ par un angle θ et en appliquant à la prédiction du neurone une contre-rotation de $-\theta$ (voir Figure 5.2). Dit autrement, $\mathbf{\Phi}(\theta)$ est la carte que l'on obtiendrait si l'image I était gardée fixe et si tous les noyaux de convolution du modèle ϕ subissaient une rotation d'un angle θ. Par définition, ϕ est équivariant par rotation si $\mathbf{\Phi}$ est constant : $\mathbf{\Phi}(\theta) = \mathbf{y}\ \forall\ \theta \in \mathbb{R}$. Notre étude pourrait s'arrêter à la vérification de cette invariance, mais le bruit et les artéfacts d'interpolation liés aux rotations rendent floue la notion de constance dans ce contexte. Il nous faut la quantifier.

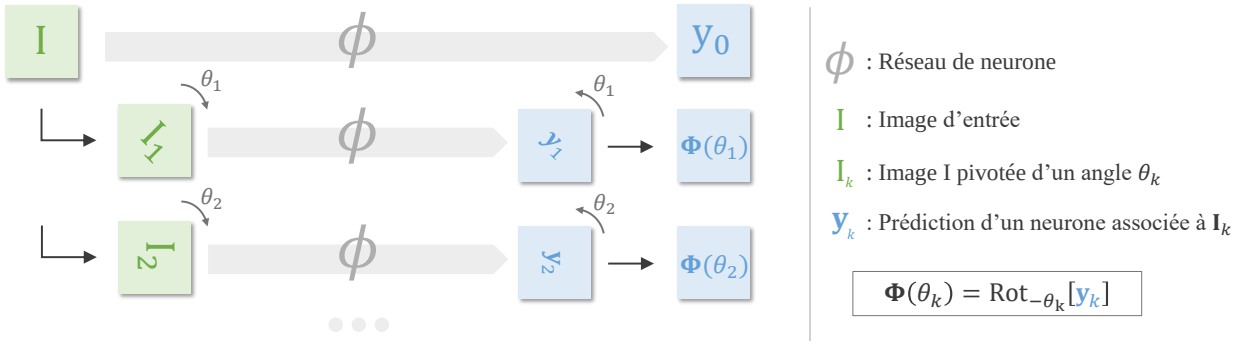


FIGURE 5.2 Schéma de la méthode de calcul de $\mathbf{\Phi}(\theta)$.

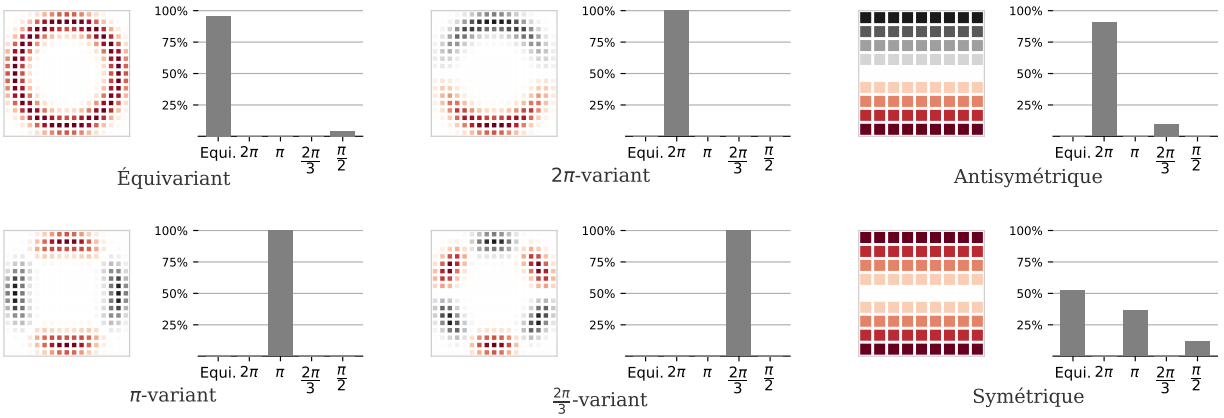


FIGURE 5.3 Exemples de décomposition polaire de la densité spectrale de puissance (DSP) pour des harmoniques polaires de rang 0, 1, 2 et 3 ; et pour deux noyaux de convolution symétrique et anti-symétrique.

Soit $\mathbf{R}(\Delta\theta)$ la fonction d'autocorrélation de $\mathbf{\Phi}(\theta)$ définie pour chaque pixel par :

$$\mathbf{R}(\Delta\theta) = \sum_{\theta} \mathbf{\Phi}(\theta) \cdot \mathbf{\Phi}(\theta - \Delta\theta)$$

On s'intéresse à sa transformée de Fourier $\mathbf{S}(\omega) = \mathcal{F}\Big[\mathbf{R}(\Delta\theta)\Big]$. L'autocorrélation permet ici d'évaluer les variations de $\mathbf{\Phi}$ induites par une rotation $\Delta\theta$, et ce, quelle que soit l'orientation de départ, supprimant ainsi la primauté de $\theta = 0$ comme orientation de référence. La transformée de Fourier décompose ensuite ces variations par harmoniques angulaires. Ainsi, $\mathbf{S}(\omega)$ distingue par leur tropisme les différentes composantes de la fonction réalisée par ϕ : $\mathbf{S}(0)$ en identifie la composante équivariante, $\mathbf{S}(\frac{1}{2\pi})$ la composante anti-symétrique, $\mathbf{S}(\frac{2}{2\pi})$ celle symétrique, etc. (voir Figure 5.3). Le théorème de Wiener-Khinchin indique que $\mathbf{S}(\omega)$ est équivalente à la densité spectrale de puissance et en donne une méthode de calcul plus commode qui se passe de l'autocorrélation : $\mathbf{S}(\omega) = \Big|\mathcal{F}\Big[\mathbf{\Phi}(\theta)\Big]\Big|^2$.

Analyse de l'équivariance par rotation de la classification A/V

En calculant $\mathbf{S}(\omega)$ sur la carte de probabilité prédite par un CNN, on obtient alors une série de spectres (un pour chaque pixel) qui, agrégés et normalisés, offrent un bon outil de compréhension du tropisme du modèle. La première composante de ce spectre indique la proportion équivariante du signal prédit, ou dit autrement, la proportion de fonctions isotropiques parmi l'ensemble des fonctions du modèle. Ce spectre renseigne aussi sur la proportion de composantes 2π-variantes ou π-variantes, associées aux fonctions antisymétriques et symétriques.

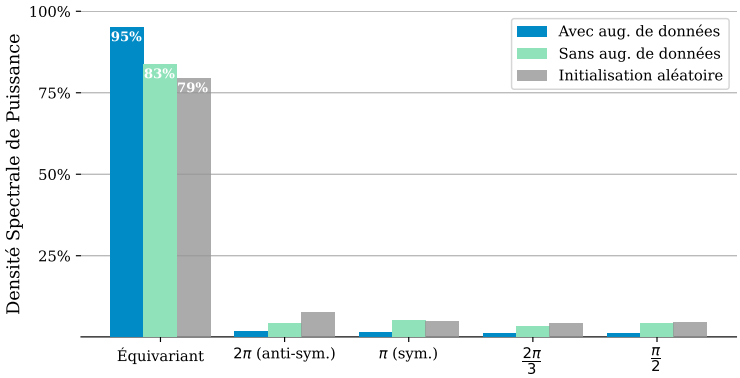


FIGURE 5.4 Comparaison du tropisme de la prédiction de CNN entraîné avec et sans augmentation de données géométriques.

Dans une première expérience de cette étude préliminaire, nous comparerons l'équivariance de modèles entraînés avec et sans augmentation de données géométriques. La Figure 5.4 présente pour ces deux cas, les DSP de modèles [1] évalués sur 25 images du jeu de test de MAPLES-DR. La DSP d'un modèle non entraîné et initialisé aléatoirement a aussi été ajoutée à la Figure à titre de référence. D'après ces résultats, les variations de prédiction liées à un changement d'orientation de l'image que l'on avait rapportées en introduction de ce chapitre (cf. Figure A) concerneraient moins de 5% de la puissance totale du signal prédit. En tout cas lorsque le modèle est entraîné avec des rotations aléatoires de ses échantillons d'entrée, car lorsque cette augmentation de données est désactivée, cette proportion monte à 17%, se rapprochant de sa distribution avant entraînement. En première approche, cette méthode simple semble donc efficace pour induire une équivariance, même imparfaite, à la rotation.

Analyse de l'équivariance par rotation de chaque couche convolutive

Si les CNN sont bien globalement isotropiques, qu'en est-il de leurs couches individuelles ? La Figure 5.5 présente, pour 8 CNN entraînés avec augmentation de données et 8 CNN entraînés sans, la DSP des composantes équivariantes, anti-symétriques et symétriques de chacune de leurs couches convolutives. L'architecture et le protocole d'entraînement de ces modèles seront présentés dans la section de validation ; indiquons seulement qu'il s'agit d'un U-Net comptant 5 étages d'encodeur (Conv 1 à 5) et 4 étages de décodeur (Conv 6 à 9).

Plusieurs observations sont à tirer de cette Figure. Elle confirme d'abord l'équivariance incomplète mais significativement améliorée par l'augmentation de données géométriques des modèles entraînés à la classification des vaisseaux rétiniens. Elle signale surtout que la prédiction finale n'est pas la seule à être globalement équivariante à la rotation. En fait, les seules couches anisotropiques du réseau sont les plus profondes, au centre du U-Net ; et même pour

1. Le protocole d'entraînement de ces modèles est le même que celui qui sera décrit dans la section de validation de ce chapitre. Les spectres de la Figure sont une moyenne obtenue par l'agrégation des prédictions de 8 modèles entraînés avec augmentation de données géométriques et de 8 modèles entraînés sans.

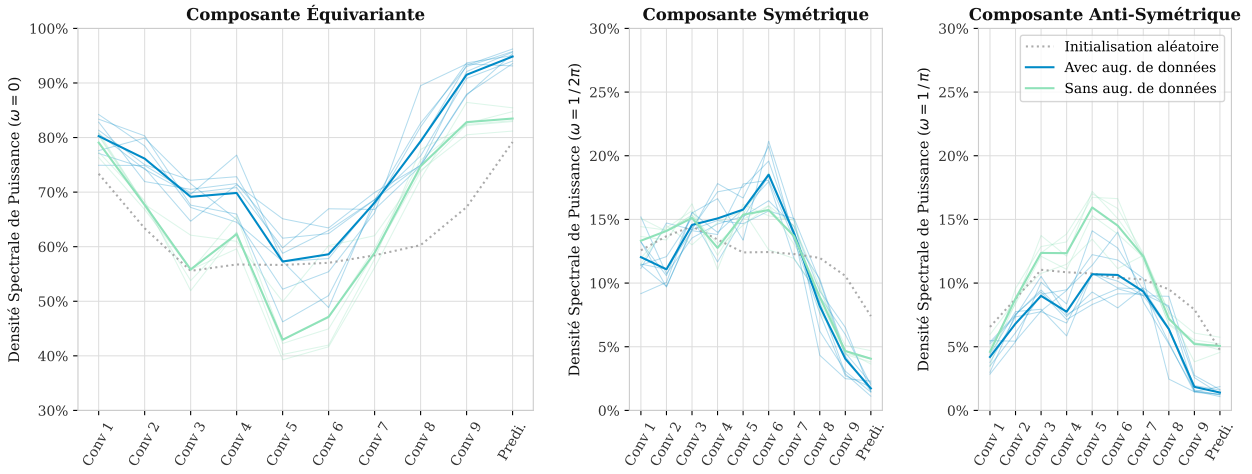


FIGURE 5.5 Effet de l'augmentation de données sur les composantes polaires de la Densité Spectrale de Puissance pour chaque couche convolutive d'un CNN classifieur d'A/V.

Les lignes opaques sont la moyenne des huits meilleurs modèles entraînés avec (*bleu*) et sans (*vert*) augmentation de données. Une ligne propre à chaque modèle est affichée semi-transparente. La ligne pointillée est une référence correspondant à un modèle initialisé aléatoirement et non entraîné.

celles-ci, la composante équivariante reste la plus importante de toutes. En comparant les courbes obtenues avec et sans augmentation de données, il apparaît clairement que si cette dernière améliore bel et bien l'équivariance globale du modèle, son action ne se limite pas à la couche finale : c'est l'ensemble des neurones qui deviennent plus isotropiques et voient s'atténuer leurs composantes anti-symétriques. En analysant individuellement chaque neurone de chaque couche, le constat est encore plus frappant : en moyenne, si l'on exclut les deux couches centrales (Conv 5 et Conv 6), plus de 90% des neurones des CNN entraînés avec l'augmentation de données sont principalement isotropiques !

Ce résultat est doublement paradoxal : si autant de paramètres sont naturellement isotropiques, est-il vraiment nécessaire de développer un formalisme équivariant pour dédupliquer les quelques paramètres restants ? En un sens, s'ils sont « insensibles » à la rotation, ces paramètres sont déjà mutualisés entre les différentes orientations vasculaires. Certes. Mais comment expliquer alors que les approches traditionnelles pour la classification des vaisseaux rétiniens trouvent généralement plus pertinentes des caractéristiques anisotropiques [78, 82] ?

La théorie des réseaux neuronaux démontre que, si le nombre de paramètres et d'échantillons d'entraînement est suffisant, un réseau peut théoriquement approximer n'importe quelle fonction. Néanmoins, en pratique, l'architecture et la procédure d'apprentissage favorisent certaines approximations, au détriment d'autres. Aussi, je postule que la régularisation par augmentation de données détourne le modèle de l'apprentissage des noyaux de convolution anisotropiques. Ainsi, le défi à relever par les *steered* CNN ne se limite pas à l'implémentation d'un modèle complètement équivariant à la rotation, il doit aussi permettre de la réaliser sans imposer l'apprentissage de noyaux de convolution isotropiques.

5.3 Méthode

5.3.1 Objectif et vue schématique des *steered* CNN

Dans un CNN standard, les couches convolutives suivent la forme mathématique suivante :

$$\mathrm{y}_n(i,j) = \sum_m \mathbf{x}_m \underset{(i,j)}{\circledast} \mathbf{W}_{n,m} \tag{5.4}$$

où $\mathrm{y}_n(i,j)$ est le pixel (i,j) de la carte $\mathbf{y}_n$ produite par le $n^{\text{ième}}$ neurone de la couche ; $\mathbf{x}_m$ est la $m^{\text{ième}}$ carte en entrée de la couche ; $\mathbf{W}_{n,m}$ est le noyau de convolution reliant l'entrée $\mathbf{x}_m$ à la sortie $\mathbf{y}_m$; et où $\underset{(i,j)}{\circledast}$ est l'opérateur de corrélation croisée [2] calculé au pixel (i,j).

L'objectif des *steered* CNN est de pouvoir choisir pour chaque pixel (i,j), l'angle $\alpha(i,j)$ sous lequel le noyau doit être orienté avant sa convolution avec l'image. Autrement dit, on souhaite transformer l'équation 5.4 en :

$$\mathrm{y}_n(i,j) = \sum_m \mathbf{x}_m \underset{(i,j)}{\circledast} \mathrm{Rot}_{\alpha(i,j)}\left[\mathbf{W}_{n,m}\right] \tag{5.5}$$

Du point de vue du traitement des signaux, cette formulation découple deux tâches qui sont normalement intriquées pour un CNN standard : d'une part la reconnaissance et classification des vaisseaux et d'autre part la détection de leur orientation. Dans les *steered* CNN, cette seconde tâche est « externalisée » et devient un signal d'entrée du modèle sous la forme d'un champ d'orientation [3] $\vec{\boldsymbol{\alpha}}$ indiquant, pour chaque pixel de l'image, la direction du vaisseau le plus proche (voir section 5.3.2 pour son calcul). Les *steered* CNN ne sont donc pas équivariants par rotation au sens strict du terme : pour réaliser cette propriété, il repose sur l'hypothèse que lors d'une rotation de l'image, les directions des vecteurs de $\vec{\boldsymbol{\alpha}}$ subiront la même rotation, et ils la répercuteront sur les noyaux de convolution du modèle. À tout le moins, l'équation 5.5 assure la mutualisation des paramètres entre les différentes orientations vasculaires (sous réserve que $\vec{\boldsymbol{\alpha}}$ traduise bien la direction locale des vaisseaux).

Cette reformulation des couches convolutives qui rend leur orientation pilotable localement, est la seule modification qui distingue les *steered* CNN des CNN standards. Ainsi, n'importe quel modèle de classification vasculaire reposant sur des neurones convolutifs peut bénéficier de ce formalisme en remplaçant simplement ses couches convolutives par leur version « steered » et en lui ajoutant une entrée secondaire : le champ d'orientation $\vec{\boldsymbol{\alpha}}$.

---

2. Les implémentations des couches convolutives ont généralement recours à la corrélation croisée plutôt qu'à la convolution. Ce détail d'implémentation n'ayant que très peu d'importance, on continuera d'utiliser le terme *convolution* dans le reste du chapitre.

3. Pour le champ d'orientation on utilisera la nomenclature suivante : $\vec{\boldsymbol{\alpha}}$ est le champ de vecteurs (dotés d'une norme et d'une direction) et $\boldsymbol{\alpha}$ est une matrice d'angle : $\boldsymbol{\alpha}(i,j) = \arg\left(\vec{\boldsymbol{\alpha}}(i,j)\right)$

5.3.2 Calcul du champ de direction $\vec{\alpha}$

Les *steered* CNN ajustent l'orientation de leur noyau de convolution selon un champ de direction $\vec{\boldsymbol{\alpha}}$ calculé en amont. Dans le contexte de la classification des vaisseaux en artères ou veines, ce champ doit suivre la direction locale des vaisseaux.

Les méthodes de segmentation des vaisseaux du fond d'œil ont aujourd'hui atteint un niveau de performance suffisant (cf. section 2.2.5) pour que la procédure de calcul du champ $\vec{\boldsymbol{\alpha}}$ puisse raisonnablement s'appuyer sur elles. Cette procédure débute par une squelettisation de la segmentation vasculaire. On attribue ensuite à chaque pixel v du squelette un champ vectoriel radial $\vec{\gamma}_v$ orienté dans la direction opposée à v et dont la norme est pondérée par une gaussienne centrée sur v. Cette pondération s'assure que l'intensité du champ décroisse à mesure que la distance à v augmente. Son écart-type a été fixé empiriquement à 75 pixels pour une image de fond d'œil de résolution 1500×1500 pixels. Le champ d'orientation $\vec{\boldsymbol{\alpha}}$ est alors défini comme la somme des influences de l'ensemble des pixels du squelette vasculaire : $\vec{\boldsymbol{\alpha}} = \sum_v \vec{\gamma}_v$. Orienter ce champ dans une direction normale aux vaisseaux plutôt que tangentielle évite d'avoir à leur attribuer un sens. La Figure 5.6 présente un exemple de champ d'influence $\vec{\gamma}_v$ et du champ d'orientation total $\vec{\boldsymbol{\alpha}}$.

Afin de préserver la cohérence spatiale au sein d'un *steered* CNN, la résolution spatiale du champ $\vec{\boldsymbol{\alpha}}$ est adaptée à celle des cartes d'entrées de chaque couche convolutive. En particulier, lorsque ces cartes sont sous-échantillonnées (suite à une couche de pooling ou un pas de convolution), $\vec{\boldsymbol{\alpha}}$ subit le même sous-échantillonnage précédé d'un filtre moyenneur.

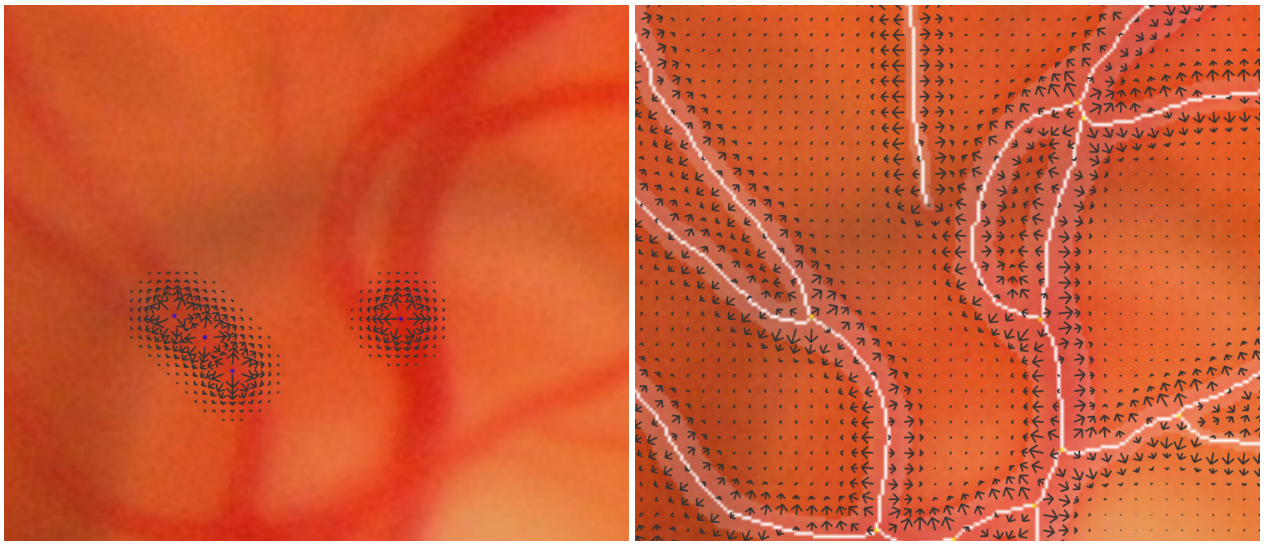


FIGURE 5.6 Champ d'orientation $\vec{\boldsymbol{\alpha}}$ calculé à partir de la squelettisation de la segmentation vasculaire. *À gauche* : Champ d'influence généré par 4 pixels individuels du squelette (dessinés en bleu). *À droite* : Champ complet, superposé au squelette et à la segmentation vasculaire. (Pour des raisons de lisibilité, les champs affichés sont sous-échantillonnés.)

5.3.3 Reparamétrisation des noyaux de convolution

La mise en œuvre de la formule 5.5 n'est pas triviale : la valeur de $\alpha(i,j)$ et donc celle de $\mathrm{Rot}_{\alpha(i,j)}\left[\boldsymbol{W}_{n,m}\right]$ varie à chaque pixel (i,j). Il faudrait donc, en théorie, recalculer une rotation des filtres de convolution pour chacun des pixels de l'image. Cette solution est évidemment irréaliste pour des réseaux qui convoluent des milliers de filtres avec des millions de pixels. La présente section montre comment résoudre ce problème par une reparamétrisation des filtres de convolution $\mathbf{W}_{n,m}$ sous une forme « steerable ».

Pour commencer, il nous faut choisir une famille de steerable filters et le jeu de filtres élémentaires associé. Ce sera le jeu de filtres utilisé par les Steerable CNN, formés par une bande gaussienne dans la dimension radiale et une série harmonique dans leur composante polaire. Ces filtres élémentaires, définis dans le domaine complexe, sont générés par :

$$\Psi_{k,r}(\rho,\theta) = G_r(\rho) \ \cdot \ e^{\mathrm{i}k\theta} \tag{5.6}$$

dont la composante radiale $G_r(\rho) = e^{2(\rho-r)^2}$ décrit un anneau de rayon r pondéré par une gaussienne, et la composante polaire est l'harmonique angulaire de rang k (voir Figure 5.7).

Considérons maintenant un noyau de convolution standard $\mathbf{W}(p,q) \in \mathbb{R}$ défini dans le système de coordonnées cartésien $(p,q) \in \mathbb{Z}^2$. À la manière d'une décomposition en ondelette polaire, on peut approximer $\mathbf{W}(p,q)$ par une combinaison linéaire de filtres élémentaires $\boldsymbol{\Psi}_{k,r}(p,q)$ produits par le rééchantillonnage en coordonnées cartésiennes des filtres $\Psi_{k,r}(\rho,\theta)$:

$$\mathbf{W}(p,q) \approx \mathrm{Re}\Big(\sum_{k=0}^{K}\sum_{r=0}^{R} \omega_{k,r} \ \boldsymbol{\Psi}_{k,r}(p,q)\Big) \tag{5.7}$$

L'analogie de la décomposition en ondelette pour interpréter cette équation est d'autant plus adaptée que les filtres $\boldsymbol{\Psi}_{k,r}$ constituent, dans leur composante angulaire, les bases d'une décomposition en série de Fourier polaire. Les filtres générés par $k = 0$ codent les composantes purement radiales de $\mathbf{W}$ et ceux générés avec $k \geq 1$ représentent chacun une fréquence angulaire propre. Naturellement, la qualité de l'approximation de $\mathbf{W}$ par la décomposition 5.7 dépend des paramètres générant la famille de filtres élémentaires $\boldsymbol{\Psi}_{k,r}$: le rang harmonique maximum K limite la résolution polaire du filtre reconstruit, quant au rayon R il en définit la taille maximale. Le réglage de ces paramètres sera discuté dans la section 5.3.6. Notons que, pour approximer le filtre carré $\mathbf{W}$ de taille $P \times P$, les filtres de convolution élémentaires $\boldsymbol{\Psi}_{k,r}$ doivent être de taille $\sqrt{2}P \times \sqrt{2}P$, afin de modéliser la rotation des pixels aux extrémités des diagonales de $\mathbf{W}$. Ainsi, la décomposition d'un filtre de convolution 5×5 nécessite des filtres élémentaires de taille 7×7 (voir Figure 5.7).

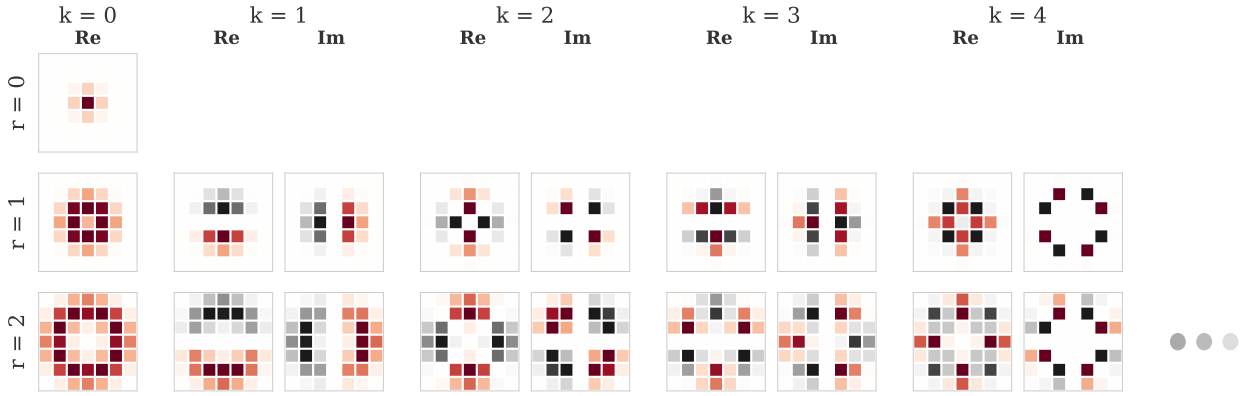


FIGURE 5.7 Le jeu de filtres élémentaires $\Psi_{k,r}$ échantillonné en coordonnées cartésiennes.
Les valeurs positives sont colorées en rouge et celles négatives en noir.

Les coefficients $\omega_{k,r}$ sont (comme les filtres élémentaires $\mathbf{\Psi}_{k,r}$) définis dans le domaine complexe et en ce sens chaque coefficient est double [4] : $\omega_{k,r} = \omega_{k,r}^{\mathrm{Re}} + \mathrm{i}\,\omega_{k,r}^{\mathrm{Im}}$. Pour interpréter la valeur de ces coefficients, la formulation polaire est plus adaptée : $|\omega_{k,r}|$ pondère l'intensité de $\mathbf{\Psi}_{k,r}$, et $\arg(\omega_{k,r})$ définit le déphasage de $\mathbf{\Psi}_{k,r}$ relativement aux autres filtres élémentaires. Vu de $\mathbf{W}$ – qui ne considère que la partie réelle de la combinaison linéaire des filtres élémentaires – ce déphasage se traduit par une rotation. En modulant ainsi l'intensité et l'orientation de chaque filtre $\mathbf{\Psi}_{k,r}$, les coefficients linéaires $\omega_{k,r}$ sont l'équivalent pour les couches convolutives *steered* de ce que sont les poids pour les couches convolutives classiques : ils définissent quelle forme géométrique est détectée par le noyau de convolution. Sous cette perspective, l'équation 5.7 consiste donc en une reparamétrisation des couches convolutives, dont les implications seront explorées dans la prochaine section.

Mais revenons auparavant à la rotation du filtre de convolution $\mathbf{W}_{n,m}$. On peut montrer que, suite à sa décomposition suivant l'équation 5.7, et par un jeu de phases et de trigonométrie, le filtre $\mathbf{W}_{n,m}$ est « steerable » selon la formule :

$$\mathrm{Rot}_\alpha\Big[\mathbf{W}_{n,m}\Big] = \sum_{r=0}^{R}\sum_{k=0}^{K}\Big[\cos(k\alpha)\mathbf{W}_{n,m,k,r}^{\mathrm{Re}} + \sin(k\alpha)\mathbf{W}_{n,m,k,r}^{\mathrm{Im}}\Big] \tag{5.8}$$

où $\mathbf{W}_{n,m,k,r}^{\mathrm{Re}}$ et $\mathbf{W}_{n,m,k,r}^{\mathrm{Im}}$ sont la partie réelle et imaginaire de $\mathbf{W}_{n,m,k,r} = \omega_{n,m,k,r} \cdot \mathbf{\Psi}_{k,r}$ qui s'écrivent sous leur forme développée :

$$\begin{aligned}\mathbf{W}_{n,m,k,r}^{\mathrm{Re}} &= \omega_{n,m,k,r}^{\mathrm{Re}}\mathbf{\Psi}_{k,r}^{\mathrm{Re}} + \omega_{n,m,k,r}^{\mathrm{Im}}\mathbf{\Psi}_{k,r}^{\mathrm{Im}}\\ \mathbf{W}_{n,m,k,r}^{\mathrm{Im}} &= \omega_{n,m,k,r}^{\mathrm{Re}}\mathbf{\Psi}_{k,r}^{\mathrm{Im}} - \omega_{n,m,k,r}^{\mathrm{Im}}\mathbf{\Psi}_{k,r}^{\mathrm{Re}}\end{aligned}$$

4. À l'exception des coefficients $\omega_{0,r}$ qui sont réels purs car $\mathbf{\Psi}_{k,r}$ n'a pas de partie imaginaire pour $k = 0$.

La combinaison de 5.5 et 5.8 donne alors la définition des SCN :

$$\begin{aligned} y_n(i,j) = \sum_{r=0}^{R} \sum_{k=0}^{K} \Bigg[& \cos\Big(k\,\alpha(i,j)\Big) \sum_m \Big(\mathbf{x}_m \underset{(i,j)}{\circledast} \mathbf{W}^{\mathrm{Re}}_{n,m,k,r}\Big) \\ & + \sin\Big(k\,\alpha(i,j)\Big) \sum_m \Big(\mathbf{x}_m \underset{(i,j)}{\circledast} \mathbf{W}^{\mathrm{Im}}_{n,m,k,r}\Big)\Bigg] \end{aligned} \tag{5.9}$$

Autrement dit, en convoluant dans un premier temps $\mathbf{x}_m$ avec les noyaux $\mathbf{W}^{\mathrm{Re}}_{n,m,k,r}$ et $\mathbf{W}^{\mathrm{Im}}_{n,m,k,r}$ (pour toutes les valeurs de k et r), puis en multipliant élément par élément les résultats de ces convolutions avec les cartes $\cos(k\,\boldsymbol{\alpha})$ et $\sin(k\,\boldsymbol{\alpha})$, il est bien possible de sélectionner, a posteriori et pour chaque pixel (i,j), l'orientation dans laquelle $\mathbf{W}_{n,m,k,r}$ est appliqué.

5.3.4 Différences théoriques entre les CNN standards et les *steered* CNN

À première vue, la capacité d'adapter l'orientation de ces filtres peut sembler être l'unique distinction entre un *steered* CNN et un CNN standard. Mais en y regardant de plus près, la reparamétrisation des noyaux de convolution implique une utilisation bien différente de leurs paramètres. Dans un CNN standard, chaque poids correspond à un unique pixel d'un noyau de convolution, sa fonction géométrique ne prend donc sens que par sa relation avec les poids adjacents et sa valeur ne peut être interprétée que relativement à ces derniers. À l'inverse, les paramètres $\omega_{n,m,k,r}$ des *steered* CNN pondèrent chacun un filtre élémentaire $\boldsymbol{\Psi}_{k,r}$ complet. Ils ont donc leur propre signification géométrique caractérisée par une dispersion spatiale r et une fréquence angulaire k.

Cette différence conceptuelle du rôle déféré à chaque paramètre est encore plus évidente lorsqu'on s'intéresse à l'entraînement d'un *steered* CNN et aux formules de rétropropagation du gradient qui le régissent. On peut en effet montrer que le gradient de la fonction de coût $\mathcal{L}$ relativement aux paramètres $\omega_{n,m,k,r}$ s'écrit :

$$\mathrm{d}\omega_{n,m,k,r} = \frac{\partial \mathcal{L}}{\partial \omega_{n,m,k,r}} = \Big(\mathbf{x}_m \underset{(i,j)}{\circledast} \mathrm{Rot}_{\alpha(i,j)}\Big[\boldsymbol{\Psi}_{k,r}\Big]\Big) \circledast \mathrm{d}\mathbf{y}_n \tag{5.10}$$

où $\mathrm{d}\mathbf{y}_n = \frac{\partial \mathcal{L}}{\partial \mathbf{y}_n}$ est le signal de gradient associé à la prédiction $\mathbf{y}_n$ de la couche convolutive. Pour rappel, cette même formule est réduite dans un CNN standard à $\mathrm{d}\mathbf{W}_{n,m} = \mathbf{x}_m \circledast \mathrm{d}\mathbf{y}_n$. En d'autres termes, les mises à jour accumulées dans les paramètres pendant l'entraînement sont obtenues dans les deux cas par la corrélation du signal de gradient $\mathrm{d}\mathbf{y}_n$ avec le signal d'entrée $\mathbf{x}_m$. Mais dans les *steered* CNN, $\mathbf{x}_m$ a préalablement été filtré avec le filtre élémentaire $\boldsymbol{\Psi}_{k,r}$ associé au paramètre $\omega_{n,m,k,r}$. Au contraire, dans un CNN standard, les mises à jour des poids d'un même noyau de convolution ne diffèrent que par une translation par quelques pixels de $\mathbf{x}_m$ avant sa corrélation avec $\mathrm{d}\mathbf{y}_n$. Les paramètres des *steered* CNN sont donc bien plus indépendants les uns des autres pendant l'apprentissage.

Spécialiser chaque paramètre à une fonction géométrique permet aussi d'élaguer les fonctions – et donc les paramètres – qui contribuent le moins à la qualité des prédictions du modèle. Cette idée sera explorée plus en détail dans la section 5.3.6. On y montrera que pour approximer un filtre de taille 5×5 pixels, la plage d'harmonique polaire peut être limitée à $K = 2$ sans perte de performance. La paramétrisation *steered* compte alors 3 paramètres réels purs et 4 paramètres complexes, soit $3 + 4 \times 2 = 11$ degrés de liberté, contre 25 pour une paramétrisation convolutive standard.

Ainsi, à nombre de neurones égal, les paramètres des *steered* CNN sont moins nombreux, plus spécifiques et plus interprétables que leur alter ego standard. Tous ces avantages de modélisation ont cependant un coût computationnel : le nombre de convolutions nécessaires au calcul de $\mathbf{y}_n$ lors de l'inférence ou de $\mathrm{d}\mathbf{x}_n$ lors de la propagation arrière du gradient augmente d'un facteur de $2KR$. En effet, la carte $\mathbf{x}_m$ n'est plus simplement convoluée avec les filtres de convolution standard $\mathbf{W}_{n,m}$, mais avec chacune de leurs composantes $\mathbf{W}^{\mathrm{Re}}_{n,m,k,r}$ et $\mathbf{W}^{\mathrm{Im}}_{n,m,k,r}$ (cf. équation 5.9). La prochaine section montrera comment alléger ce coût additionnel en temps de calcul.

5.3.5 Implémentation optimisée des *steered* CNN

Réarrangements des termes de convolutions

La décomposition de chaque noyau de convolution $\mathbf{W}_{n,m}$ en une combinaison linéaire de filtres élémentaires $\mathbf{\Psi}_{k,r}$ démultiplie le nombre de convolutions dans les *steered* CNN. Les différentes composantes radiales k de ces filtres doivent en effet impérativement être convoluées séparément pour permettre la sélection de leur orientation a posteriori par leur multiplication respective avec $\cos(k\boldsymbol{\alpha})$ et $\sin(k\boldsymbol{\alpha})$. Néanmoins, cette contrainte ne s'applique pas aux composantes radiales r de $\mathbf{W}_{n,m,k,r}$. Une première approche consiste donc à agréger ces composantes en un filtre composite $\mathbf{W}_{nmk} = \sum_{r=0}^{R} \omega_{n,m,k,r} \mathbf{\Psi}_{k,r}$ et ainsi diviser le nombre de convolutions par R. On appellera cette approche *composite* :

$$\begin{aligned} y_n(i,j) = \sum_m \left(\mathbf{x}_m \underset{(i,j)}{\circledast} \mathbf{W}^{\mathrm{Re}}_{nm0} \right) \; + \; \sum_{k=1}^{K} \Bigg[& \cos\Big(k\,\alpha(i,j)\Big) \sum_m \left(\mathbf{x}_m \underset{(i,j)}{\circledast} \mathbf{W}^{\mathrm{Re}}_{n,m,k} \right) \\ & + \; \sin\Big(k\,\alpha(i,j)\Big) \sum_m \left(\mathbf{x}_m \underset{(i,j)}{\circledast} \mathbf{W}^{\mathrm{Im}}_{n,m,k} \right) \Bigg] \end{aligned} \tag{5.11}$$

Par ailleurs, notons dans cette formule que les harmoniques $k = 0$ sont traitées séparément des autres pour tenir compte de leur particularité : elles sont les seules à être réelles pures.

Une autre approche d'optimisation consiste à convoluer en amont les cartes $\mathbf{x}_m$ avec chaque filtre élémentaire $\mathbf{\Psi}_{k,r}$, afin d'obtenir les cartes intermédiaires $\mathbf{z}_{m,k,r} = \mathbf{x}_m \circledast \mathbf{\Psi}_{k,r}$. On recompose alors $\mathbf{y}_n$ par une multiplication matricielle de ces cartes $\mathbf{z}_{m,k,r}$ avec les poids $\omega_{n,m,k,r}$. On désignera cette approche comme *pré-convoluée*. Elle se calcule par :

$$\begin{aligned}\mathbf{y}_n = \sum_m \sum_{r=0}^{R} \left(\omega^{\mathrm{Re}}_{n,m,0,r} \cdot \mathbf{z}^{\mathrm{Re}}_{m,0,r} \right) \; + \; \sum_{k=1}^{K} \Bigg[\; & \cos\left(k\,\boldsymbol{\alpha}\right) \sum_m \sum_{r=1}^{R} \mathrm{Re}\Big(\omega_{n,m,k,r} \cdot \mathbf{z}_{m,k,r} \Big) \\ & + \; \sin\left(k\,\boldsymbol{\alpha}\right) \sum_m \sum_{r=1}^{R} \mathrm{Im}\Big(\omega_{n,m,k,r} \cdot \mathbf{z}_{m,k,r} \Big) \Bigg]\end{aligned} \tag{5.12}$$

Dans ce cas, le nombre de convolutions par couche est donc toujours égal au nombre de filtres élémentaires $\mathbf{\Psi}_{k,r}$: soit $2KR + 1$ convolutions, et ce, quel que soit le nombre de caractéristiques d'entrée ou de sortie de la couche. Pour accélérer encore le calcul, $\mathbf{z}_{m,k,r}$, la taille de chaque filtre élémentaire $\Psi_{k,r}$, peut aussi être limitée à sa valeur minimale imposée par r (3×3 pour $r = 0$; 5×5 pour $r = 1$; etc.) . Néanmoins, cette optimisation des opérations de convolutions implique une augmentation conséquente du nombre de multiplications matricielles.

Ces deux méthodes sont mutuellement exclusives et leur gain respectif en temps de calcul dépend des valeurs de K, R, n, m, de la taille de l'image, etc. Testée en contexte réel d'un U-Net complet où $K = 2$ et $R = 2$, l'inférence d'une image 512×512 pixels est plus rapide avec l'approche préconvoluée (166 ms plutôt que 178 ms), mais la rétropropagation du gradient bénéficie davantage de l'approche composite (263 ms contre 288 ms). Dans les deux cas, ces optimisations réduisent efficacement le temps de calcul ajouté par la formulation *steered*.

Calcul rapide de $\cos(k\boldsymbol{\alpha})$ et $\sin(k\boldsymbol{\alpha})$

Au-delà des convolutions supplémentaires, les calculs trigonométriques de $\cos(k\boldsymbol{\alpha})$ et $\sin(k\boldsymbol{\alpha})$ sur l'ensemble des pixels de l'image sont aussi coûteux, mais peuvent être contournés. En effet, le champ d'orientation est communiqué au réseau sous la forme d'un champ de vecteurs $\vec{\boldsymbol{\alpha}}$. En assumant que ses vecteurs soient *unitaires*, alors les composantes de $\vec{\boldsymbol{\alpha}}$ dans un repère cartésien s'écrivent :

$$\vec{\alpha}(i,j) = \begin{bmatrix} \cos\big(\alpha(i,j)\big) \\ \sin\big(\alpha(i,j)\big) \end{bmatrix}$$

On notera $\vec{\boldsymbol{\alpha}}_1$ ce champ de vecteur contenant les valeurs de $\cos(k\boldsymbol{\alpha})$ et $\sin(k\boldsymbol{\alpha})$ pour $k = 1$. Par identité trigonométrique, on peut montrer que le champ $\vec{\boldsymbol{\alpha}}_k$ est égal à une combinaison

linéaire des composantes cartésiennes de $\vec{\boldsymbol{\alpha}}_{k-1}$ et $\vec{\boldsymbol{\alpha}}_1$.

$$\begin{aligned}
\vec{\boldsymbol{\alpha}}_k &= \begin{bmatrix} \boldsymbol{\alpha}_k^{(x)} \\ \boldsymbol{\alpha}_k^{(y)} \end{bmatrix} = \begin{bmatrix} \cos\left(k\,\boldsymbol{\alpha}\right) \\ \sin\left(k\,\boldsymbol{\alpha}\right) \end{bmatrix} \\
&= \begin{bmatrix} \cos\left((k-1)\,\boldsymbol{\alpha}\right)\cos\left(\boldsymbol{\alpha}\right) - \sin\left((k-1)\,\boldsymbol{\alpha}\right)\sin\left(\boldsymbol{\alpha}\right) \\ \cos\left((k-1)\,\boldsymbol{\alpha}\right)\sin\left(\boldsymbol{\alpha}\right) + \sin\left((k-1)\,\boldsymbol{\alpha}\right)\cos\left(\boldsymbol{\alpha}\right) \end{bmatrix} \\
\vec{\boldsymbol{\alpha}}_k &= \begin{bmatrix} \boldsymbol{\alpha}_{k-1}^{(x)}\,\boldsymbol{\alpha}_1^{(x)} & - & \boldsymbol{\alpha}_{k-1}^{(y)}\,\boldsymbol{\alpha}_1^{(y)} \\ \boldsymbol{\alpha}_{k-1}^{(x)}\,\boldsymbol{\alpha}_1^{(y)} & + & \boldsymbol{\alpha}_{k-1}^{(y)}\,\boldsymbol{\alpha}_1^{(x)} \end{bmatrix}
\end{aligned}$$

On a donc une définition récursive de $\cos(k\boldsymbol{\alpha})$ et $\sin(k\boldsymbol{\alpha})$ nécessitant uniquement des opérations de multiplication-accumulation dans lesquelles excellent les cartes graphiques.

5.3.6 Réglage du jeu des filtres élémentaires

Choix du rang harmonique maximum K

Malgré les optimisations présentées à la section 5.3.5, le temps de calcul nécessaire à un *steered* CNN reste supérieur à celui d'un CNN standard d'un facteur $2K+1$. Il y a donc un compromis sur le choix du rang harmonique maximum K. Une valeur élevée assure une résolution angulaire supérieure, mais au prix d'un temps de calcul allongé. La question est donc de savoir, à budget de temps de calcul constant, à partir de quelle valeur de K le gain de performance lié à la hausse de la résolution angulaire ne compense plus la perte de gain potentiel qu'aurait permise une optimisation plus longue des hyper-paramètres.

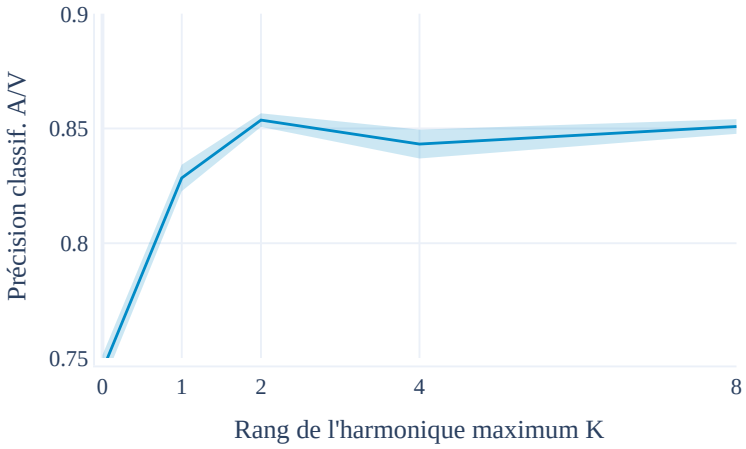


FIGURE 5.8 Précision moyenne d'un *steered* CNN sur son ensemble de test en fonction du rang harmonique maximum K de ces filtres élémentaires $\boldsymbol{\Psi}_{k,r}$.

Pour répondre à cette question, on a entraîné à la classification de vaisseaux rétiniens [5] une série de modèles avec des rangs harmoniques maximum variables. Leurs performances ont ensuite été évaluées sur un ensemble de test, elles sont présentées sur la Figure 5.8. Les résultats de cette série d'entraînement sont univoques : au-delà de $K = 2$, la précision sur l'ensemble de test stagne et régresse même légèrement. Tous les *steered* CNN présentés dans les expériences de validation seront donc limités à des harmoniques de rang 0, 1 et 2.

Par ailleurs, la Figure 5.8 révèle que, comme nous l'avions pressenti, une modélisation purement isotropique n'est pas suffisante pour classifier des artères et des veines rétiniennes. En effet, les modèles n'ayant accès qu'à des filtres élémentaires où $k = 0$ voient leurs précisions moyennes sur l'ensemble de test chuter de 10 points : 0.75 au lieu de 0.85.

Normalisation de $\mathbf{\Psi}_{n,m}$ et initialisation de $\omega_{n,m,k,r}$

L'architecture de *steered* CNN est inédite, elle ne dispose évidemment d'aucun poids préentraîné sur ImageNet, et ses poids doivent être initialisés aléatoirement avant chaque entraînement. Généralement, les noyaux de convolutions $\mathbf{W}_{n,m}$ dans les CNN standard sont initialisés par une distribution gaussienne de moyenne nulle et dont le paramétrage de la variance est crucial pour garantir la stabilité du signal à travers le modèle. En effet, un réglage légèrement trop faible ou trop élevé induirait une petite amplification ou atténuation qui, réverbérée par chacune des couches, pourrait provoquer l'explosion ou l'effondrement du signal. Dans un *steered* CNN, $\mathbf{W}_{n,m}$ est une composition de $\omega_{n,m,k,r}$ et de $\mathbf{\Psi}_{n,m}$. Ce réglage induit donc des contraintes sur l'initialisation aléatoire de $\omega_{n,m,k,r}$, ainsi que sur la normalisation des $\Psi_{n,m}$.

En suivant la démarche de He et al. [217] on peut montrer que les filtres élémentaires $\mathbf{\Psi}_{n,m}$ doivent être normalisés de sorte que $\sum_{p,q} \Psi^2_{n,m}(p,q) = 1$ et que la gaussienne initialisant les paramètres $\omega_{n,m,k,r}$ doit avoir comme variance :

$$Var\Big[\omega_{n,m,k,r}\Big] = \begin{cases} \dfrac{2}{m \cdot \mathrm{N}_{k,r}} \cdot \dfrac{1}{1 + \frac{2}{\pi}\Big(1 - \Big(\sum_{p,q} \mathbf{\Psi}_{k,r}(p,q)\Big)^2\Big)} & \text{si } k = 0 \\ \dfrac{2}{m \cdot \mathrm{N}_{k,r}} \cdot \dfrac{\pi}{\pi + 2} & \text{si } k > 0 \end{cases}$$

où $N_{k,r}$ est le nombre de filtres élémentaires $\mathbf{\Psi}_{n,m}$ et m est le nombre d'entrées de la couche. La démonstration de ces formules est détaillée en Annexe A.

5. Le protocole d'entraînement est celui partagé par toutes les expériences de validation (voir section 5.4.1).

5.4 Évaluation expérimentale des *steered* CNN

L'étude expérimentale d'une nouvelle formulation de neurone convolutif s'est révélée exaltante par l'espoir et la curiosité qu'elle suscite. Dans cette section, nous évaluerons les performances des *steered convolutional neurons (SCN)* comparées à celles des neurones convolutifs standards et nous identifierons leurs propriétés uniques, en particulier leur capacité de généralisation.

5.4.1 Protocole expérimental

Modèle de référence

Récemment, plusieurs stratégies ont été proposées pour améliorer l'entraînement de modèles convolutifs à la classification des vaisseaux rétiniens, mais il n'existe pas encore de consensus sur laquelle est la meilleure. Pour valider les SCN, nous avons décidé d'ignorer les méthodes les plus récentes et de nous concentrer plutôt sur leur point commun : l'architecture U-Net. De cette façon, les résultats des expériences présentés dans cette section ne sont pas propres à une stratégie d'apprentissage spécifique et peuvent théoriquement être reproduits avec toutes.

Hemeling et al. [39] ont étudié en détail une série de variations d'architecture U-Net pour la classification A/V . Nous retenons de cette étude le modèle générique présenté sur la Figure 5.9 : un U-Net travaillant à 5 échelles obtenues en divisant puis multipliant successivement la résolution des prédictions neuronales par un facteur de 2 grâce à un rééchantillonnage bilinéaire. Chaque résolution est analysée par deux couches convolutives 5×5, chacune suivie d'une *Batch Norm* et d'une fonction d'activation ReLU. Le nombre de caractéristiques prédites par ces couches augmente de façon inversement proportionnelle à leur résolution,

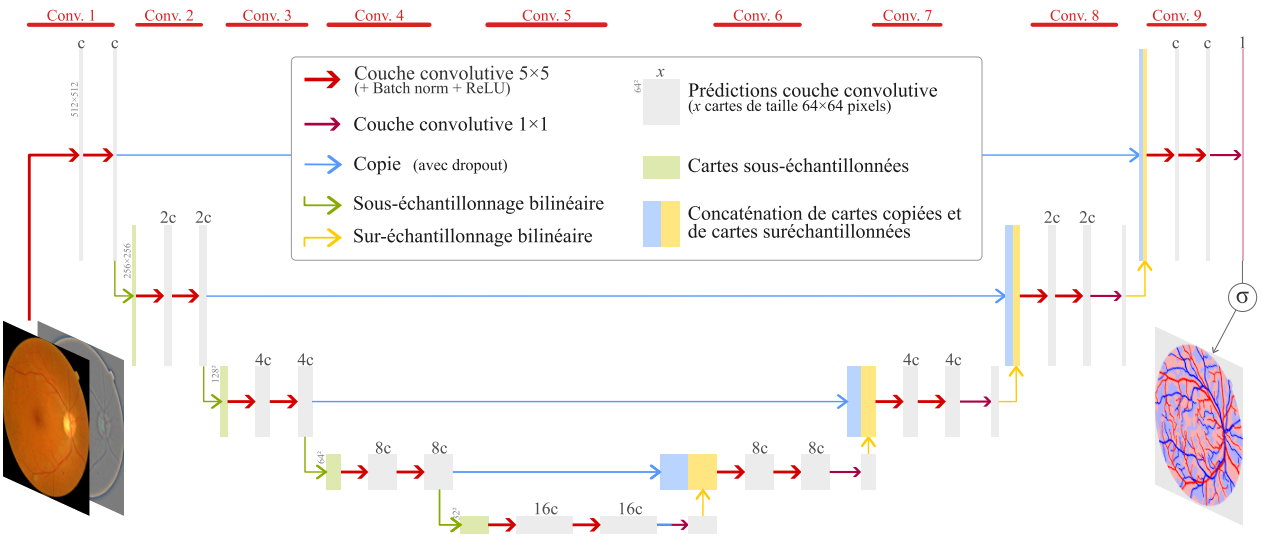


FIGURE 5.9 Architecture de modèle U-Net retenue pour évaluer les *steered* CNN.

passant de c pour les premières couches à $16c$ pour les couches les plus profondes. On étudiera plus loin l'effet de l'hyperparamètre c, mais indiquons tout de suite que les meilleurs modèles ont été obtenus avec $c = 12$ ou $c = 16$.

Notre étude se limite à la classification artères/veines des vaisseaux rétiniens dont on suppose connaître la segmentation. Ainsi, et contrairement à Hemeling et al., la dernière couche de notre modèle ne produit qu'une unique carte dont la sigmoïde indique, pour chaque pixel, la probabilité que le vaisseau le plus proche soit une artère plutôt qu'une veine. Évidemment, la fonction de coût et les métriques de classification du modèle ont été réglées pour ne considérer que les pixels appartenant à la segmentation vasculaire. Par ailleurs, une étude préliminaire sur le pré-traitement de l'image de fond d'œil nous a révélé que le modèle performe mieux lorsqu'on lui présente une image dont l'illumination a été localement corrigée par soustraction de la médiane, plutôt que l'image brute. Néanmoins, les meilleures performances sont observées lorsqu'il a accès aux deux. C'est cette dernière option qui a donc été retenue.

Dans la suite de cette section, deux architectures de modèles seront comparées. La première reprendra à la lettre la description du modèle présentée ci-dessus et servira de base de référence aux expériences. La seconde verra toutes ses couches convolutives 5×5 (en rouge sur la Figure 5.9) remplacées par leur formulation *steered* dans laquelle les noyaux seront orientés selon un champ de vecteur $\vec{\boldsymbol{\alpha}}$ calculé à partir de la segmentation (voir la section 5.3.2).

Procédure d'entraînement

Dans toutes les expériences qui suivront, les deux architectures ont été entraînées et évaluées sur un sous-ensemble de 65 images de MAPLES-DR réannotées avec une classification artère/veine (A/V) [6]. La moitié de ces images (35) a été dédiée à l'entraînement du modèle, 5 à sa validation et 25 à son test. Tous les modèles ont été initialisés aléatoirement selon He et al. [217] (cf. section 5.3.6) puis entraînés par un optimiseur Adam et pour un maximum de 500 époques. La fonction de coût minimisée était une simple entropie croisée binaire. Des fonctions de coût Dice, smooth et focale ont aussi été étudiées, mais aucune n'a amélioré les résultats de l'entropie croisée.

Pour chaque expérience, l'optimisation des hyperparamètres (pas d'apprentissage, augmentation de données, etc.) a été conduite par un algorithme *Tree-structured Parzen Estimator* [218] avec pour objectif de maximiser l'aire sous la courbe ROC (AUC-ROC) sur l'ensemble de validation. Le temps d'entraînement sur une carte graphique NVidia RTX2080 Ti varie de 10 minutes à 5 heures en fonction des configurations du modèle.

6. Au moment où ces expériences ont été conduites, seulement 65 images avait été annotées A/V.

Mesure des performances de classification

Les performances comparées des *steered* CNN et des CNN standard seront estimées par deux métriques de classification. La première est l'*exactitude* (*accuracy* en anglais) qui mesurera leurs performances lorsque leur prédiction est seuillée à $\frac{1}{2}$. Cette métrique est courante pour évaluer la classification des vaisseaux rétiniens : elle rapporte le nombre de pixels correctement classifiés comme artères ou veines, au nombre de pixels total de la segmentation vasculaire. L'aire de cette segmentation étant globalement répartie équitablement entre artères et veines (i.e. les classes sont équilibrées), la valeur de l'exactitude est très proche d'autres métriques courantes de classification comme le score F1 ou l'indice de Jaccard que l'on peut donc omettre. La seconde métrique est l'AUC-ROC. Elle témoigne de la force discriminative des cartes de probabilités sans imposer un point de fonctionnement.

La capacité de généralisation des modèles a été mesurée en évaluant les modèles sur un jeu de données qui n'a pas servi à leur entraînement, soit le jeu de données HRF, publié par Hemeling et al. [39] qui contient 45 images de résolution comparable à celle de MAPLES-DR.

Les expériences qui vont suivre tentent de mesurer l'impact de certains paramètres de l'entraînement (par ex. le nombre de poids du modèle ou le volume d'entraînement) sur les performances de classification de modèles. Cependant, les performances fluctuent d'un entraînement à l'autre et l'optimisation des hyperparamètres reste le facteur le plus déterminant sur leur valeur. Pour mieux isoler l'effet d'une variable spécifique dans le bruit lié à cette optimisation, chaque expérience a été répétée 20 fois et on a sélectionné les 5 meilleurs essais selon l'AUC sur l'ensemble de validation. Tous les résultats présentés ci-dessous indiquent la moyenne et l'écart-type de ces 5 essais en mesurant leur performance sur les ensembles de test de MAPLES-DR et HRF. Au final, cette section de validation s'appuie sur les résultats collectés parmi 600 entraînements de *steered* CNN et 900 entraînements de CNN standard.

5.4.2 Performance générale de classification

Parmi ce millier de modèles, analysons dans un premier temps les 5 meilleurs *steered* CNN avec les 5 meilleurs CNN standard, toutes expériences confondues. Leurs performances de classification A/V sur MAPLES-DR et HRF sont rapportées dans le tableau 5.1.

Tableau 5.1 Performances (AUC-ROC) des 5 meilleurs modèles *steered* et standard

Ensemble de test :	MAPLES-DR		HRF (généralisation)	
Métrique :	Exactitude	AUC-ROC	Exactitude	AUC-ROC
CNN standard	88.2±0.3%	95.2±0.2%	82.3±1.4%	91.2±1.2%
Steered CNN	**89.3±0.8%**	**95.9±0.4%**	82.4±1.9%	91.7±0.7%

Ces premiers résultats sont rassurants : les *steered* CNN apprennent bien à classifier les vaisseaux de MAPLES-DR, ils obtiennent même de meilleurs résultats que les CNN standard sur l'ensemble de test de ce jeu de données : l'AUC-ROC augmente de 0.7% et l'exactitude de 1%. Sur le jeu de test HRF, les deux modèles perdent en performance et, bien que les *steered* CNN conservent une avance sur les CNN standard, celle-ci n'est pas significative étant donné l'écart-type plus grand dont souffrent les mesures. On reviendra plus en détail sur les capacités de généralisation des *steered* CNN plus loin dans cette section.

Qualitativement, les erreurs des CNN standard sont concentrées aux terminaisons et aux croisements de vaisseaux les plus fins, comme nous l'avions vu en introduction. Les *steered* CNN ne corrigent pas toutes ces erreurs : certaines petites branches émergant d'intersections restent mal classifiées. Cependant, leurs prédictions sont moins sujettes à des changements soudains ou à l'apparition de taches au milieu d'une branche, et elles sont globalement plus cohérentes (voir Figure 5.10). L'équivariance par rotation complète des *steered* CNN n'est probablement pas étrangère à la disparition de ces artefacts, qui apparaissaient et disparaissaient chez les CNN standard, au gré de l'orientation initiale de l'image (voir Figure 5.1).

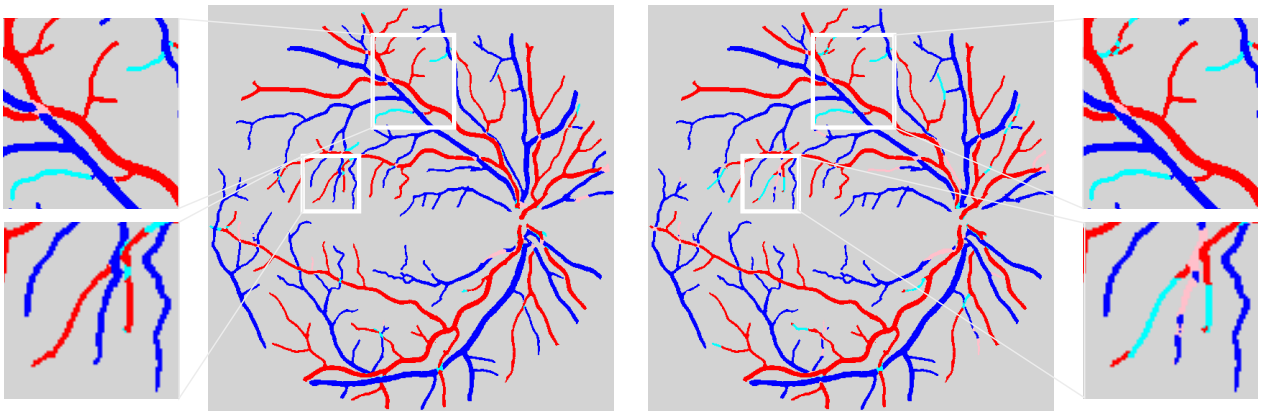


FIGURE 5.10 Classification A/V par un *steered* CNN (gauche) et un CNN standard (droite).
(L'image est issue de l'ensemble de test de MAPLES-DR. Les erreurs de classification d'artères et de veines apparaissent respectivement en cyan et rose..)

5.4.3 Équivariance à la rotation et anisotropisme des SCN

Par définition, chaque couche convolutive d'un *steered* CNN est complètement équivariante à la rotation (sous réserve que le champ $\vec{\boldsymbol{\alpha}}$ pivote avec l'image). Pour comprendre le tropisme des fonctions qu'elle modélise, on ne peut donc pas mener une analyse par DSP comme on l'a fait avec les CNN standards : toute la puissance spectrale serait, par construction, concentrée sur la composante équivariante. On lui a préféré deux approches : la première, indirecte, évalue la dépendance à l'augmentation de données géométriques ; la seconde étudie directement la distribution des poids $\omega_{k,r}$.

Dépendance à l'augmentation de données géométrique

Si l'équivariance par rotation des CNN standard s'appuie largement sur l'augmentation de données géométriques (voir section 5.2), qu'en est-il pour les *steered* CNN ? Le tableau 5.2 compile les performances des CNN *steered* et standard sans augmentation de données géométriques sur MAPLES-DR et sur HRF. Cette augmentation de données est en réalité composée de deux opérations dans notre protocole d'entraînement : une rotation aléatoire et une déformation élastique. Notons que les directions des vecteurs d'orientation $\vec{\boldsymbol{\alpha}}$ ne tiennent compte que de la première : la carte de leur composante verticale et horizontale est déformée par interpolation bilinéaire par la déformation élastique, mais leurs directions ne sont pas corrigées pour compenser les petites rotations locales qu'elle induit.

La désactivation complète de ces augmentations de données détériore les performances des CNN standard et *steered*, mais les *steered* CNN en souffrent moins. Sur le jeu de test de MAPLES-DR, ils perdent 2% de AUC-ROC contre 3% pour les CNN standard ; et sur HRF, ils ne perdent que 3% là où les CNN standard en perdent 5%. Si on ne désactive que les rotations aléatoires mais qu'on conserve les déformations élastiques, la perte de performances sur MAPLES-DR disparaît chez les *steered* CNN mais reste partiellement présente chez les CNN standard à hauteur de 1% (et 2% sur HRF).

Les *steered* CNN ne dépendent donc plus du tout de la rotation des échantillons d'entraînement pour apprendre à classer les artères et veines rétiniennes quelle que soit leur orientation. La déformation élastique reste néanmoins un moyen efficace pour augmenter artificiellement le nombre d'échantillons d'entrée, même si elle est moins nécessaire aux *steered* CNN qu'aux CNN standards.

Ensemble de test :	MAPLES-DR		HRF (généralisation)	
Aug. de données Géométrique :	élastique seule	aucune	élastique seule	aucune
CNN standard	94.2±0.4%	92.2±0.5%	89.7±1.3%	86.5±1.5%
Steered CNN	95.7±0.3%	94.1±0.1%	90.8±1.0%	89.6±0.7%

Tableau 5.2 Effet de la présence ou l'absence de rotation aléatoires des échantillons d'entraînement sur la métrique d'AUC calculée sur les ensembles de test.

Analyse des poids $\omega_{k,r}$ appris par les SCN

La rotation aléatoire des échantillons d'entraînement augmente les performances des CNN standards et leur assure une meilleure équivariance par rotation. Mais elle induit un isotropisme sur l'ensemble des noyaux de convolution, à rebours de la littérature sur la classification A/V par des algorithmes traditionnels qui reposent plutôt sur des filtres anisotropiques. Dans cette alternative, quel choix est fait par les *steered* CNN ? Répondre à cette question est bien plus simple pour les *steered* CNN que pour les CNN standards : les paramètres $\omega_{k,r}$ étant associés à des fonctions géométriques spécifiques, on peut directement comparer la pondération des filtres élémentaires $\mathbf{\Psi}_{k,r}$ isotropiques (pour $k = 0$) avec celles anisotropiques ($k = 1$ ou $k = 2$).

La Figure 5.11 montre la distribution des poids $\omega_{k,r}$ du *steered* CNN ayant obtenu la meilleure AUC-ROC sur l'ensemble de validation (les poids des 5 modèles suivants dans le classement observent les mêmes tendances). Les poids sont regroupés par rang harmonique k et par paire de couches convolutives : où Conv. 1-5 sont les paires de l'encodeur du U-Net, et Conv. 6-9 celles du décodeur. Les trois premières paires donnent autant ou plus de poids aux filtres élémentaires anisotropiques qu'à ceux isotropiques. Puis, cette tendance s'atténue pour Conv. 4 et elle disparaît définitivement à partir de Conv. 5 et pour les couches du décodeur. Notons que cette distinction n'existe pas à l'initialisation du modèle où tous les filtres élémentaires ont la même pondération relative ; elle apparaît à l'entraînement. La solution optimale d'après les *steered* CNN pour classifier les vaisseaux rétiniens est donc conforme à notre intuition : extraire des caractéristiques anisotropiques de l'image de fond d'œil, puis les recombiner isotropiquement pour calculer la prédiction.

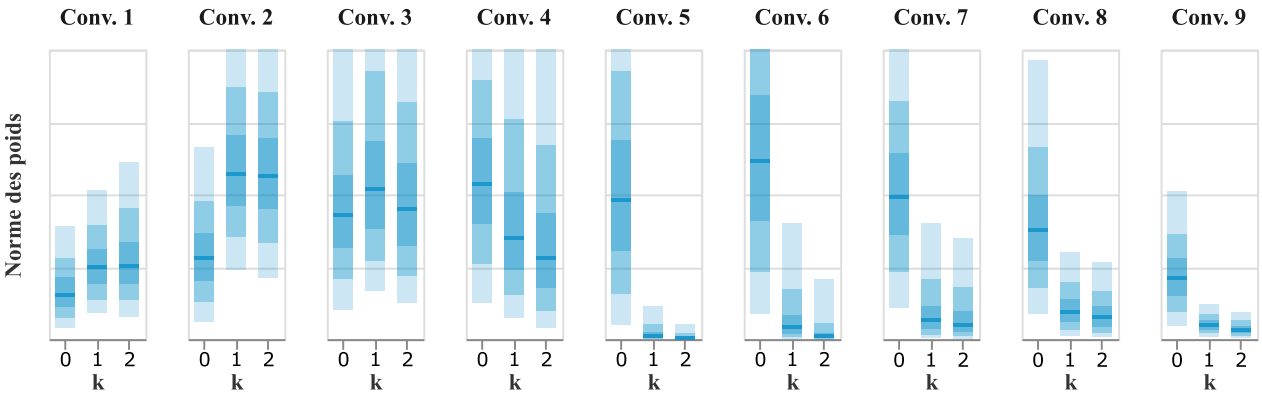


FIGURE 5.11 Distribution des poids $\omega_{k,r}$ d'un *steered* CNN regroupés par harmonique polaire k et par couche convolutive.

La médiane de chaque distribution est représentée par un trait horizontal. Les gradients de couleur indiquent la position des terciles autour de la médiane.

5.4.4 Efficacité de la modélisation

Efficacité paramétrique

Pour évaluer l'efficacité paramétrique des SCN (c'est-à-dire la quantité d'information que chaque paramètre peut encoder), j'ai effectué plusieurs entraînements, en réduisant progressivement le nombre de caractéristiques du modèle : en commençant par une moyenne de 134 caractéristiques par couche, pour finir à une seule (voir Figure 5.12). Dans l'ensemble, l'architecture que nous proposons permet d'obtenir une AUC-ROC plus élevée que celle de référence en utilisant moins de paramètres. En dessous de 20k paramètres en particulier, les performances chutent rapidement pour les CNN standards, alors qu'elles restent stables pour les *steered* CNN sur les deux jeux de données. Même lorsque chaque couche est contrainte à n'utiliser qu'une seule caractéristique – cas complètement irréaliste où leurs mécanismes de combinaison linéaire seraient complètement désactivés – les *steered* CNN atteignent toujours 0,88 AUC sur MAPLES-DR et 0,84 sur HRF.

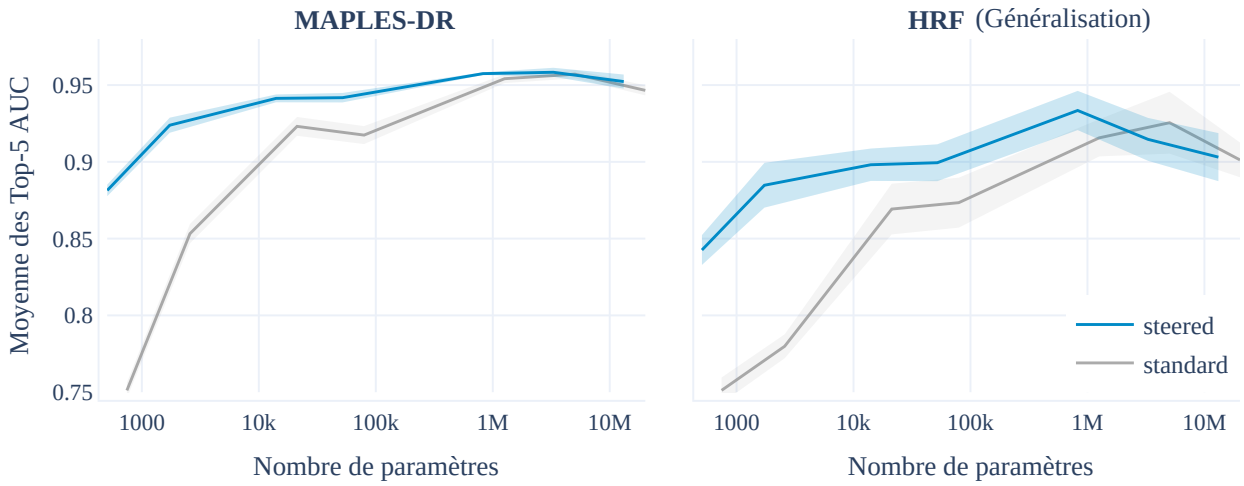


FIGURE 5.12 Performances comparées du *steered* CNN avec un CNN standard en fonction du nombre de paramètres.

Il est intéressant de noter que les courbes de performance sur l'ensemble de données de généralisation HRF illustrent bien l'intuition du rasoir d'Ockham : au-delà de 830k paramètres, les modèles *steered* commencent à se spécialiser sur MAPLES-DR et leur AUC-ROC sur HRF recule. Les modèles standards sont également sujets à ce phénomène, mais uniquement après avoir été entraînés avec 10 fois plus de paramètres, ce qui indique une fois de plus la meilleure efficacité paramétrique des *steered* CNN.

Capacité de généralisation avec peu d'échantillons d'entraînement

Pour évaluer plus en profondeur les capacités de généralisation des *steered* CNN lorsque peu de données sont disponibles, une série d'entraînements fut conduite en ne conservant qu'une partie des 35 images d'entraînement de MAPLES-DR (cf. Figure 5.13). Alors que l'AUC-ROC des CNN standards chute de manière significative lorsque le nombre d'échantillons d'entraînement est considérablement réduit, les *steered* CNN conservent des performances raisonnables sur MAPLES-DR et se détériorent moins rapidement sur HRF. Ainsi, même lorsqu'ils sont entraînés sur seulement 7 images (20% de l'ensemble d'entraînement de MAPLES-DR), les *steered* CNN atteignent une AUC-ROC de 0,91 alors que l'AUC-ROC des CNN standards tombe à 0,86. Ce résultat concorde avec la plus faible dépendance des *steered* CNN à l'augmentation de données géométrique : la mutualisation paramétrique induite par le formalisme *steered* permet à chaque filtre convolutif d'être entraîné sur l'ensemble des vaisseaux d'une image sans les distinguer par leur orientation. En quelque sorte, à nombre d'images d'entraînement égal, chaque filtre d'un *steered* CNN bénéficie de plus « d'exemples de vaisseaux » desquels il peut apprendre.

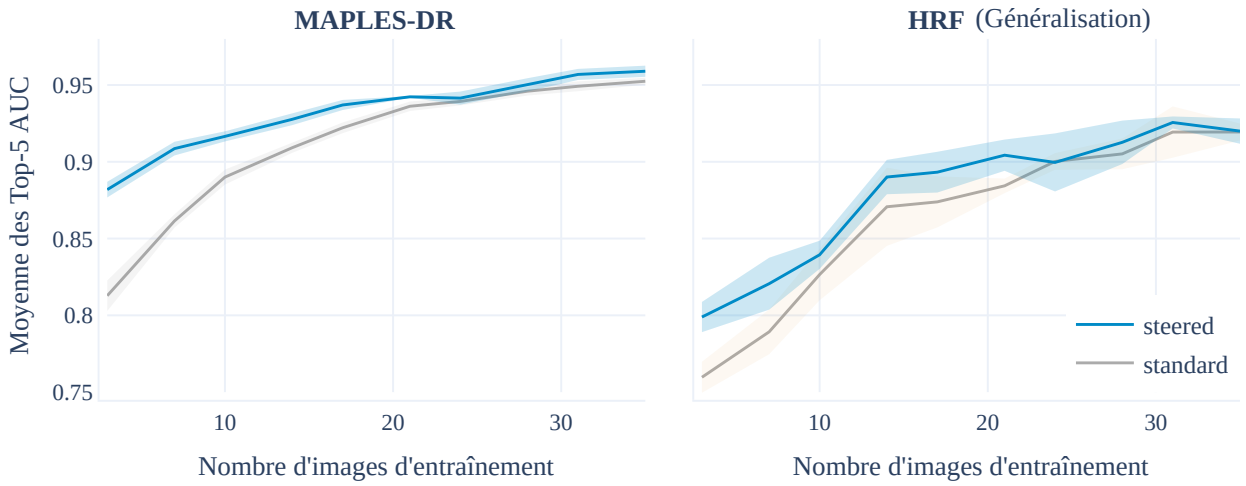


FIGURE 5.13 Performances comparées du *steered* CNN avec un CNN standard en fonction du volume d'entraînement.

Rapidité de convergence du modèle

Non seulement les SCN nécessitent moins d'échantillons de données pour converger, mais ils requièrent également moins d'itérations d'entraînement. Pour mesurer cette vitesse d'apprentissage accrue, nous avons modélisé l'évolution de l'AUC de validation au cours des entraînements par des courbes exponentielles : $\text{AUC}_{\text{val}} = b \cdot \exp(-a \cdot t) + c$ où t est le nombre d'itérations d'entraînement et a, b et c les paramètres de la courbe. En posant $\lambda = \frac{\log(2)}{|a|}$, on peut extraire de ces courbes un indicateur de « demi-vie », c'est-à-dire le nombre d'itérations d'entraînement nécessaires pour réduire de moitié la différence entre la valeur courante de l'AUC de validation et sa valeur finale (voir Figure 5.14).

La distribution de ces indicateurs de demi-vie est représentée sur la Figure 5.15. Toutes les configurations testées pour les expériences précédentes y sont agrégées. En moyenne, l'AUC-ROC de validation des modèles *steered* converge en 2,4 fois moins d'itérations que les modèles standards. En d'autres termes, les SCN facilitent l'exploration de l'espace des paramètres pour y chercher le modèle optimal.

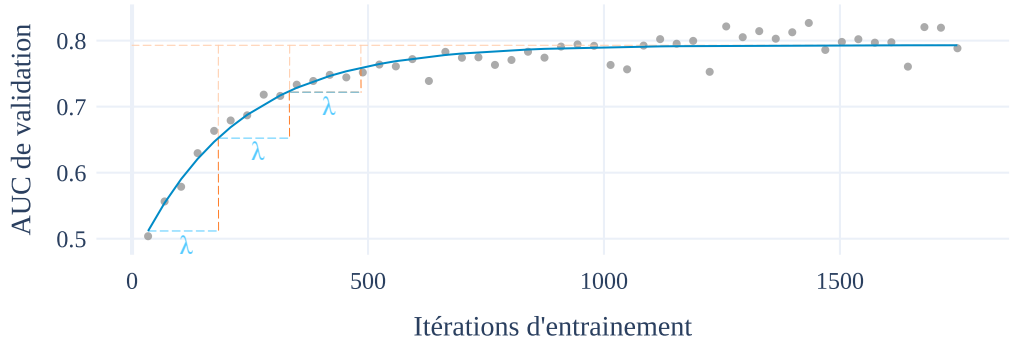


FIGURE 5.14 Exemple de regression par une courbe exponentielle de l'évolution de l'AUC-ROC de validation ; et de mesure géométrique de la demi-vie λ.

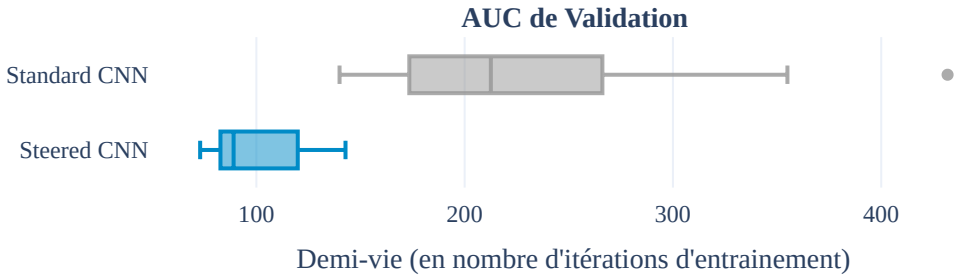


FIGURE 5.15 Rapidité de convergence lors de l'entraînement de CNN *steered* et standard.
La « vitesse » est ici quantifiée par la demi-vie de l'AUC-ROC calculée sur l'ensemble de validation. Plus la demi-vie est faible, plus la convergence est rapide. Les distributions représentées agrègent toutes les courbes d'entraînement des expériences précédentes.

5.5 Discussion et Contributions

Toutes nos expériences montrent que le remplacement des neurones convolutifs standard par des SCN augmente les performances de classification A/V des vaisseaux rétiniens. Elles mettent aussi en évidence une amélioration de la modélisation de cette tâche qui se traduit par une efficacité paramétrique accrue, une convergence plus rapide et une moindre vulnérabilité au surapprentissage lorsque peu de données d'entraînement sont disponibles. Par ce dernier point, les *steered* CNN ont démontré une meilleure capacité de généralisation.

La nature de la tâche de classification A/V n'est probablement pas étrangère à ces gains de performance. En effet, avant les modèles entraînés par apprentissage profond, les meilleurs algorithmes de classification vasculaires reposaient sur des descripteurs dépendants de l'orientation des vaisseaux (par exemple, les caractéristiques des dérivées de gaussiennes et la mesure de l'épaisseur des vaisseaux [78] ou l'analyse des reflets le long des artères [82]). On a d'ailleurs vu que ces descripteurs sensibles à l'orientation sont essentiels aux *steered* CNN : lorsqu'on limite les filtres élémentaires à des composantes purement radiales, l'exactitude du modèle chute de 10% (voir Figure 5.8). Pour apprendre de tels descripteurs, les CNN standards doivent cependant consacrer plusieurs paramètres à l'apprentissage de filtres anisotropiques selon différentes orientations et doivent être entraînés sur de nombreux exemples de vaisseaux dans chaque orientation pour apprendre ces paramètres. En pratique, et à plus forte raison lorsqu'ils sont entraînés sous augmentation de données géométriques de rotation, les filtres de convolution des CNN standards convergent plus facilement vers des filtres isotropiques. Nous pensons que la formulation de la rétropropagation du gradient dans les CNN standards, où les signaux de gradient générés par chaque vaisseau présent dans un mini-batch sont agglomérés ensemble – y compris des vaisseaux ayant des orientations mutuellement orthogonales – explique ce biais isotropique. Ainsi, les bénéfices des SCN dépassent la simple amélioration de l'efficacité paramétrique : la reparamétrisation des neurones convolutifs permet surtout l'apprentissage de descripteurs anisotropiques, là où les CNN favorisent ceux isotropiques.

Enfin l'étude qualitative des *steered* CNN semble indiquer que leur équivariance à la rotation les immunise contre les variations soudaines de prédiction A/V au sein d'un même vaisseau, et induit une meilleure cohérence topologique de leur prédiction.



CHAPITRE 6 FUNDUS VESSELS TOOLKIT : MODÉLISATION GÉOMÉTRIQUE ET TOPOLOGIQUE DE L'ARBRE VASCULAIRE RÉTINIEN

Ce chapitre vise à modéliser topologiquement puis à paramétrer l'arbre vasculaire rétinien à partir de cartes de segmentation sémantiques imparfaites prédites par CNN. Il a été implémenté dans un paquet python accessible sur Github et Pip nommé fundus-vessels-toolkit (FVT). Ce nom sera utilisé pour désigner la méthode décrite dans ce chapitre.

6.1 Enjeux et Motivations

Standardisation de la topologie de l'arbre vasculaire

Comme on l'a vu dans la revue de littérature, la paramétrisation de l'arbre vasculaire rétinien suppose une représentation topologique structurée : la mesure de la tortuosité repose sur l'identification de branches et la paramétrisation des bifurcations nécessite de les distinguer des croisements. Plus précisément, la modélisation topologique finale doit idéalement se conformer aux règles suivantes pour permettre une paramétrisation exhaustive :

- Le réseau vasculaire rétinien est modélisé par une forêt d'arbres.
- Dans cette forêt, chaque arbre représente un vaisseau unique émanant du disque optique et est affublé d'un unique label : artère ou veine.
- Chaque noeud de ces arbres représente soit la racine d'un vaisseau, soit une de ses terminaisons, soit un de ses embranchements (e.g. une branche donnant naissance à 2 branches ou plus). Les coordonnées spatiales de chaque noeud sont connues.
- Chaque arc représente une branche d'un vaisseau et est associé à une séquence de pixels appartenant au squelette vasculaire et formant une courbe dont on peut dériver les tangentes. Le calibre du vaisseau est défini pour chaque point de cette courbe.

Artefacts des segmentations sémantiques par CNN

Les modèles CNN atteignent aujourd'hui de bonnes performances d'exactitude sur la segmentation et la classification des vaisseaux sur images de fond d'œil. Mais ils peinent à apprendre leur structure topologique, de sorte que leurs prédictions contiennent toujours plusieurs types d'artefacts connus :

- des irrégularités de la surface des vaisseaux segmentés induisant des pointes (spurs) sur le squelette vasculaire ;

— des fragmentations de la segmentation sous la forme : soit de trous responsables de l'apparition de cycles érronnés dans le squelette, soit de discontinuités qui isolent certains vaisseaux du reste du graphe vasculaire (voir Figure 6.1 droite) ;
— des erreurs de la classification A/V locales ou pour une branche complète ;
— des erreurs de classification aux croisements de vaisseaux : au lieu d'annoter ces pixels comme veinules ET artérioles, le modèle choisi un type unique, induisant une discontinuité dans la carte de segmentation de l'autre.

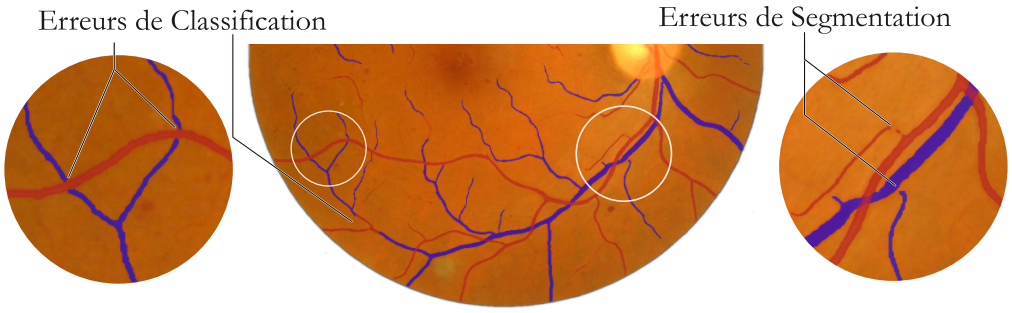


FIGURE 6.1 Artefacts de segmentation sémantique des vaisseaux rétiniens par CNN.

Travaux connexes

Parmi les logiciels récents qui utilisent des cartes de segmentation prédites par CNN pour paramétriser automatiquement le réseau vasculaire rétinien, trois sur quatre abandonnent la modélisation en arbre, au détriment des paramètres qui en dépendent. Le quatrième, RMHAS (Shi et al. 2022 [149]) implémente une méthode pour cette modélisation mais ne tient pas compte des artefacts identifiés plus haut. Ainsi, à la moindre erreur de segmentation ou de classification, la représentation topologique produite est erronée.

Dans leur récente revue de littérature sur les algorithmes de classification A/V du fond d'œil, Chen et al. [26] ont conclu à la nécessité du développement de méthodes de post-traitement en s'inspirant de la littérature antérieure aux CNN (en 2015) et qui abordait la classification A/V comme un problème de partitionnement du graphe vasculaire. C'est-à-dire le problème de l'identification des différents vaisseaux dans le graphe non orienté du squelette où tous, artérioles comme veinules, s'entremêlent. La méthode présentée dans ce chapitre s'inspire particulièrement des travaux d'Estrada et al. [219] qui formulaient le problème comme une recherche de l'arborescence la plus plausible parmi l'ensemble des arborescences générables par le graphe non orienté. On conservera de leur approche le choix d'orienter les branches vasculaires et d'énumérer les arborescences possibles pour trouver la plus plausible. Cependant, leur méthode ne peut opérer que sur un squelette de qualité parfaite : en particulier,

elle ne contient aucun mécanisme pour corriger les éventuelles déconnexions de branche. Elle présuppose aussi que les nœuds racines du graphe sont connus.

À l'inverse, la méthode que je propose ici tient compte des artefacts inhérents à la segmentation sémantique par CNN et même cherche à les corriger. Elle vise à identifier quelle arborescence conforme aux règles topologiques énoncées plus haut pourrait avoir généré l'observation bruitée de la segmentation sémantique qu'on lui présente. Pour ce faire, elle traduit les connaissances cliniques sur la structure des vaisseaux rétiniens en une série d'heuristiques topologiques et géométriques.

6.2 Méthode

La séquence d'algorithme présentée dans ce chapitre a pour but l'extraction d'une topologie standardisée du graphe vasculaire rétinien à partir de cartes de segmentation et de classification imparfaites produites par un réseau de neurones. Pour ce faire, une succession de représentations est construite, chacune s'appuyant sur la précédente pour en corriger les artefacts et standardiser la structure. Le diagramme de la méthode est présenté sur la figure 6.2

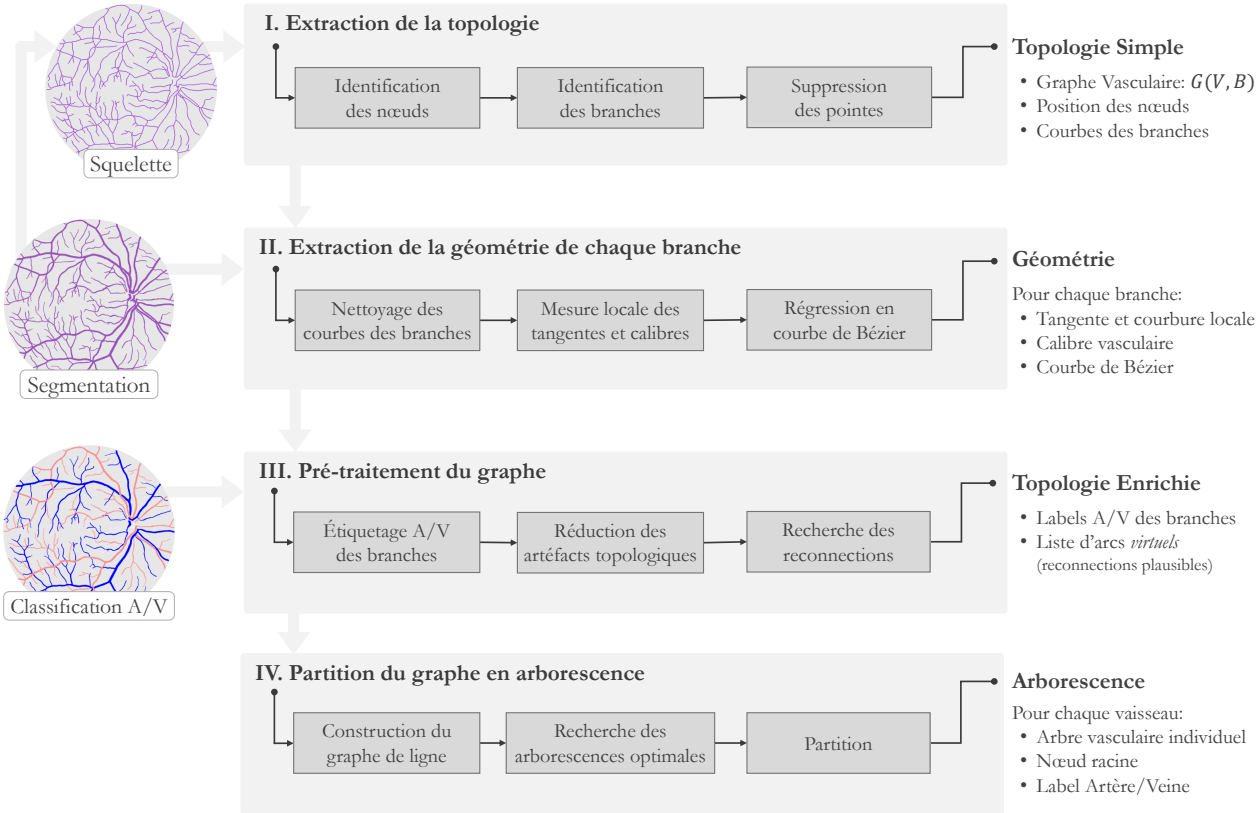


FIGURE 6.2 Vue d'ensemble de l'extraction de l'arbre vasculaire dans FVT.

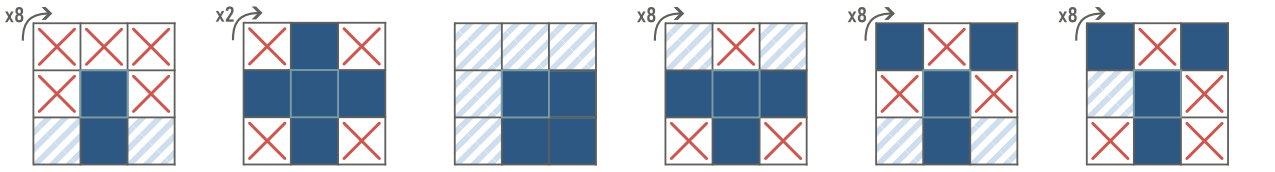
6.2.1 Extraction initiale de la topologie

La première phase consiste à convertir le squelette des vaisseaux en un graphe de branches vasculaires reliées par des nœuds pour obtenir une première modélisation topologique des vaisseaux rétiniens. Pour ce faire, la carte du squelette est dérivée de la segmentation des vaisseaux grâce à la méthode de squelettisation de Lee et al. [220], telle qu'implémentée dans `skimage.morphology`. Puis, chaque pixel du squelette est attribué à une branche ou, s'il est à l'intersection de plusieurs, à un nœud. Cette phase est très similaire aux travaux de la littérature, bien qu'elle s'en distingue déjà par la correction de certains artefacts du squelette.

Détection des noeuds du graphe : les terminaisons et les jonctions

Pour identifier les terminaisons et jonctions vasculaires dans le squelette, une majorité de travaux récents comptent, pour chacun de ses pixels, le nombre de voisins qui appartiennent aussi au squelette : si ce nombre est de 3 ou plus, le pixel est considéré comme une jonction ; s'il est exactement égal à un, le pixel est identifié comme une terminaison [149, 165]. Cette méthode est algorithmiquement simple, mais produit des faux positifs aux jonctions.

Nous avons donc préféré une détection explicite de chaque type de nœud par les détecteurs *hit-or-miss* présentés sur la figure 6.3. Ces détecteurs sont appliqués à chaque pixel du squelette. Ceux répondant aux filtres *(a)* sont identifiés comme des terminaisons et ceux activant les filtres *(b-f)* sont annotés comme jonctions. Tous les autres pixels du squelette sont considérés comme appartenant à une branche, on y reviendra.



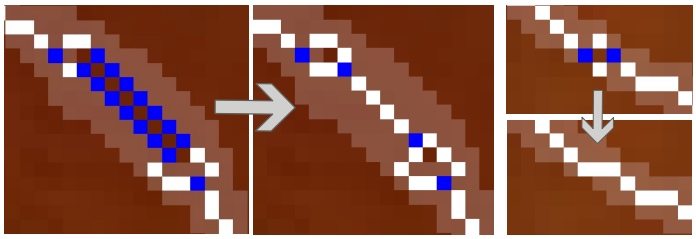
(a) Terminaison (b) Jonction X (c) Jonction carrée (d) Jonction T (e) Jonction Y (f) Jonction Y (bis)

FIGURE 6.3 Masques de détecteurs *hit-or-miss* pour la détection des noeuds du graphe.
Pour activer un détecteur, tout les pixels en bleu foncé doivent appartenir au squelette et tout ceux en rouge doivent en être absent. Les masques (a), (d), (e), et (f) sont appliqués 8 fois, en subissant une rotation de 45° entre chaque application. Le masque (b) est de la même manière appliqué deux fois.

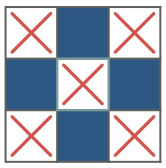
Correction des croix creuses

Une première étape de correction d'artefacts intervient immédiatement après la détection des nœuds. En effet, il arrive qu'un pixel soit absent de la segmentation vasculaire, ou que deux vaisseaux mitoyens soient correctement détectés mais séparés par une ligne diagonale

d'un unique pixel de large. La squelettisation produit alors une série de croix creuses qui sont interprétées de manière erronée comme une succession de jonctions en T (voir figure 6.4a). Pour corriger cette erreur, un détecteur *hit-or-miss* supplémentaire est chargé de leur identification, puis ces croix sont effacées et le squelette est reformé. Suite à ces corrections, tous les nœuds restants sont ajoutés au graphe, et leurs coordonnées spatiales sont sauvegardées.



(a) Exemples de CC et de leurs corrections



(b) Détecteur de CC

FIGURE 6.4 Illustration de la correction des croix creuses (CC).

Découverte des branches qui relient les noeuds

Il reste à découvrir les branches qui relient les nœuds précédemment identifiés. Pour ce faire, le voisinage de chaque jonction est inspecté. On cherche dans un premier temps les paires de nœuds directement adjacents afin d'ajouter immédiatement les arcs correspondants au graphe. Puis, une branche est créée pour chaque pixel du voisinage qui appartient au squelette et qui n'a pas encore été étiqueté ni comme nœud, ni comme branche. Le pixel est annoté du numéro de la branche pour ne pas être visité à nouveau, et on cherche parmi les 5 pixels de son voisinage diamétralement opposés au nœud de départ celui qui appartient au squelette. Cette dernière étape est répétée en analysant itérativement le voisinage des pixels identifiés, et ce jusqu'à trouver un nœud. On ajoute alors au graphe un arc reliant le nœud de départ à celui d'arrivée, et on lui associe des informations complémentaires liées à la branche. Pour l'instant, ces informations comprennent uniquement le numéro de la branche et la séquence de pixels définissant sa courbe.

Une seconde correction du squelette intervient alors : on supprime toutes les branches dont au moins une des extrémités est une terminaison vasculaire et dont la longueur est inférieure à 2. Elles correspondent en effet à des aberrations (dites *spurs* ou pointes) produites par la squelettisation lorsque la surface d'un vaisseau varie rapidement. Les nœuds qui ne sont connectés qu'à deux branches après cette étape sont supprimés et leurs branches sont fusionnées. Cette correction ne supprime pas toutes les *spurs* mais simplifie le graphe avant la conduite d'une analyse poussée de la géométrie des branches.

6.2.2 Extraction de la géométrie de chaque branche

À la fin de la première phase du traitement, la vasculature rétinienne est représentée par un graphe non orienté de branches vasculaires dont on connaît les courbes que forment leurs squelettes. La seconde phase du traitement combine cette représentation topologique simple avec la segmentation des vaisseaux pour enrichir le graphe d'une modélisation géométrique de chacune de ses branches.

Nettoyage des extrémités des branches

Aux extrémités des branches (aux abords des nœuds), le squelette vasculaire n'est pas fiable : attiré par les branches émergentes, il s'écarte du centre des vaisseaux principaux et prend la forme d'un H aux intersections (cf. Figure 6.5). Les algorithmes d'extraction de graphes vasculaires remédient généralement à cette faillibilité en masquant tous les pixels situés dans un rayon fixe autour des jonctions. Mais ce faisant, ils risquent de supprimer des pixels valides du squelette (si le rayon est trop grand) ou de conserver des pixels erronés (s'il est trop étroit). Nous proposons plutôt de nettoyer chaque extrémité de branche en utilisant ses caractéristiques géométriques.

Soit les coordonnées $p_i = \left(p_i^{(x)}; p_i^{(y)}\right)$ du i^{ieme} pixel de la courbe d'une branche, nous définissons la tangente à l'extrémité du squelette comme suit :

$$\hat{t}_i = \sum_{j=1}^{10} \mathcal{G}(j) \cdot \left(p_{i+j} - p_i\right) \tag{6.1}$$

où $\mathcal{G}(j)$ est une gaussienne centrée en 0 et avec un écart-type de 3 pixels.

Soit aussi dL_i et dR_i : les distances séparant p_i des deux bords les plus proches du vaisseau dans une direction orthogonale à $\hat{t}_i$ et soit le calibre du vaisseau $c_i = dR_i + dL_i$. En partant de l'extrémité p_0 d'une branche, chaque pixel du squelette p_i est testé en fonction des quatre conditions suivantes :

1. Les bords du vaisseau R_i et L_i ont été trouvés ;
2. Le squelette se trouve près du centre du vaisseau ;
3. Les calibres c_i et c_{i+1} sont approximativement les mêmes ;
4. Le pixel p_i est le pixel du squelette le plus proche de R_i et L_i.

Le premier pixel qui remplit toutes les conditions est choisi comme l'extrémité de la branche, et tous ceux qui le précèdent sont supprimés (voir Figure 6.5). Si l'extrémité de la branche correspond à une terminaison vasculaire, la condition 4 est remplacée par une vérification de la stabilité de la tangente, garantissant que $\hat{t}_i \approx \hat{t}_{i+1}$.

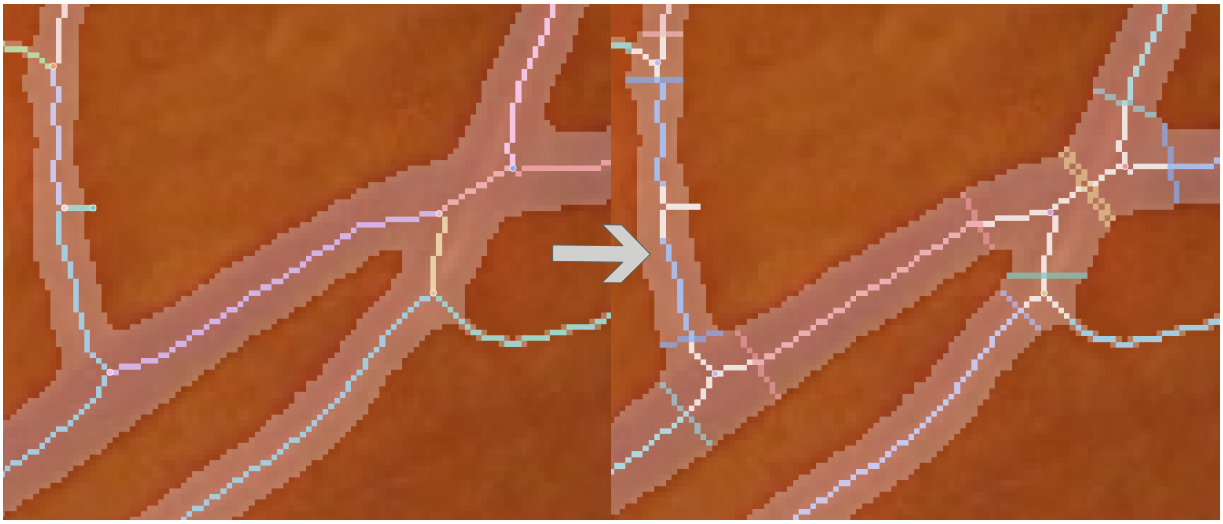


FIGURE 6.5 Nettoyage du squelette aux extrémités des branches.

À droite, les extrémités des branches sont matérialisées par des traits orthogonaux aux vaisseaux. Les portions du squelette qui apparaissent en blanc sur cette image n'appartiennent à aucune branche.

Suite au nettoyage de leurs extrémités, on élimine toutes les branches terminales dont la longueur est inférieure au calibre de la branche adjacente la plus large. Cette méthode de détection et suppression des *spurs* est plus fiable que celle généralement utilisée qui consiste à supprimer, de façon indifférenciée, toutes les branches de longueur inférieure à un seuil fixe. Notre méthode permet en effet de conserver les petites branches fines en périphérie de la vasculature tout en supprimant les longs *spurs* des vaisseaux les plus larges. Le nettoyage du squelette proposé plus haut est clé dans cette démarche car il permet de mesurer la longueur réelle des branches plutôt que la longueur de leur squelette qui comprend en plus la portion entre le centre et la surface du vaisseau parent (si le vaisseau est large, cette différence est importante). Les nœuds orphelins – qui ne sont connectés à aucune branche – sont également supprimés du graphe.

Calcul des propriétés géométriques des courbes des branches

Une fois nettoyé, le squelette de chaque branche est analysé point par point, pour en extraire les propriétés géométriques. À commencer par la tangente au point p_i, calculée par addition des points subséquents p_{i+j} et soustraction des points précédents p_{i-j} du squelette :

$$t_i = \sum_{j=0}^{J} \frac{\mathcal{G}(j)}{2} \cdot \left(p_{i+j} - p_{i-j} \right) \tag{6.2}$$

Ici, le réglage de J et de l'écart-type de la gaussienne $\mathcal{G}$ définit la sensibilité de t_i aux changements d'orientation : plus ils sont faibles, plus la tangente sera à même de capter les

changements rapides de la direction du vaisseau. Néanmoins, pour assurer une mesure fiable, l'écart-type doit être d'au moins 3 pixels. En effet, la position relative de p_{i+n} et p_{i-n} dans le voisinage proche de p_i est très sensible au bruit de discrétisation du squelette : par exemple, une mesure de la tangente qui ne tiendrait compte que de p_{i+1} verrait sa résolution angulaire limitée à 45°.

À partir de la tangente ainsi lissée, on définit la courbure $\mathcal{K}_i$ au point p_i comme une fonction de la tangente t_i et de sa dérivée $\delta t_i = t_i - t_{i-1}$. D'après Grisan et al. [153] :

$$\mathcal{K}_i = \frac{t_i^{(x)} \delta t_i^{(y)} - t_i^{(y)} \delta t_i^{(x)}}{\|\delta t_i\|^3} \tag{6.3}$$

Enfin, le calibre du vaisseau est mesuré à chaque point p_i par la même procédure que celle utilisée pour nettoyer les extrémités des branches : En partant de p_i, deux rayons sont émis dans des directions opposées, perpendiculairement à t_i. On nomme p'_i et p''_i les points d'intersection de ces rayons avec la bordure du vaisseau (selon la segmentation vasculaire). Le calibre de la branche au point p_i est alors défini comme la distance entre p'_i et p''_i.

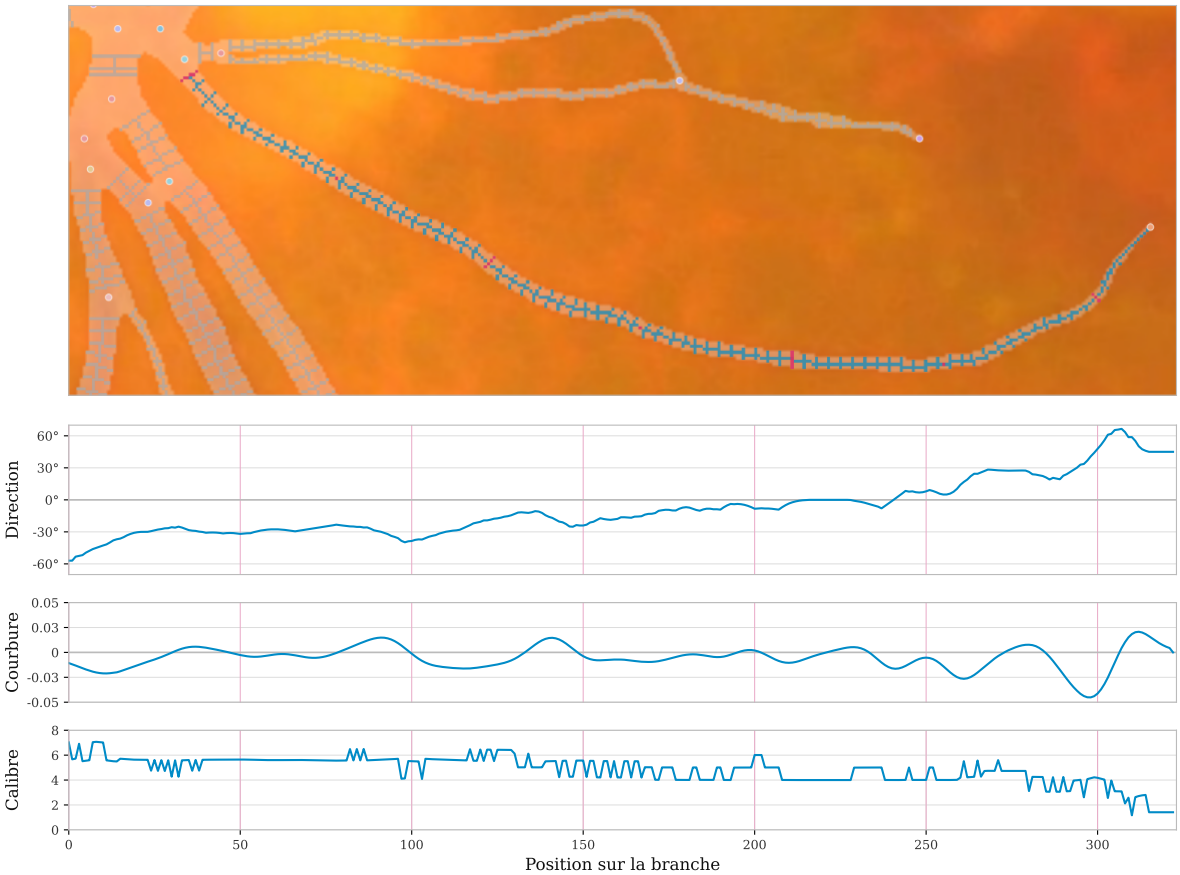


FIGURE 6.6 Informations géométriques extraites pour une branche vasculaire.

Modélisation des courbes vasculaires par des B-splines cubiques

La modélisation géométrique de branches est enrichie d'une dernière propriété : une paramétrisation de leur courbe par des B-splines cubiques. L'objectif est double : d'une part, faciliter la manipulation mathématique des courbes des branches, par exemple pour des calculs de distances ou d'intersections ; et d'autre part, obtenir un descripteur stable de la géométrie vasculaire qui, comparé au squelette brut, soit moins sensible aux variations de segmentation.

Notre méthode pour cette modélisation s'inspire de l'algorithme proposé par Philip J. Schneider [221] qui convertit une courbe discrétisée en une spline de Bézier cubique. Soit une courbe discrète formée par la succession de n points : $p_i = \left(p_i^{(x)}, p_i^{(y)}\right)$ avec $i \in \mathbb{N}$ et $i < n$, l'algorithme commence par l'approximer avec un unique polynôme de Bézier cubique :

$$\mathrm{P}(u) = \left(\mathrm{P}^{(x)}; \mathrm{P}^{(y)}\right) = (1-u)^3\mathrm{P}_0 + 3u(1-u)^2\mathrm{P}_1 + 3u^2(1-u)\mathrm{P}_2 + u^3\mathrm{P}_3 \tag{6.4}$$

Où P_0 et P_3 sont les points de départ et d'arrivée, P_1 et P_2 sont les points de contrôle et $u \in [0\,;1]$ est le paramètre de la courbe. Soit la discrétisation $D : i \to u_i$ qui associe à chaque point p_i de la courbe sa valeur interpolée $\mathrm{P}(u_i)$, on définit l'erreur quadratique moyenne de reconstruction par : $\varepsilon_{\mathrm{MSE}} = \sum_{i=0}^{n} \|\mathrm{P}(u_i) - p_i\|^2$. Si P_0, P_3 et leurs tangentes sont fixées, le choix optimal de P_1 et P_2 pour minimiser l'erreur quadratique $\varepsilon_{\mathrm{MSE}}$ est solvable exactement et avec une complexité $O(n)$. Néanmoins, la qualité de la modélisation dépend de la discrétisation D qui doit aussi être optimisée par une routine de Newton-Raphson afin de trouver, pour chaque valeur de i, le paramètre u_i qui minimise la distance $\varepsilon_i = \|p_i - \mathrm{P}(u_i)\|$. L'algorithme optimise donc alternativement D puis les coefficients P_1 et P_2, jusqu'à ce que l'erreur maximale de reconstruction : $\varepsilon_{\max} = \max_i \|\mathrm{P}(u_i) - p_i\|$ soit inférieure à l'erreur désirée $\varepsilon_{\text{cible}}$. Si cette erreur cible n'est pas atteinte après dix itérations, ou si $\varepsilon_{\max}$ est supérieure au double de $\varepsilon_{\text{cible}}$, la courbe est divisée en deux et l'algorithme est appliqué individuellement aux deux sections. Si, à nouveau, une section ne peut être approximée par une unique courbe de Bézier cubique, elle est elle-même sous-divisée et ce récursivement jusqu'à ce que $\varepsilon_{\max}$ soit inférieure à $\varepsilon_{\text{cible}} = 2$ pixels (déterminé empiriquement). La B-spline peut alors être formée par l'assemblage des courbes de Bézier ainsi obtenues.

Dans son implémentation initiale, Philip J. Schneider [221] subdivise les courbes au point où ε_i est maximale. Mais ces points sont particulièrement sensibles aux variations de segmentation, et ils sont donc insatisfaisants pour devenir les nœuds d'une B-spline stable. Nous choisissons plutôt de forcer leur placement sur les points d'inflexion de la courbe. Dans notre implémentation, les racines de la courbure de la branche sont identifiées en amont comme de potentiels nœuds de la B-spline. Lorsqu'il doit subdiviser une courbe, l'algorithme choisit

le point d'inflexion pour lequel l'erreur cumulée est répartie le plus équitablement de part et d'autre du point. Deux erreurs interviennent dans notre calcul : la distance euclidienne ε_i et l'erreur cosinus entre les tangentes des courbes réelle et approximée (voir l'illustration Figure 6.7).

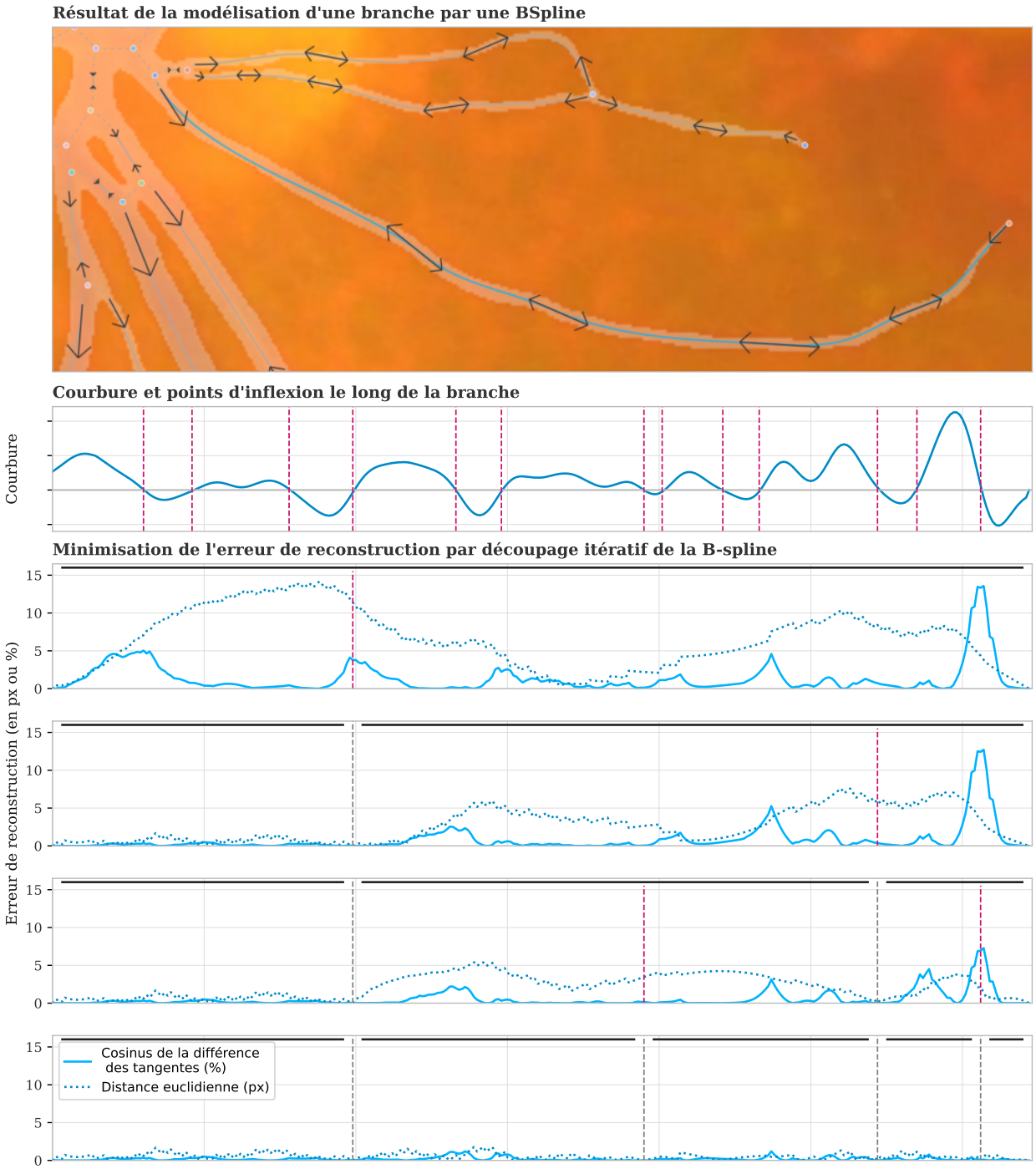


FIGURE 6.7 Découpage itératif de la courbe d'une branche en une B-spline cubique.
Le premier graphique indique la position des points d'inflexions en rouge. Les quatres graphiques suivants présentent, de haut en bas, chaque subdivision de la courbe et la diminution de l'erreur de reconstruction.

6.2.3 Intégration de la classification A/V au graphe vasculaire

Pour l'instant, seule la segmentation des vaisseaux a été utilisée pour construire le graphe et ses propriétés géométriques. La troisième étape du traitement enrichit la représentation du graphe vasculaire à partir de l'information fournie par la classification A/V.

Coloration des branches et noeuds comme artériole ou veinule

On souhaite colorier chaque branche et chaque nœud du graphe d'une couleur définissant leur classe : **ART** pour les artérioles et **VEI** pour les veinules. Cette tâche est moins triviale qu'il n'y paraît car la segmentation vasculaire sur laquelle notre travail (et l'ensemble de la littérature) s'appuie n'est pas une segmentation par instance. Par conséquent, lorsque deux branches de vaisseaux se rapprochent suffisamment l'une de l'autre, leurs segmentations s'unissent et leurs squelettes fusionnent. Dans le graphe $G(V, B)$ une seule branche rend alors compte de ces deux vaisseaux, qui sont pourtant bien souvent de classes différentes (puisque les artérioles et les veinules s'alternent sur la rétine). Une troisième couleur **BOTH** est donc définie pour ces branches à la fois **ART** et **VEI**.

Le coloriage des branches débute par l'assignation d'une couleur à chacun des pixels p_i de leur squelette. Pour ce faire, trois pixels de la carte de classification A/V sont étudiés : p_i lui-même ainsi que les points p'_i et p''_i placés sur la bordure du vaisseau, à l'intersection d'une droite orthogonale à t_i (voir section 6.2.2). Si ces trois pixels sont **ART** (resp. **VEI**), p_i est étiqueté **ART** (resp. **VEI**) ; sinon, il est étiqueté **BOTH**. Une fois que l'ensemble de ses pixels a ainsi été colorié, la couleur globale de la branche est déduite comme suit : si plus de 2/3 des pixels d'une branche partagent une couleur commune, cette dernière est attribuée à la branche ; sinon, la branche est divisée en sections de couleur homogènes. Ce seuil permet de corriger les petites "taches" qui apparaissent parfois sur les cartes de classification A/V prédites par CNN. Enfin, les branches de moins de deux pixels après l'étape de nettoyage du squelette sont coloriées d'une quatrième couleur ***UNK*** : leur classe A/V étant inconnue.

Les nœuds sont alors initialisés avec la classe **UNK**, puis leur couleur est définie en fonction de celles de leurs branches incidentes :

- Si toutes sont **ART** (resp. **VEI**), le noeud est **ART** (resp. **VEI**) ;
- Si certaines sont **ART** et d'autres **VEI** ou si au moins une est **BOTH**, le nœud est **BOTH** ;
- Si un nœud a deux branches incidentes, dont une **UNK**, la couleur de l'autre branche est propagée au noeud et à la branche **UNK**.

Puisque cette dernière règle modifie la couleur d'une branche, la procédure de coloriage des

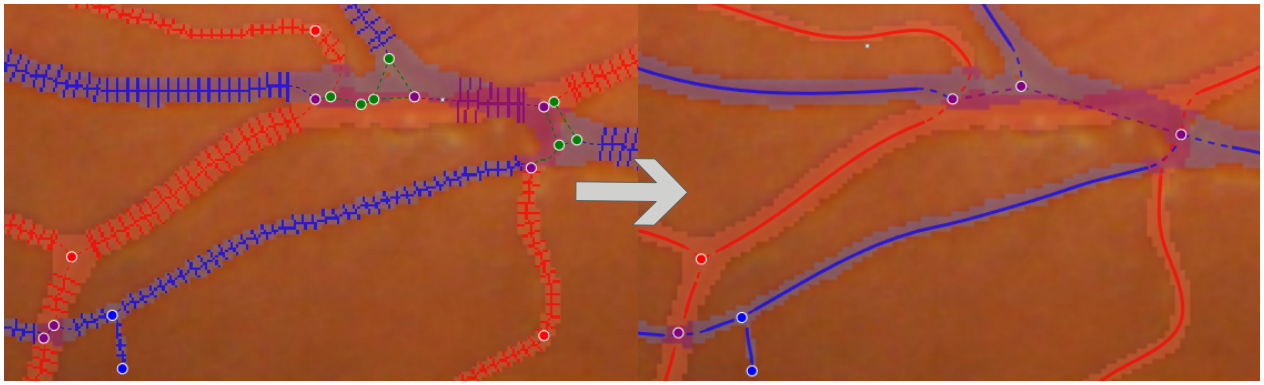


FIGURE 6.8 Correction des artefacts topologiques. **Gauche :** Graphe colorié par les classes A/V (*rouge : **ART** ; bleu : **VEI** ; violet : **BOTH** ; vert : **UNK***). **Droite :** Graphe simplifié.

nœuds est répétée itérativement jusqu'à ce que leur couleur soit stable.

Enfin, on identifie par un algorithme de composantes connexes les groupes de nœuds **UNK**. Si toutes les branches incidentes d'un groupe sont **ART** (resp. **VEI** ou **BOTH**), ses nœuds sont recolorés **ART** (resp. **VEI** ou **BOTH**). La figure 6.8 montre un exemple de graphe A/V colorié selon cette méthode.

Simplification topologique du graphe A/V

Le graphe vasculaire extrait à partir d'une segmentation automatique contient parfois des structures topologiques inutilement complexes, en particulier lorsqu'il est construit sans distinction des artérioles et des veinules. Ces artefacts topologiques ont été conservés jusqu'ici pour ne pas perturber la coloration du graphe. Celle-ci étant faite, on peut dorénavant les simplifier par la procédure suivante (illustrée par la Figure 6.8) :

1. Tout nœud reliant exactement deux branches de la même couleur qui forment un angle obtu est supprimé puisqu'il est topologiquement inutile, et ses deux branches incidentes sont fusionnées en une seule.
2. Les branches orphelines (connectées à aucune autre branche) dont la longueur de la courbe est inférieure au calibre D_{max} du vaisseau le plus large sont aussi supprimées.
3. Les jonctions adjacentes de même couleur sont fusionnées lorsque leur distance est inférieure à $D_{\max}$. Si une jonction pourrait ainsi s'associer avec plusieurs autres, toutes sont regroupées et fusionnées ensemble. Le nœud résultant est placé à leur barycentre et est connecté à toutes les branches incidentes au groupe, tandis que les branches qui reliaient les jonctions entre-elles sont effacées.
4. L'étape 1. de cette procédure est exécutée à nouveau pour supprimer tout noeud qui

aurait été rendu topologiquement inutile par les simplifications précédentes.

5. La routine de coloriage des noeuds est exécutée à nouveau pour que la coloration A/V du graphe soit conforme à sa nouvelle topologie.

Recensement des reconnections potentielles

Il n'est pas rare que les modèles de segmentation vasculaire probabilistes fractionnent par erreur la segmentation des petits vaisseaux, en particulier proches des intersections. Ce fractionnement déconnecte certaines branches du reste du graphe vasculaire et induit ainsi des artefacts topologiques : les arcs qui auraient dû matérialiser ces connexions sont absents, et des nœuds terminaux apparaissent à l'emplacement des fins prématurées des vaisseaux. Suite à la simplification du graphe à l'étape précédente, une recherche est conduite parmi tous les nœuds terminaux pour identifier ceux qui résulteraient d'une telle déconnexion.

Pour ce faire, la tangente à l'extrémité de chaque terminaison vasculaire est prolongée dans un cône large de 20° et long de 100 pixels. Toutes les branches dont le squelette intersecte ce cône sont subdivisées en insérant un nœud au point d'intersection, puis un arc *virtuel* est créé entre ce nouveau nœud et celui au sommet du cône. Si l'intersection est proche de l'extrémité d'une branche (au maximum $D_{\max}$), elle n'est pas subdivisée et l'arc *virtuel* est connecté à son nœud existant. Ces arcs *virtuels* ne sont pas directement ajoutés au graphe : ils sont transmis en tant qu'arcs hypothétiques à l'algorithme d'optimisation d'arborescence vasculaire décrit dans la section suivante. C'est ce dernier qui évaluera la plausibilité de chacun et retiendra uniquement ceux qui améliorent la vraisemblance globale du graphe.

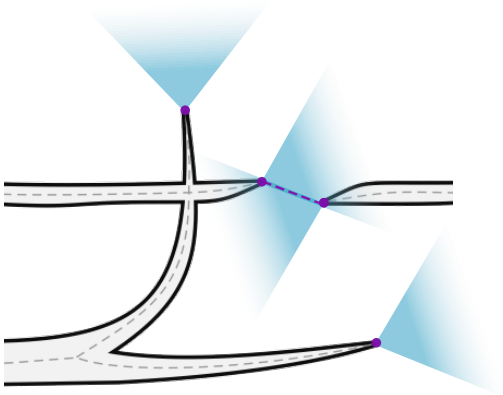


FIGURE 6.9 Illustration de la recherche des reconnections possibles.

6.2.4 Résolution de l'arborescence propre à chaque vaisseau

À ce stade, une majorité des informations présentes dans les cartes de segmentation et de classification vasculaire ont été traduites par des attributs du graphe $G(V, B)$. Mais cette représentation s'arrête à la notion de branches : artérioles et veinules sont enchevêtrées au sein du même graphe où les différents vaisseaux de la rétine se confondent. L'étape finale du traitement vise à identifier dans ce graphe les sous-arbres propres à chaque vaisseau, puis à les représenter par une arborescence composée d'une branche principale et de branches secondaires, tertiaires, etc. C'est véritablement la clé de voûte de la méthode : elle s'appuie sur toutes les informations extraites par les étapes précédentes pour désentrelacer les vaisseaux, trouver leur racine et leurs terminaisons, tout en corrigeant les artefacts résiduels de segmentation ou de classification.

Pour y parvenir, cette ultime étape ajoute aux informations précédentes la position du disque optique et de la macula fournies par une segmentation de ces structures. Elle s'appuie ensuite sur les connaissances cliniques de la structure des graphes vasculaires rétiniens, traduites en heuristiques, pour estimer la plausibilité des différentes hypothèses d'arborescences et sélectionner la plus crédible. Mais auparavant, elle débute par une énumération exhaustive de toutes les arborescences que $G(V, B)$ peut générer. Cette énumération prend la forme d'un graphe de lignes.

Construction du graphe de lignes

La réduction du graphe $G(V, B)$ en une forêt de sous-arbres enracinés consiste, mathématiquement parlant, à trouver le morphisme $\mathcal{P} : B \to B \cup \varnothing$ qui associe à chaque branche $b_i \in B$ sa branche parente b_j tel que :

$$\mathcal{P} : b_i \mapsto \begin{cases} b_j \in B & \text{si } b_i \text{ est une branche secondaire} \\ \varnothing & \text{si } b_i \text{ est primaire et n'a pas de parent} \end{cases}$$

Si on contraint $\mathcal{P}$ à ne contenir aucun cycle, alors le morphisme génère bien un arbre orienté dont la racine est $\varnothing$ et où toute branche $b_p \in B$ telle que $\mathcal{P}(b_p) = \varnothing$ est une branche principale d'un vaisseau rétinien. Bien sûr, le choix de $\mathcal{P}$ peut drastiquement modifier l'allure du graphe vasculaire. Pour s'en convaincre, prenons le graphe vasculaire $G(V, B)$ à 5 nœuds et 4 branches présenté sur la figure 6.10. Fixons $\mathcal{P}(\overline{13}) = \mathcal{P}(\overline{23}) = \varnothing$, si $\mathcal{P}(\overline{34}) = \overline{13}$ et $\mathcal{P}(\overline{35}) = \overline{23}$ on obtient l'arborescence *b1*, si $\mathcal{P}(\overline{34}) = \overline{23}$ et $\mathcal{P}(\overline{35}) = \overline{13}$ on obtient celle *b2*, si enfin $\mathcal{P}(\overline{34}) = \mathcal{P}(\overline{35}) = \overline{13}$ on a celle *b3*. Au total, plus de 30 arborescences différentes peuvent être générées par cette simple intersection. Évidemment, toutes ne sont pas pertinentes : l'arborescence *b*4 est par exemple très peu probable.

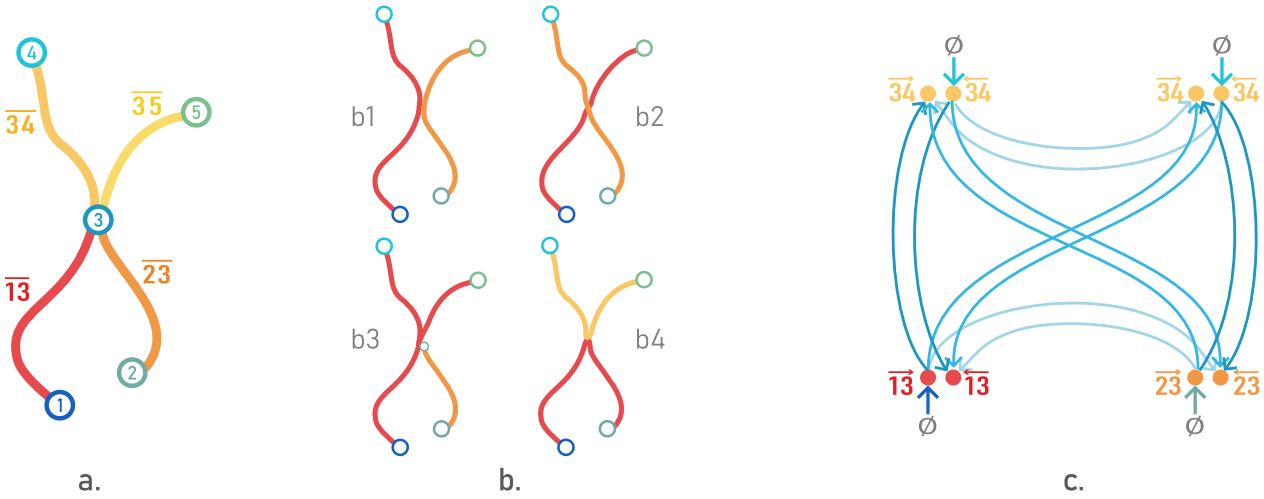


FIGURE 6.10 Illustration de la construction du graphe de ligne. **a.** Graphe vasculaire $G(V, B)$; **b.** Exemple d'arborescence que G peut générer ; **c.** Graphe de ligne listant toutes les arborescences possibles.

L'ensemble des arborescences peut être représenté par un graphe de ligne dirigé $L(G)$ dérivé du graphe vasculaire $G(V, B)$ comme suit. Pour chaque branche $b_{uv} \in B$ reliant les nœuds $u, v \in V$, deux nœuds sont ajoutés à $L(G)$: $\bar{b}_{\overrightarrow{uv}}$ et $\bar{b}_{\overrightarrow{vu}}$. Le premier suppose que b_{uv} est orientée de u vers v tandis que le second suppose l'inverse. Les nœuds de $L(G)$ sont coloriés de la même classe A/V que la branche qu'ils représentent, à l'exception près que les branches **BOTH** sont dédoublées dans le graphe de ligne : une paire de nœuds la représentant dans sa version **ART**, et une autre dans sa version **VEI**.

On considère ensuite un nœud $u \in V$ du graphe vasculaire et ses nœuds adjacents $\mathrm{Adj}(u) \subset V$: pour chaque combinaison de nœuds adjacents $(v_n, v_m) \in \mathrm{Adj}(u)$ avec $n \neq m$, un arc de ligne est ajouté à $L(G)$ dirigé de $\bar{b}_{\overrightarrow{v_n u}}$ vers $\bar{b}_{\overrightarrow{uv_m}}$. Si u est connecté à N branches vasculaires (et donc à autant de nœuds), on ajoute ainsi $N \times (N-1)$ arcs dans le graphe de ligne. Chacun de ces arcs représente une hypothèse de relation hiérarchique entre branches : en l'occurrence que $b_{\overrightarrow{v_n u}}$ donne naissance à $b_{\overrightarrow{uv_m}}$ à travers le nœud u.

Les points d'extrémité et les nœuds proches des limites du disque optique ou du fond d'œil sont considérés comme des racines plausibles de leurs sous-arbres. Pour chaque nœud u parmi ces candidats, un arc est créé dans $L(G)$, dirigé de $\varnothing$ vers la branche $\bar{b}_{\overrightarrow{ub}}$,

Pondération du graphe de ligne

Chaque arc du graphe de ligne est affublé d'un score de plausibilité calculé par des heuristiques qui traduisent la plausibilité de l'hypothèse portée par l'arc selon les connaissances cliniques.

— Ces heuristiques favorisent l'hypothèse que $b_{\overrightarrow{v_n u}}$ soit la branche parente de $b_{\overrightarrow{uv_m}}$ si **1.** leur tangentes se font face ; **2.** elles ont la même classe A/V, ou **3.** la branche mère a un calibre supérieur à la branche fille. Plus précisément, le score de l'arc qui porte cette hypothèse est calculé par :

$$\rho\left(b_{\overrightarrow{v_n u}}, b_{\overrightarrow{uv_m}}\right) = \omega_{\hat{t}} \cdot \hat{t}(b_{\overrightarrow{v_n u}}) \circ \hat{t}(b_{\overrightarrow{uv_m}}) \; + \; \omega_{\text{av}} \cdot \mathbf{1}^{\text{av}}(b_{v_n u}, b_{uv_m}) + \; \omega_{\hat{c}} \cdot \sigma\left(\hat{c}(b_{\overrightarrow{v_n u}}) - \hat{c}(b_{\overrightarrow{uv_m}})\right)$$

Où $\hat{t}(b_{\overrightarrow{v_n u}})$ et $\hat{c}(b_{\overrightarrow{v_n u}})$ sont les tangentes et les calibres à l'extrémité u de $b_{v_n u}$; $\sigma(\cdot)$ est la fonction sigmoïde ; et $\mathbf{1}^{\text{av}}(b_{v_n u}, b_{uv_m})$ vaut 1 si $b_{v_n u}$ et b_{uv_m} ont la même classe A/V et 0 sinon. Les paramètres $\omega_{\hat{t}}$, ω_{av} et $\omega_{\hat{c}}$ pondèrent les différents termes de l'heuristique.

— Les arcs portant l'hypothèse que la branche $b_{\overrightarrow{uv}}$ soit la racine d'un vaisseau sont favorisés si u est proche du disque optique ou du bord de la région d'interêt de l'image de fond d'oeil.

— Enfin la plausibilité de la direction de chaque branche b_{uv} est aussi estimée en ajoutant à tous les arcs de lignes adjacents aux noeuds $\bar{b}_{\overrightarrow{uv}}$ et $\bar{b}_{\overrightarrow{vu}}$, les scores suivants :

$$\rho\left(b_{\overrightarrow{uv}}\right) = -\rho\left(b_{\overrightarrow{vu}}\right) = \omega_t \cdot \left\langle t_i\left(b_{\overrightarrow{uv}}\right) \circ \tilde{t}\left(p_i\right) \right\rangle \; + \; \omega_c \cdot \sigma\left(c_{\text{half}}(b_{\overrightarrow{uv}}) - c_{\text{half}}(b_{\overrightarrow{vu}})\right)$$

où $c_{\text{half}}(b_{\overrightarrow{uv}})$ et $c_{\text{half}}(b_{\overrightarrow{vu}})$ sont le calibre médian sur la première et seconde moitié de b_{uv}, et où $\left\langle t_i\left(b_{\overrightarrow{uv}}\right) \circ \tilde{t}\left(p_i\right) \right\rangle$ est la moyenne de distance cosinus entre la tangente de la branche $t_i\left(b_{\overrightarrow{uv}}\right)$ et celle attendue $\tilde{t}\left(p_i\right)$, calculée pour tous les points p_i de la branche b_{uv}. La direction $\tilde{t}$ est une estimation de l'orientation probable des branches en tout point du fond d'œil : elle pointe à l'opposé du disque optique et vers son symétrique autour de la macula [1]. La direction favorisée pour la branche est donc celle pour laquelle **1.** les tangentes s'éloignent du disque optique et tournent autour de la macula, **2.** le calibre décroit.

1. L'influence de ces deux points est définie comme inversement proportionnelle au carré de leur distance à la manière d'un champ magnétique.

Résolution de l'arborescence la plus plausible

Une fois que $L(G)$ est défini, son arborescence maximale (MSA) est calculée, de sorte à ne conserver qu'un seul arc de ligne dirigé vers chaque branche $\bar{b}_{\overrightarrow{uv}}$ et $\bar{b}_{\overrightarrow{vu}}$. Ensuite, chaque paire de nœuds linéaires $\bar{b}_{\overrightarrow{uv}}$ et $\bar{b}_{\overrightarrow{vu}}$ est fusionnée, formant un graphe linéaire plus simple L^* où les branches vasculaires sont représentées par des nœuds linéaires uniques. La représentation arborescente la plus plausible pour le graphe vasculaire est obtenue en calculant le MSA de L^* (cf. Figure 6.11c). Les sous-arbres de vaisseaux sont ensuite développés en parcourant la MSA à partir de chaque arête sortante de la branche racine ∅. Une fois séparé, chaque sous-arbre se voit attribuer son étiquette A/V majoritaire, ce qui permet de corriger les erreurs de classification A/V locales.

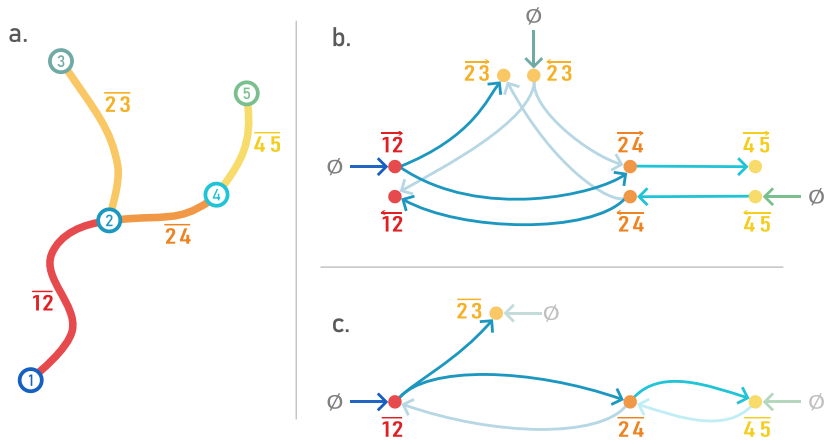


FIGURE 6.11 Illustration de l'optimisation d'arborescence **a.** Graphe vasculaire G ; **b.** Graphe de ligne orienté $L(G)$; **c.** Graphe de ligne simplifié $L^*(G)$. (Les arcs de lignes correspondants aux hypothèses rejetées par la MSA sont semi-transparents.)

Heuristique «garde-fou»

L'étape d'optimisation décrite ici accorde plus d'importance aux caractéristiques géométriques et topologiques extraites de la segmentation des vaisseaux qu'à la classification A/V. Cependant, il arrive qu'un vaisseau semble « émerger » d'un autre parce que son segment en amont a été occulté ou n'a pas été segmenté. Ces rares cas conduisent à des affiliations de branches erronées et à la fusion de vaisseaux distincts en un seul sous-arbre. Une dernière heuristique de sauvegarde détecte ces cas en recherchant les branches croisées qui sont affiliées au même sous-arbre. Si les branches croisées avaient initialement des étiquettes A/V différentes, leurs sous-arbres sont déconnectés.

6.3 Évaluation de la modélisation topologique

L'évaluation de la méthode d'extraction de la topologie par FVT est divisée en deux parties. Dans un premier temps, on évaluera la capacité de l'algorithme de partitionnement du graphe à corriger les artefacts de segmentation et de classification. Puis, dans un second temps, on évaluera la pertinence des paramètres vasculaires extraits avec la librairie sous la forme d'une étude pilote sur la base de données canadienne CLSA.

6.3.1 Correction des artéfacts topologiques

Pour évaluer la capacité de corrections des artefacts de FVT, on appliqua sa méthode de modélisation topologique à des cartes de segmentation A/V prédites par le modèle de segmentation sémantique Automorph [164]. Ce modèle a été spécifiquement choisi car, d'une part, il est public, de sorte que n'importe qui pourrait reproduire les expériences ci-dessous ; et d'autre part, ses cartes de segmentation A/V sont généralement de bonne qualité mais contiennent les artefacts topologiques typiques des CNN.

Nous avons comparé les arbres vasculaires extraits par le FVT avec ceux obtenus par la méthode de RMHAS [149], c'est-à-dire la méthode « naïve » qui analyse séparément les cartes de segmentation des artérioles et veinules, extrait de chacune le graphe **ART** et **VEI** individuellement, puis identifie le sous-graphe propre à chaque vaisseau par une analyse en composantes connectées et finalement place la racine de chaque vaisseau sur le nœud le plus proche du disque optique. Les deux méthodes reçoivent les mêmes segmentations A/V prédites par Automorph de sorte à pouvoir comparer les propriétés des arborescences extraites par chacune. La Figure 6.12 présente une vue comparative de ces arborescences pour trois images extraites du sous-ensemble réannoté de MAPLES-DR.

Analyse qualitative

Cette comparaison qualitative des deux méthodes (qui me rappelle étrangement le jeu des 7 différences) révèle plusieurs qualités de notre approche :
Premièrement, les principales erreurs de classification d'Automorph où une section de vaisseau change soudainement de classe sont corrigées par FVT et l'intégrité du vaisseau est correctement restaurée A/V (voir agrandissement **a.**, **b.** et **c.**).
Deuxièmement, une majorité des déconnexions de vaisseaux semble aussi être identifiée et corrigée correctement par FVT (voir agrandissement **a.** et **c.**). Concernant ces reconnexions, j'attire l'attention du lecteur sur l'agrandissement **a.** et ses nombreuses interruptions de veinules aux croisements avec l'artériole horizontale. Ces interruptions sont causées par une mauvaise

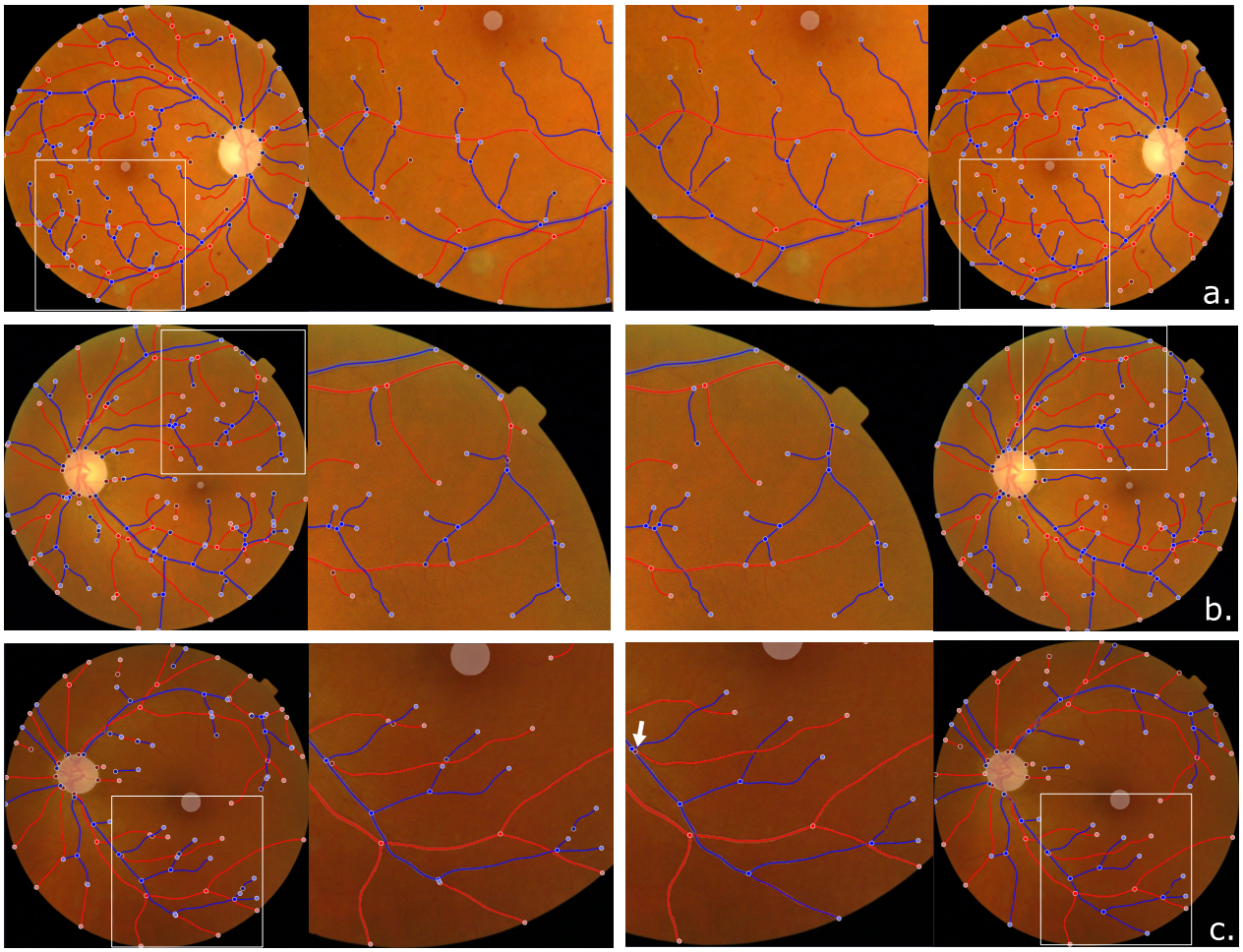


FIGURE 6.12 Arborescence extraite par RMHAS (gauche) et par FVT (droite).

Les noeuds racines apparaissent plus sombres, les noeuds feuilles plus clair. Toutes les images présentent des reconnexions ; **a.** et **b.** montrent des corrections de classification A/V ; **b.** présente le positionnement valide d'une branche réentrant l'image ; **c.** illustre l'effet de l'heuristique « garde-fou » sur une artère qui semble *émerger* d'une veine.

classification des pixels de croisements par Automorph : au lieu d'être classés **BOTH**, ils ont été identifiés uniquement **ART**, et ont ainsi été exclus de la carte de segmentation **VEI**. Après l'optimisation d'arborescence de FVT, les branches de veinules au-dessus de cette artériole ont toutes été reconnectées avec leur vaisseau mère. Dans ces trois images, je ne distingue qu'une déconnexion de veinules non identifiée par FVT, en haut au milieu de l'agrandissement **a.** Ici, les tangentes aux extrémités des nœuds terminaux ne se font pas suffisamment face pour que la reconnexion soit envisagée.

Troisièmement, FVT identifie correctement la direction des branches y compris pour la veine de l'agrandissement **b.** qui provient de l'extérieur de l'image et se dirige vers son centre. Le nœud racine (plus foncé) est bien placé à la frontière de la région d'intérêt et non proche du disque optique, en bas à gauche de l'agrandissement comme RMHAS.

Quatrièmement, l'agrandissement **c.** montre un exemple d'application de l'heuristique « garde-

fou ». La branche mère de l'artériole indiquée d'une flèche blanche n'a pas été segmentée par Automorph, ainsi l'artériole semble émerger de la bifurcation de la veinule. Puisque ce nœud n'est ni à proximité du disque optique, ni de la bordure de la région d'intérêt, il n'est pas considéré dans le graphe de ligne comme un potentiel nœud racine. Lors de l'optimisation de l'arborescence, l'artériole qui en émane est donc considérée comme appartenant au même vaisseau que la veinule. Cependant, puisque ces deux vaisseaux se recroisent plus en aval, l'heuristique « garde-fou » identifie ce cas peu probable et déconnecte l'artériole a posteriori, de sorte qu'elle apparaît bien détachée sur l'agrandissement **c**.

Analyse quantitative

Pour quantifier l'étendue de ces corrections, nous avons testé ces deux méthodes de modélisation topologique sur les 100 images de MAPLES-DR annotées A/V. L'extraction du graphe vasculaire est toujours réalisée à partir des segmentations sémantiques imparfaites d'AutoMorph, mais les annotations de MAPLES-DR nous permettent de mesurer combien de branches d'artérioles et de veinules sont annotées de la bonne classe. Les résultats quantitatifs confirment l'analyse qualitative : le partitionnement du graphe réalisé par FVT est plutôt robuste aux erreurs de classification : l'exactitude de classification A/V des branches passe ainsi de 93,7% pour RMHAS (sans correction d'artefacts) à 95,2% avec FVT. En particulier, le nombre de branches artériolaires correctement classées augmente de près de 3% (cf. Tableau 6.1). Ces résultats confirment également l'intuition de Chen et al. (2024 [26]) selon laquelle le post-traitement des prédictions des CNN à l'aide de méthodes basées sur les graphes peut améliorer la classification A/V. Je tiens ici à rappeler que FVT n'utilise pas d'informations extraites de l'image pour réaliser ces corrections : il analyse simplement la propriété géométrique et topologique extraite de la segmentation sémantique, et les rend conformes à une arborescence vasculaire plausible selon des connaissances cliniques a priori.

Tableau 6.1 Correction par FVT des artefacts de segmentation sémantique sur les images de MAPLES-DR.

Partionnement	Simple	Optimisé
Exactitude de classification A/V (en % du nombre de branches)		
Artérioles	90.29%	**93.11%**
Veinules	96.41%	**96.84%**
Nombre d'artefacts topologiques par image		
Racines de vaisseau mal placées	6.5	**2.2**
Discontinuités de vaisseaux	15.7	**3.7**

Pour aller plus loin, nous avons aussi compté manuellement le nombre de nœuds identifiés comme nœuds racines à tort ainsi que le nombre de discontinuités de vaisseaux dans les images de MAPLES-DR. L'attention portée au placement des nœuds racines peut apparaître saugrenue, mais elle est capitale pour déterminer la direction des branches du graphe vasculaire et ainsi identifier quelle est la branche principale de chaque bifurcation. Ici encore, FVT montre sa robustesse aux artefacts de segmentation et de classification d'AutoMorph, en réduisant le nombre de racines invalides par un facteur de 3, et le nombre de discontinuités de vaisseaux par un facteur de 4 (cf. Tableau 6.1).

6.3.2 Exemple d'application : Étude pilote sur CLSA

L'objectif final de FVT est l'extraction des paramètres vasculaires, dont un certain nombre a déjà été implémenté (tortuosité, et paramètres de bifurcations). Dans cette section, nous proposons une étude pilote pour montrer la validité de ces paramètres. Nous les comparons à ceux extraits par AutoMorph (AVR, tortuosité et dimension fractale). Pour que la comparaison soit juste, nous n'étudierons que le paramètre commun à ces deux logiciels : la tortuosité des artérioles et des veinules. Les deux logiciels s'appuient sur les mêmes cartes de segmentation sémantique prédite par AutoMorph pour réaliser leur mesure. Ainsi, les différences entre les valeurs mesurées sont à imputer aux corrections A/V et à la minutie de la modélisation géométrique de FVT (la tortuosité s'appuie sur la mesure de la courbure). Ils mesurent cette tortuosité par branche, puis les agrègent pour toute l'image par une moyenne pondérée par la taille de chaque branche.

Les images utilisées pour mener cette comparaison sont celles de l'Étude longitudinale canadienne sur le vieillissement (CLSA) [222], une étude à grande échelle qui suit les données de santé de 20 423 participants à travers le Canada. Parmi les informations de santé disponibles, nous nous sommes intéressés aux 272 patients ayant déclaré avoir eu des antécédents d'AVC. Cet antécédent a été associé à une évolution de la tortuosité vasculaire de la rétine par des

Tableau 6.2 Associations entre les tortuosités calculées par FVT ou Automorph, et les AVC dans la cohorte *Baseline* de CLSA. (Les p-values significatives sont indiquées en gras.)

Logiciel	Paramètre vasculaire	Odds Ratio	p-value
FVT	Tortuosité **ART**	3.6e-15	**0.031**
	Tortuosité **VEI**	6.1e+25	**0.006**
AutoMorph	Tortuosité **ART**	1.02	0.232
	Tortuosité **VEI**	1.08	**0.011**

études cliniques (Cheung et al., 2013 [223] ; Sandoval-Garcia et al., 2021 [224]). Pour reproduire ce résultat, nous avons appliqué une régression logistique aux mesures de tortuosité et ajusté nos modèles en fonction de l'âge, du diabète, du cholestérol et de l'hypertension des participants. Le tabagisme, l'IMC et l'origine culturelle n'ont pas démontré d'associations significatives et ont été exclus du modèle final.

Contrairement à Automorph, FVT a fourni une tortuosité artériolaire qui est significativement associée aux accidents vasculaires cérébraux. En ce qui concerne la tortuosité veineuse, le FVT et l'Automorph présentent tous deux des associations significatives avec les accidents vasculaires cérébraux, mais le FVT présente une corrélation plus forte avec un rapport de cotes de 6,1e+25. La correction des artefacts effectuée par le FVT sur l'arbre vasculaire semble donc améliorer la fiabilité et la qualité de l'évaluation de la tortuosité des vaisseaux.

En outre, Automorph – dont l'implémentation est en Python pur – requiert 715 ms par image pour calculer les biomarqueurs à partir des cartes de segmentation A/V précalculées, alors que FVT – dont l'extraction bas niveau de la topologie et de la géométrie est implémentée comme une extension C++ – n'a eu besoin que de 214 ms en comptant la correction des artefacts de segmentation ! Au total, l'analyse de 20 000 scans rétiniens de la cohorte du CLSA n'aura pris que 73 minutes, démontrant la capacité du FVT à traiter de grands ensembles de données.

6.4 Discussion et Contributions

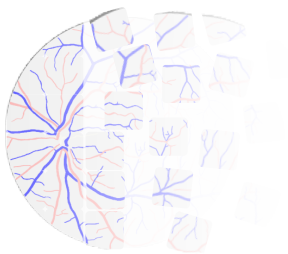
La méthode de modélisation géométrique et topologique implémentée dans FVT se déploie en quatre étapes, les trois premières préparant le terrain à la quatrième : d'abord l'extraction naïve de la topologie du squelette, puis la modélisation de la géométrie de chaque branche à partir de la segmentation, ensuite la prise en compte de la classification A/V et l'identification des artefacts topologiques (simplification et déconnexions), chaque étape enrichissant la représentation du graphe vasculaire tout en prenant en considération les potentiels artefacts présents dans la carte du squelette, de la segmentation ou de la classification. L'extraction topologique initiale à partir de la segmentation pure assure notamment que les erreurs de classification ne soient pas converties en erreurs de segmentation comme c'est le cas pour RMHAS. À partir de cette modélisation prudente, la dernière étape de la méthode recense toutes les hypothèses concernant les directions de branches, leurs connexions et reconnexions, et le positionnement des nœuds racines, et leur attribue un score de plausibilité. Et finalement, la partition du graphe est réalisée par une approche globale qui élimine les hypothèses les moins plausibles tout en assurant que celles restantes forment une arborescence cohérente.

Comparé aux autres logiciels de paramétrisation vasculaire complètement automatiques RM-

HAS [149] et AutoMorph [164], cette méthode confère à FVT une plus grande robustesse aux artefacts de segmentation et de classification régulièrement présents dans les cartes prédites par CNN. Par cet atout, la modélisation topologique de FVT peut servir autant à la correction a posteriori des segmentations sémantiques de ces modèles, qu'à la mesure de paramètres vasculaires fiables.

En outre, le projet de FVT a été pensé non pas comme une simple implémentation de partition du graphe à des fins de recherche théorique, mais comme un outil public et documenté mis à disposition de la communauté dans l'espoir qu'elle s'en empare. Ainsi, l'API de la librairie s'assure que l'utilisateur puisse non seulement accéder ergonomiquement à l'ensemble des composantes de la représentation riche du graphe vasculaire, mais puisse aussi manipuler ce graphe : ajouter, supprimer ou même scinder des branches, fusionner des nœuds, etc. La méthode décrite ici utilise ces fonctions pour simplifier le graphe, en corrigeant les artefacts ou en mesurant les paramètres vasculaires. Mais l'utilisateur peut choisir d'utiliser ces fonctions pour développer sa propre méthode, ou ses propres paramètres. Par cette librairie, j'espère humblement fournir un outil qui puisse, un tant soit peu, rapprocher la communauté de chercheurs en segmentation vasculaire de leurs acolytes cliniciens.

Le paquet FVT est aussi accompagné d'une librairie de visualisation permettant d'inspecter et de comparer dans le détail et directement depuis un notebook Jupyter ; toutes les représentations du graphe vasculaire énoncées dans ce chapitre (segmentations, squelettes, tangentes, graphes, etc.) ont d'ailleurs été générées avec son aide. Cet outil de visualisation peut paraître anecdotique, mais je suis pourtant convaincu que sa capacité d'inspecter aisément le moindre détail de la modélisation topologique ou géométrique, doublée du choix d'une implémentation en C++, a grandement conditionné la méthode que je propose, m'invitant à porter attention aux détails. La multitude de petites optimisations et les multiples règles et heuristiques qui jalonnent ce chapitre en sont autant de traces.



CHAPITRE 7 CHAPITRE COMPLÉMENTAIRE : LE PARADOXE DE LA LITTÉRATURE SUR LA SEGMENTATION DES VAISSEAUX RÉTINIENS

Dans ce chapitre complémentaire, je propose de prendre du recul sur la littérature de modélisation du graphe vasculaire pour étudier le paradoxe suivant. Ces dernières années, la segmentation sémantique des vaisseaux rétiniens reçoit une attention comme jamais auparavant : depuis 2020, le nombre de publications annuelles pour la segmentation seule dépasse les 150 chaque année (cf. Figure 7.1). C'est trois fois plus qu'en 2015, avant l'avènement des CNN. Pourtant, depuis 2020, cette vague de publications n'a pas suscité d'amélioration significative des performances. On va ici investiguer ce paradoxe par une méthode de revue systématique, automatique et expérimentale que je n'aurais pas pu utiliser au début de mon doctorat car elle exploite un LLM pour le triage des papiers.

7.1 Méthode de revue systématique et automatique

La revue systématique débute par un recensement exhaustif des papiers liés au sujet. Une première collecte de références est menée sur Scopus, Pubmed et Google Scholar. La requête est choisie volontairement large pour ne pas négliger d'articles. Par exemple, pour Scopus :

```
TITLE-ABS-KEY((fundus OR retinal) AND (vessels OR vascular OR arteries OR artery
OR arteriole OR arterioles OR vein OR veins OR veinule OR veinules)
AND (segmentation OR classification))
```

Cette recherche initiale aboutit à 8183 références uniques, principalement issues de Scopus.

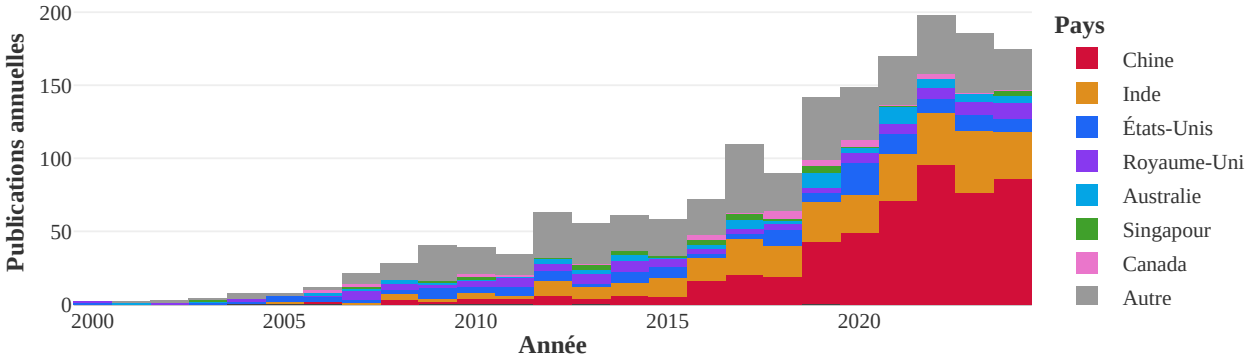


FIGURE 7.1 Nombre de publications par an et par pays traitant de la segmentation des vaisseaux rétiniens sur images de fond d'oeil.

Parmi elles, 705 furent exclues par impossibilité de télécharger le pdf automatiquement [1]. Dans les 7478 références restantes, une majorité est hors sujet. Elles furent donc filtrées par un LLM (le Phi-3-mini de Microsoft) à qui on confia le titre et l'abstract de chaque papier ainsi qu'une série de questions à choix multiples vérifiant les critères d'exclusion suivants :

- La langue de l'article est l'anglais (choix possibles : `english`, `chinese`, `german`, et `other`) ;
- L'article traite d'images de fond d'oeil couleur (choix : `color fundus`, `OCT or OCTA`, `angiography or angiograms`, `slit lamp`, `ultrasound`, `3D images`, `MR`, `CT`, et `other`) ;
- La contribution du papier concerne une nouvelle méthode de segmentation ou de classification d'artère veine (choix : `novel segmentation method`, `novel artery and vein classification method`, `review other research papers`, `novel diagnosis method`, et `other`) ;
- Les structures segmentées sont les vaisseaux (choix : `retinal vessels`, `arteries and veins`, `optic disc`, `macula or fovea`, `lesions`, et `microaneurysms or hemorrhages`).

Avec cette méthodologie, j'ai recensé 1738 publications concernant la segmentation des vaisseaux rétiniens et 126 concernant leur classification. Le regroupement par pays sur la Figure 7.1 révèle que l'accroissement spectaculaire du nombre de publications depuis 2020 est principalement dû à un intérêt grandissant en Chine et dans une moindre mesure en Inde. Les universités de ces deux pays totalisent près de deux tiers des publications de ces cinq dernières années.

Ces références furent agglomérées pour former un corpus dont le texte et les tableaux furent extraits du pdf des articles à l'aide de GrobID [225]. Puis, une série d'heuristiques analysant le titre des sections classa chacune selon leur fonction : `Introduction`, `Travaux Connexes`, `Méthode`, `Évaluation`, et `Conclusion`.. À la fin de cette procédure, on obtient donc un corpus soigneusement organisé contenant la plupart des articles publiés traitant de la segmentation des vaisseaux rétiniens sur image de fond d'œil et dont on peut sonder les titres et abstracts des articles, ainsi que le texte et les tableaux de leurs différentes sections. J'ai analysé ce corpus par des méthodes lexicographiques pour peindre un tableau général de l'évolution de la littérature sur la segmentation sémantique du graphe vasculaire rétinien.

7.2 Sonder la littérature de segmentation vasculaire rétinienne

Évolution des tendances méthodologiques Pour débuter l'analyse, étudions simplement les termes récurrents dans les titres des articles du corpus. Les nuages de mots présentés sur la Figure 7.2 regroupent ces termes selon quatre périodes : 2000-2009, 2010-2014,

1. Le téléchargement automatique des pdfs a été confié à un script python qui renseigne l'adresse DOI d'un article dans la barre d'un navigateur puis déclenche le téléchargement de sa référence et de son pdf à l'aide du plugin Zotero. Le script est disponible à l'adresse : github.com/gabriel-lepetitaimon/bibtool.

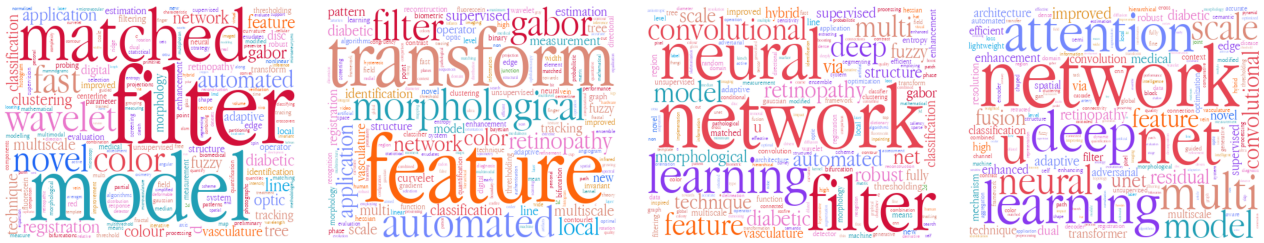


FIGURE 7.2 Nuages de mots des titres d'articles sur la segmentation des vaisseaux rétiniens.

Les mots communs à tous les papiers sont omis (`fundus`, `vessel`, `segmentation`, etc.).

2015-2019 et 2020-2025. On y voit apparaître dans les titres : les mots clés `convolutional neural network` et `deep learning` dans la période charnière 2015-2019. Auparavant, on voit la prédominance des `matched filter` et des ondelettes (`wavelet`) entre 2000 et 2009, puis des filtres de `Gabor`, des approches morphologiques (`morphological`) entre 2010 et 2015. Le concept de `feature` qui est présent dans les titres à cette période est généralement accolé aux termes `multiscale`, `invariant`, ou `landmark`. Après 2020, les réseaux de neurones perdent leur adjectif `convolutional`, au profit de `u net` et d'`attention`. Toutes ces observations corroborent le récit de l'évolution des méthodes décrites dans la revue de littérature.

Ces observations peuvent être approfondies en mesurant la proportion de papiers qui mentionnent, dans leur titre ou dans leur méthode, les termes associés aux différentes approches recensées à la section 2.2. La Figure 7.3 présente ces proportions année après année. Elle révèle que l'intérêt pour les méthodes non supervisées est en perte de vitesse. Les `matched filter`, les filtres de `Gabor` et les ondelettes ont résisté le plus longtemps, mais leur occurrence décroît monotonément depuis 2017. Même les opérateurs de morphomat pourtant utiles comme post-traitement et qui participaient à un tiers des méthodes proposées en 2014, ne sont mentionnés que par 7% des papiers en 2024. Seules les approches `multi-scale` connaissent un intérêt croissant puisque ce concept transcende les approches non supervisées et reste pertinent à la conception de réseaux de neurones. Concernant les classifieurs traditionnels, deux se démarquent sur la période 2005-2015 : les SVM qui sont étudiés dans 12% des papiers en 2011 ; et les réseaux de neurones qui, même en l'absence d'apprentissage profond et réduits à une ou deux couches cachées, étaient déjà les classifieurs les plus plébiscités de cette période. Cette Figure nous apprend surtout que, depuis son introduction dans le domaine en 2015, la proportion de papiers proposant des méthodes entraînées par apprentissage profond n'a cessé de croître. Ainsi, en 2024, 81% des articles mentionnaient dans leur méthode les termes de `Deep Learning`, `CNN` ou `U-Net` ! Ce dernier modèle est spécifiquement mentionné dans 71% des publications cette année-là, il est sans conteste le socle commun de l'état de l'art. Les recherches sur les approches génératives `GAN` ou `Transformer` sont en croissance mais restent

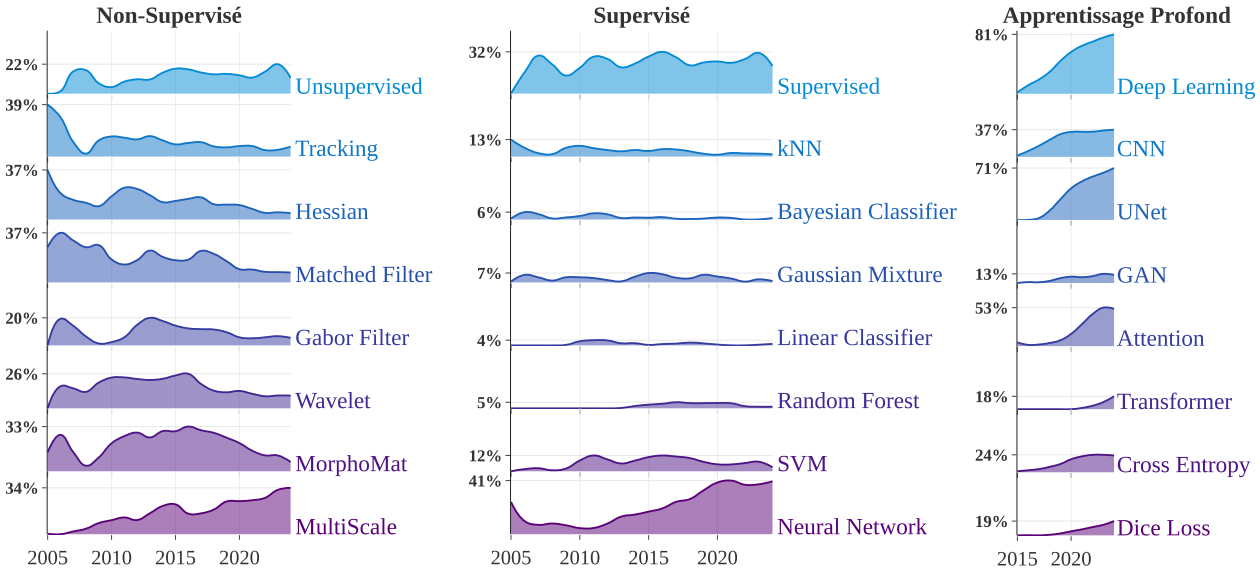


FIGURE 7.3 Occurence entre 2005 et 2024 des termes méthodologiques dans le titre et les sections de méthodes des articles de segmentation des vaisseaux rétiniens.

L'axe des ordonnées indique le pourcentage de papier contenant chaque terme parmi les publications de l'année. L'échelle de cet axe varie d'une colonne à l'autre.

minoritaires (respectivement 13% et 18% en 2024) ; contrairement au concept d'`attention` qui apparaît dans la moitié des papiers publiés en 2023 et 2024. Enfin, bien qu'il soit difficile de catégoriser les fonctions de coûts utilisées par une analyse lexicographique, on constate néanmoins que le nombre de mentions de la `dice loss` rattrape celui de la `cross-entropy` et pourrait l'égaler en 2025 si la tendance se maintient.

Évolution des méthodes de validation La méthode d'analyse par occurrence est parfaitement adaptée pour étudier l'évolution des pratiques de validation des algorithmes : il suffit de mesurer l'occurrence des noms des métriques et des jeux de données dans les sections de validation et dans les tableaux des papiers. Ces mesures sont présentées sur la Figure 7.4. Concernant les métriques, elles mettent en exergue la prédominance de l'exactitude (utilisée par 83% des papiers en 2024) et du couple sensibilité/spécificité. La Figure montre aussi une nette progression du score Dice à partir de 2015, de sorte qu'il est mentionné pour valider 35% des articles en 2024. Le score de détection de vaisseau CL Dice proposé par Shit et al. en 2021 [136] peine à s'implanter : en 2024, seulement 4% des papiers l'utilisent. Pour les métriques sensibles à la topologie vasculaire, c'est encore pire : les termes `topolog*` n'apparaissent que dans 9% de la littérature en 2014 et les métriques topologiques telles que le nombre de Betti ne sont tout simplement pas mentionnées. Concernant les jeux d'entraînement et de validation, DRIVE [31] et STARE [32] sont les plus utilisés (par respectivement 70% et 50% des articles publiés en 2024) et sont rattrapés par CHASE DB [33] (47% en 2024)

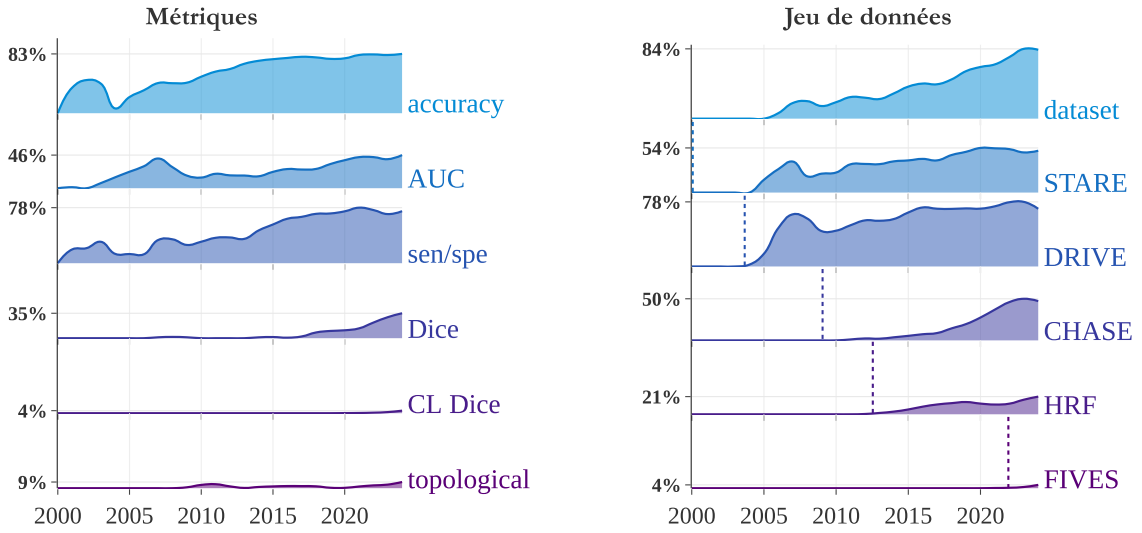


FIGURE 7.4 Occurence des métriques et jeu de données annotées pour la segmentation des vaisseaux rétiniens dans les sections de validation et les tableaux des articles du corpus.

Les traits verticaux pointillés sur le graphe des jeux de données indiquent leurs dates de publication.

bien que sa publication ait été plus tardive. L'adoption des bases de données haute résolution HRF [226] et FIVES [36] semble par contre ralentie. Certes, elles ont été publiées plus récemment (en 2013 et 2021), mais je reste étonné qu'une base aussi complète que FIVES tant en termes de résolution que de nombre d'échantillons (800 images !) ne soit toujours exploitée que par 4% des publications de 2024. Plus généralement, on constate un certain conservatisme dans les protocoles de validation : suite à l'adoption rapide de DRIVE, les autres jeux de données peinent à y trouver leur place. Une dizaine d'années se seront ainsi écoulées entre la publication de CHASE DB et sa démocratisation dans la littérature.

7.3 Pistes d'explications du paradoxe

Revenons au paradoxe qui nous préoccupe : l'envolée du nombre de publications qui contraste avec la stagnation des performances. La vague de publications ces dernières années traitant de la segmentation des vaisseaux rétiniens est quasiment exclusivement dédiée à l'entraînement par apprentissage profond de modèles convolutifs dérivés de la famille des U-Net. Cette application clinique est devenue un laboratoire d'où les équipes de recherche redoublent de propositions pour construire des réseaux toujours plus performants et toujours plus complexes. Cette exploration des différentes structures de modèles est catalysée par les outils de développement à la disposition des chercheurs (PyTorch, Keras, etc.) qui rendent aisées les itérations architecturales à partir de modèles existants.

Mais pour évaluer la pertinence de ces modèles, les protocoles de validation actuels sont insuffisants. D'une part, les jeux de données publics pour l'entraînement et la validation

des algorithmes sont trop limités : la majorité des papiers publiés n'exploitent que les deux ou trois jeux de données les plus anciens, dont la résolution n'est pas représentative des technologies d'imagerie actuelles et qui totalisent seulement 88 images à diviser en ensembles d'entraînement, de validation et de test. Pour Li et al. (2025 [25]), c'est ce point qui entrave la progression des modèles et réduit surtout leur valeur d'usage dans des scénarios cliniques. La sous-exploitation des jeux de données récents est d'autant plus problématique qu'elle aggrave le risque de surapprentissage déjà entretenu par la complexité croissante des modèles.

D'autre part, les métriques de validation sont en retard par rapport aux performances élevées des modèles. Si l'exactitude de segmentation était pertinente pour mesurer la progression des algorithmes il y a 15 ans, elle montre aujourd'hui ses limites. En effet, les modèles actuels dépassent systématiquement de 2 à 3 % l'exactitude obtenue par le second annotateur de DRIVE (94.73%) en comparaison avec la vérité terrain annotée par le premier. J'ai régulièrement vu ce résultat interprété dans la littérature comme une attestation de la supériorité de la machine sur un annotateur humain. C'est, il me semble, un contresens. Les segmentations fournies comme vérité terrain par DRIVE ne sont pas plus valides que celles du second observateur : l'une comme l'autre ont été annotées par des humains et souffrent de biais et de variances d'annotations. Ainsi, une exactitude nettement supérieure à celle du second annotateur témoigne d'une sur-spécialisation du modèle sur les biais d'annotation du premier. En ce sens, une telle valeur devrait plutôt être interprétée comme étant « en dehors de la zone de validité de la métrique », à la manière des instruments de mesure indiquant « Out of Range » lorsque leurs aiguilles sortent du cadran. Pour lire une valeur fiable des performances des modèles de segmentations, il faut donc changer d'instrument de mesure et choisir une métrique qui relève plus finement les erreurs qui nous importent et ignore les variations inter-observateurs : comme le CL Dice ou les métriques topologiques.

Pourtant depuis 2015, les métriques qui ont principalement progressée et rivalisent avec l'exactitude sont celles qui souffrent des mêmes lacunes de mesure : l'AUC-ROC et la sensibilité/spécificité. Je vois deux facteurs pouvant expliquer ce phénomène. **1.** La comparaison à l'état de l'art nécessite la réutilisation des métriques déjà publiées pour évaluer les performances d'un modèle, ce qui crée une inertie dans les pratiques de validation. La progression de l'AUC-ROC et du couple sensibilité/spécificité est donc le prolongement de leur implantation déjà importante avant 2015. **2.** L'*ubiquité* des réseaux de neurones, c'est-à-dire la capacité de ce formalisme à apporter la même solution algorithmique à une vaste diversité de problèmes de traitement du signal, est certes au cœur de leur développement et de leur succès rapide, mais elle tend à uniformiser aussi les protocoles de validation. Or ceux-ci devraient rester spécifiques à chaque application. Ce constat avait d'ailleurs poussé Maier-Hein et al. (2024 [227]) à mener un vaste projet de recommandations de métriques pour les différentes applications

biomédicales. En l'occurrence, pour la segmentation des vaisseaux, ils préconisent le CL Dice.

Ainsi, si les performances de segmentation semblent stagner, c'est que leurs protocoles de validation ne sont plus en mesure de supporter les problématiques de recherche contemporaines qui font essentiellement face à des défis topologiques. Dit autrement, le problème de la segmentation des vaisseaux rétiniens – tel qu'il est formulé selon l'exactitude et sur les quelques jeux de données publics – peut être considéré comme résolu. L'accumulation des recherches sur le sujet n'induit donc pas d'amélioration des performances tant qu'un nouveau protocole de validation plus spécifique n'aura pas été développé pour les évaluer.

Ce protocole devra veiller à aligner les métriques de performances avec l'objectif sous-jacent à la segmentation et la classification vasculaire. Cet objectif est resté inchangé depuis la première publication de Chaudhuri et al. en 1989 [42] : permettre une meilleure compréhension et un meilleur dépistage des pathologies vasculaires et rétiniennes. D'ailleurs, quatre articles sur cinq de notre corpus utilisent dans leur introduction les termes `disease` et `diagnostic`, et ce, quelle que soit leur année de publication. Plus spécifiquement, la rétinopathie diabétique est la plus invoquée (dans 6/10 papiers en moyenne), suivie de l'hypertension (4/10). Dans leur revue de 2022, Khandouzi et al. [22] expliquent aussi l'intérêt croissant pour la segmentation des vaisseaux rétiniens par : "*the accuracy of segmentation algorithms has a significant effect on the early detection of ocular diseases and can make a positive difference as a result of treatment*". La motivation sous-jacente à toute cette littérature fait donc consensus.

Pourtant, très peu de papiers s'intéressent à la transposition des modèles de segmentation et de classification vers des résultats cliniques. Comme si la conception et l'amélioration de ces modèles accaparaient l'attention de la communauté de chercheurs depuis tant d'années que leur motivation initiale s'est finalement muée en une image d'Épinal : admise par tous mais rarement interrogée et étudiée en pratique. Probablement parce que, contrairement à celui de la segmentation vasculaire, le problème de sa transposition en résultats cliniques ne se pose pas en des termes bien définis.

Au fond, je réalise que le travail que j'ai rapporté dans cette thèse pourrait s'apparenter à celui d'un pionnier : allant récolter de nouvelles annotations, alertant sur la variabilité que j'y ai découverte ; puis explorant une voie d'architecture de réseaux rarement étudiée, cherchant à améliorer les modèles par leur simplification ; enfin, défrichant le champ de la modélisation topologique à partir des segmentations imparfaites, et proposant des outils pour que d'autres puissent y faire germer leurs propres idées. Bref, j'ai consacré mon doctorat à étudier et définir les termes permettant de bien poser le problème de la modélisation automatique de la topologie de l'arbre vasculaire rétinien, pour qu'un jour, l'évaluation des qualités topologiques et paramétriques des modèles de segmentation soit aussi simple qu'un calcul d'exactitude.

CHAPITRE 8 DISCUSSION GÉNÉRALE

Le projet doctoral présenté dans cette thèse a pour objectif général d'exploiter les modèles CNN de segmentation et de classification des vaisseaux sur image de fond d'œil pour extraire une modélisation topologique robuste de l'arbre vasculaire rétinien et mesurer des paramètres vasculaires fiables. Pour ce faire, j'ai réalisé trois objectifs spécifiques de recherche : **1.** la construction d'un nouveau jeu de données annotées pour enrichir l'entraînement de ces modèles ; **2.** la reparamétrisation des neurones convolutifs pour piloter l'orientation des filtres de convolution et ainsi mutualiser les paramètres et améliorer la généralisation des modèles de classification vasculaire ; **3.** la modernisation des approches de partition du graphe vasculaire rétinien en proposant une méthode de modélisation topologique de l'arbre vasculaire rétinien robuste aux artefacts des modèles CNN. Ce pénultième chapitre discute des réussites et limitations des méthodes proposées à l'aune des objectifs spécifiques et de la problématique générale de ma thèse.

8.1 MAPLES-DR

La campagne d'annotation de MAPLES-DR a permis la collecte de nouvelles segmentations de vaisseaux annotées par des rétinologues. Ces nouvelles vérités terrain rejoignent celles des autres bases de données publiques comme STARE [32], DRIVE [31], CHASEDB [33], HRF [226], doublant le nombre d'annotations disponibles : MAPLES-DR seul contient en effet plus d'images que toutes ces bases réunies. Par cette contribution, nous enrichissons donc significativement les ensembles d'entraînement et de validation des modèles de segmentation vasculaire de sorte qu'ils représentent mieux la diversité des situations cliniques. Nous avons d'ailleurs montré à travers une étude menée sur les segmentations des structures pathologiques que les modèles entraînés sur MAPLES-DR possèdent de bonnes capacités de généralisation et que ce jeu de données constitue une base solide de pré-entraînement [202].

L'utilisation de carte de pré-segmentation a joué un rôle clé pour l'annotation des cartes vasculaires en réduisant considérablement le temps de segmentation. Alors que l'annotation manuelle de DRIVE a requis 2 heures par image [31], notre équipe de rétinologue n'a eu besoin que de 6 minutes en moyenne pour ajouter les vaisseaux omis par le modèle de pré-segmentation. Certes, ces cartes présentent un risque de biais sur les annotations finales. Mais nous avons montré que ce biais se manifestait plutôt par une sensibilité accrue des annotateurs (en tout cas pour les structures pathologiques). Nous avons surtout montré que ces cartes étaient un moyen efficace pour réduire la variabilité inter-observateur, et cet effet

est d'autant plus prononcé que la qualité du modèle de pré-segmentation est bonne. C'est le premier reproche qu'on pourrait faire à MAPLES-DR : ses modèles de pré-segmentations correspondent à l'état de l'art de 2018 et sont aujourd'hui dépassés.

En outre, et bien qu'elle ne soit pas directement liée à mon projet doctoral, la variabilité inter-observateur mesurée sur les structures pathologiques de la rétine interroge. Des études mesurant la variabilité d'annotations des vaisseaux existent : pour DRIVE [31] les deux annotateurs sont d'accord sur 94,7% des pixels, plus récemment Kai et al. [36] avaient mesuré un dice inter-annotateur de 0,924. Mais, à notre connaissance, aucune étude n'avait évalué la variabilité inter-observateur pour l'annotation des lésions rétiniennes. Les mesures de variabilité présentées dans cette thèse sont donc inédites. Elles sont aussi surprenantes par leur ampleur (entre 13,3% et 61% d'accord de détection en fonction du type de lésions). Pour mieux interpréter ces valeurs, il faudrait cependant conduire de nouvelles annotations pour distinguer dans cette variabilité, la composante inter-observateur (biais) et celle intra-observateur (bruit). Cette dernière n'est en effet pas mesurée par notre étude sur les lésions et pourrait expliquer en partie la variabilité observée.

8.2 *Steered* CNN

Le formalisme des *steered* CNN proposé dans le chapitre 5 réussit l'exploit, à rebours de la littérature des méthodes de segmentation sémantique des vaisseaux rétiniens, d'améliorer ses performances suite à une réduction de son nombre de paramètres. De ce point de vue, les *steered* CNN marchent dans les pas de Gladran et al. [128] qui affirmaient obtenir les performances de l'état de l'art en 2022 avec un modèle simplifié.

Pour réaliser cet objectif, on s'est inspiré des travaux qui proposent de redéfinir les neurones convolutifs pour les contraindre à respecter des groupes de symétries, et plus particulièrement l'équivariance par rotation [211–213]. Ces travaux sont généralement motivés par l'intuition que ces symétries épargneraient au réseau la nécessité de consacrer plusieurs paramètres à la détection d'une même forme sous toutes ses orientations. À notre connaissance, cette intuition qui semble de bon sens n'avait cependant jamais été étudiée expérimentalement. L'étude du tropisme des neurones convolutifs dans un CNN que j'ai conduite à la section 5.2 indique d'ailleurs qu'elle est partiellement fausse. J'ai en effet montré que, sous l'effet de l'augmentation de données par rotation, un CNN standard ne modélise pas l'équivariance à la rotation par une duplication de ses filtres convolutifs, mais par l'apprentissage de filtres isotropiques. Ainsi, la force des *steered* CNN comparés aux CNN standards ne se limite pas à leur efficacité paramétrique accrue : elle réside aussi dans leur capacité à apprendre des filtres anisotropiques. On observe en effet que, lorsque la régularisation par augmentation de

données de rotation est levée, les filtres vers lesquels convergent naturellement les premières couches d'un *steered* CNN sont particulièrement anisotropiques : le poids de leurs composantes symétriques et anti-symétriques est nettement plus élevé que celui de leurs composantes isotropiques. Cette observation concorde avec la littérature de classification traditionnelle des vaisseaux rétiniens qui analyse le profil des vaisseaux [78, 82].

La réduction du nombre de paramètres n'est cependant pas un but en soi. Le second objectif reposait en réalité sur le pari que la réduction du nombre de paramètres se traduirait par une hausse de la capacité de généralisation du modèle (selon l'intuition du rasoir d'Ockham). Sur ce point, les conclusions sont plus mitigées. Certes lorsqu'on limite un *steered* CNN et un CNN standard à 10k paramètres ou moins, ou lorsqu'on les entraîne sur 20 images ou moins, les performances mesurées sur un jeu de données n'ayant pas servi à l'entraînement sont nettement en faveur des *steered* CNN. Mais cet écart se résorbe lorsqu'on augmente le nombre de paramètres et d'échantillons d'entraînement. Ce résultat reste honorable dans la mesure où la plupart des méthodes de la littérature sont entraînées avec 20 images ou moins, faute de données annotées.

Les limitations des *steered* CNN sont liées aux champs d'orientations $\vec{\boldsymbol{\alpha}}$ qui pilotent l'orientation de leurs filtres pour chaque pixel de l'image. Dans notre méthode, ce champ est une variable d'entrée calculée préalablement, ce qui a deux conséquences : **1.** Les *steered* CNN ne peuvent être appliqués à la segmentation vasculaire puisque le champ d'orientations des vaisseaux serait alors inconnu. Ils pourraient néanmoins être utilisés pour raffiner une segmentation préexistante ou pour délimiter les contours des vaisseaux à partir de leur squelette. **2.** La classification prédite par les *steered* CNN peut être sensible aux erreurs de la carte de segmentation vasculaire dont est dérivé le champ d'orientations $\vec{\boldsymbol{\alpha}}$. Cet effet est partiellement atténué par l'augmentation de données élastiques appliquées à l'entraînement des *steered* CNN et qui induit des petites rotations aléatoires du champ $\vec{\boldsymbol{\alpha}}$. Néanmoins, la sensibilité aux erreurs de segmentation mériterait d'être quantifiée.

8.3 Fundus Vessels Toolkit

Parmi les logiciels de paramétrisation automatique du réseau vasculaire rétinien reposant sur des modèles CNN pour la segmentation sémantique des vaisseaux, un seul (RMHAS [149]) s'essaie à la modélisation topologique du graphe vasculaire. Mais la méthode choisie n'est pas robuste aux artefacts topologiques des CNN : à la moindre erreur de classification au sein d'une branche, cela produit une déconnexion dans le graphe.

Pour pallier ces erreurs, je reformule le problème de partitionnement du graphe vasculaire

comme celui d'une énumération d'hypothèses mutuellement inclusives ou exclusives, puis d'une sélection de l'ensemble cohérent le plus plausible d'hypothèses par double application de la MSA. Ce formalisme présente plusieurs intérêts par rapport aux méthodes qu'avaient proposées la littérature pour assister la classification A/V avant les CNN. Comparé aux méthodes d'interprétation de la connectivité locale des branches [88–91], il assure la cohérence générale de l'arborescence et permet de résoudre la connectivité des nœuds incertains en s'appuyant sur ceux plus évidents. Comparé aux méthodes d'optimisation globale de l'arborescence [93–95] elle permet d'évaluer les hypothèses de reconnexions des branches et ne postule pas de la position des nœuds racines. Elle est ainsi moins sensible aux artefacts de segmentation et de classification vasculaires, et peut même permettre de corriger des erreurs de classification A/V !

Tous ces artefacts ne sont néanmoins pas corrigés pour autant. Lorsqu'une branche entière est omise par le modèle de segmentation, FVT n'a aucun moyen pour la reconstituer. De même si l'intervalle séparant deux portions d'une branche mal segmentée est trop grand ou si les extrémités des branches ne se font pas face. Dans de rares cas, il arrive que la procédure d'optimisation d'arborescence désigne la mauvaise parenté pour une branche du graphe, de sorte que la branche et toute sa descendance soient affiliées au mauvais vaisseau. Ces cas se produisent généralement suite à un faux négatif de la segmentation effaçant complètement une branche incidente à un croisement de vaisseaux. La branche émergeant du croisement donne alors l'illusion de bifurquer de l'autre vaisseau. La solution optimale serait de considérer le croisement comme la racine de la branche émergeante, mais les nœuds éloignés du disque optique et du bord de l'image ne sont pas considérés comme des racines potentielles. Ce type d'erreurs est malheureusement commun à toutes les approches de partitionnement du graphe vasculaire. L'heuristique « garde-fou » proposée à la fin de la méthode du chapitre 6 détecte et corrige a posteriori une partie de ces cas, mais n'est pas infaillible.

Il reste que la représentation géométrique et topologique de l'arbre vasculaire rétinien par FVT produit des résultats encourageants sur l'étude pilote conduite sur les 20k images de CLSA, reproduisant les résultats d'études cliniques [223, 224] par une extraction automatique de la tortuosité à partir de cartes de segmentation sémantique imparfaites.

CHAPITRE 9 CONCLUSION

9.1 Synthèse des travaux

Par la publication de MAPLES-DR [202], j'ai contribué à l'effort impulsé par la communauté, de la collecte de nouvelles données annotées pour diversifier les ensembles d'entraînement et de validation des modèles de segmentation des structures rétiniennes. L'ambition de pluralité des labels et les contraintes liées aux distances géographiques séparant les annotateurs nous ont poussés à proposer un nouveau protocole d'annotation et à développer notre propre plateforme d'annotations. Ce diptyque s'est révélé particulièrement efficace autant pour réduire le temps d'annotation que pour réduire la variabilité d'annotation inter-observateur. Par la mise à disposition des outils d'annotations, et par la documentation du protocole ayant abouti à MAPLES-DR dans une publication Scientific Data, j'espère permettre à d'autres équipes de recherche de contribuer à l'effort d'annotation. La campagne de collecte de MAPLES-DR fut aussi l'occasion de conduire une étude de variabilité inédite sur l'annotation des structures pathologiques rétiniennes, mettant en lumière de nombreux désaccords sur l'annotation de ces structures. Cette découverte soulève, plus généralement, une problématique d'alignement entre l'objectif clinique nécessitant l'identification de ces structures et les métriques de segmentation.

J'ai ensuite proposé une reparamétrisation des neurones convolutifs qui découple la détection de l'orientation des vaisseaux de leur classification en artérioles ou veinules. Le formalisme des *steered* CNN [228] ajoute une seconde entrée aux modèles de classification sous la forme d'un champ de direction α qui pilote, pour chaque pixel de l'image, la direction dans laquelle doivent s'orienter les filtres du modèle avant d'être convolués. Ainsi, les fonctions des filtres convolutifs sont mutualisées de sorte que chacun puisse participer à l'analyse de tous les vaisseaux d'une image, quelle que soit leur orientation. Pour réaliser efficacement et précisément la rotation des filtres convolutifs, ils sont reparamétrés en *steerable filters*. Outre les propriétés d'équivariance par rotation, ce reparamétrage en harmonique polaire permet une interprétation plus intuitive des paramètres appris par le modèle. Dans un article publié sur ce formalisme, j'ai montré que les modèles de classification A/V qui l'utilisaient bénéficiaient d'une meilleure capacité de généralisation – en particulier lorsque le nombre de paramètres ou d'échantillons d'entraînement est limité – et que leur entraînement nécessitait moins d'itérations.

Enfin j'ai implémenté un outil de modélisation automatique de la géométrie et de la topologie du graphe vasculaire à partir de segmentations sémantiques imparfaites : le *Fundus Vessels*

Toolkit [229]. Cet outil s'appuie d'une part sur une séquence de traitement raffinant, étape après étape, une représentation topologique et géométrique du graphe vasculaire ; et d'autre part sur une modernisation des méthodes de partitions du graphe vasculaire afin de les adapter aux défis que représentent les artefacts présents dans les cartes de segmentation sémantique produites par CNN. L'alliance des deux permet de rétablir les connexions entre des branches déconnectées dans la segmentation à cause de faux négatifs. Elle assure aussi la correction des erreurs locales de classification, en forçant le graphe vasculaire extrait de la segmentation sémantique à adopter la forme d'une arborescence plausible. Sous cette représentation, le réseau vasculaire rétinien se prête volontiers à une paramétrisation fiable de ses caractéristiques géométriques et topologiques.

9.2 Recommandation et travaux futurs

Recommandation 1 : **Étude de variabilité inter et intra-observateur sur l'annotation des vaisseaux assistée par pré-segmentation** L'étude de la variabilité sur les structures pathologiques de la rétine du chapitre 4 interroge sur l'usage de ces vérités terrain pour entraîner et valider des modèles et révèle combien l'annotation par présegmentation réduit cette variabilité. À ma connaissance, ce facteur n'a pas été étudié dans les études de variabilités d'annotations des vaisseaux. De façon plus générale, mon expérience avec MAPLES-DR m'encourage à recommander de toujours conduire une étude de variabilité inter et intra-observateur lors de l'annotation d'un jeu de données, ne serait-ce que pour définir les valeurs "maximales" des métriques au-delà desquelles elles ne mesurent que du bruit.

Recommandation 2 : **Construire une vérité terrain topologique pour la vasculature rétinienne.** Le manque d'images de fond d'œil annotées pour la segmentation et la classification des vaisseaux se résorbe peu à peu. C'est cependant loin d'être le cas pour la tâche de partitionnement du graphe vasculaire et plus généralement d'extraction de la topologie de l'arbre vasculaire rétinien. À ma connaissance, il n'existe qu'un seul dataset à proposer ce type d'annotation : RETA [40]. Les outils proposés dans cette thèse : autant la plateforme d'annotations que la librairie FVT, pourraient assister la collecte de plus de vérités terrains de ce type, en pré-annotant les images.

Recommandation 3 : **Orienter les *steered* CNN par un mécanisme d'attention** Le formalisme des steered CNN requiert aujourd'hui le calcul préalable d'un champ de direction pour définir comment orienter ses filtres. Ce qui ne permet pas son application pour une tâche de segmentation. On peut cependant imaginer prédire ce champ de vecteurs sous

la forme d'une matrice à deux composantes calculée par une branche parallèle du modèle. Cette branche parallèle serait en réalité analogue à un mécanisme d'attention dont on aurait remplacé la fonction sigmoïde par une sinusoïde et qui aurait la charge de déphaser les steerable filters plutôt que de shunter telles caractéristiques ou tels pixels. Conserver la propriété d'équivariance par rotation avec ce formalisme nécessiterait cependant quelques recherches théoriques pour s'assurer que la direction prédite par la branche suive toujours l'orientation des objets dans l'image.Une investigation plus avancée de la théorie des steerable CNN pourrait donner plus d'idées pour aborder ce problème.

Recommandation 4 : **Modéliser le score de plausibilité du graphe plutôt que de l'estimer par heuristiques** Dans FVT, la qualité du score de plausibilité d'une arborescence est totalement dépendante de la validité des heuristiques qui le calculent. Ces heuristiques doivent donc être réglées avec soin pour tenir compte de la multitude des topologies vasculaires rétiniennes. Pour contourner ce problème, on peut tout à fait imaginer entraîner un GNN à prédire la plausibilité de chaque arc du graphe de ligne en lui donnant accès aux caractéristiques géométriques (angle, calibre, courbure) de chaque branche, accompagnées de caractéristiques géométriques (par exemple extraites de la représentation latente du modèle de classification A/V). La partition du graphe serait toujours réalisée par la double application de la MSA, mais sur les scores de plausibilité prédits par le GNN. C'est d'ailleurs toute l'élégance de la méthode proposée pour le partitionnement du graphe : elle est décorrélée de l'algorithme d'énumération des hypothèses et de celui évaluant leur plausibilité. On peut ainsi faire évoluer l'un ou l'autre en gardant le même cadre général. On peut aussi envisager l'adaptation de méthodes de suivi vasculaire par renforcement développées pour la modélisation de la micro-vasculature cérébrale [230].

Recommandation 5 : **Développer des métriques de similarités topologiques** La représentation géométriquement et topologiquement riche produite par FVT peut servir de base au développement de nouvelles métriques topologiques pour évaluer les modèles de segmentation et de classification. À ce rôle, je pense que la distance d'édition de graphe serait la meilleure candidate. Certes, il faudrait étendre les algorithmes qui la calculent efficacement pour qu'ils gèrent les opérations de fusion/scission de branches et de nœuds. Ces opérations sont en effet essentielles pour identifier les déconnexions/reconnexions de branches. Mais une fois ce problème algorithmique résolu, une telle métrique serait non seulement utile pour évaluer les méthodes de segmentation vasculaires, mais aussi celles de modélisation topologique. Et elle paverait la voie vers un recalage avancé des graphes vasculaires rétiniens permettant par exemple l'analyse de leur évolution dans le temps à travers des études longitudinales.

RÉFÉRENCES

[1] C. P. Sisson, S. Farnand, M. Fairchild et B. Fischer, "Analysis of Color Consistency in Retinal Fundus Photography : Application of Color Management and Development of an Eye Model Standard," *Analytical Cellular Pathology*, vol. 2014, n°. 1, p. 398462, 2014, _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1155/2014/398462. [En ligne]. Disponible : https://onlinelibrary.wiley.com/doi/abs/10.1155/2014/398462

[2] J. Cao, T. Felfeli, R. Merritt et M. H. Brent, "Sociodemographics associated with risk of diabetic retinopathy detected by tele-ophthalmology : 5-year results of the Toronto tele-retinal screening program," *Canadian Journal of Diabetes*, vol. 46, n°. 1, p. 26–31, févr. 2022, publisher : Elsevier BV.

[3] H. Friedenwald, "Pathological changes in Retinal Blood Vessels in Arteriosclerosis and Hypertension (Doyne lecture)," *Transactions of the Ophthalmological Societies of the United Kingdom*, vol. 50, p. 452–530, 1930.

[4] M. D. Abràmoff, M. K. Garvin et M. Sonka, "Retinal imaging and image analysis," *IEEE Reviews in Biomedical Engineering*, vol. 3, p. 169–208, 2010, publisher : IEEE.

[5] M. Gunn, "On Ophthalmoscopic Evidence of General Arterial Disease." *Transactions of the Ophthalmological Societies of the United Kingdom*, n°. 18, p. 356–381, 1898.

[6] J. B. Walsh, "Hypertensive Retinopathy : Description, Classification, and Prognosis," *Ophthalmology*, vol. 89, n°. 10, p. 1127–1131, oct. 1982. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0161642082346643

[7] N. M. Keith, H. P. Wagener et N. W. Barker, "Some different types of essential hypertension," *The American Journal of the Medical Sciences*, vol. 197, n°. 3, p. 332–343, 1939, publisher : Elsevier BV.

[8] A. Kagan, E. Aureli et J. Dobree, "A note on signs in the fundus oculi and arterial hypertension : conventional assessment and significance." *Bulletin of the World Health Organization*, vol. 34, n°. 6, p. 955–960, 1966. [En ligne]. Disponible : https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2476046/

[9] J. Parr et G. Spears, "General caliber of the retinal arteries expressed as the equivalent width of the central retinal artery," *American journal of ophthalmology*, vol. 77, n°. 4, p. 472–477, 1974, publisher : Elsevier.

[10] J. C. Parr, "Hypertensive generalised narrowing of retinal arteries," *Transactions of the Ophthalmological Society of New Zealand*, vol. 26, n°. 0, p. 55–60, janv. 1974.

[11] M. D. Knudtson, K. E. Lee, L. D. Hubbard, T. Y. Wong, R. Klein et B. E. Klein, "Revised formulas for summarizing retinal vessel diameters," *Current Eye Research*, vol. 27, n°. 3, p. 143–149, janv. 2003, publisher : Taylor & Francis _eprint : https ://doi.org/10.1076/ceyr.27.3.143.16049. [En ligne]. Disponible : https://doi.org/10.1076/ceyr.27.3.143.16049

[12] L. D. Hubbard, R. J. Brothers, W. N. King, L. X. Clegg, R. Klein, L. S. Cooper, A. R. Sharrett, M. D. Davis et J. Cai, "Methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the atherosclerosis risk in communities study1," *Ophthalmology*, vol. 106, n°. 12, p. 2269–2280, déc. 1999. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0161642099905250

[13] L. M. Sherry, J. Wang, E. Rochtchina, T. Wong, R. Klein, L. Hubbard et P. Mitchell, "Reliability of computer-assisted retinal vessel measurement in a population," *Clinical & Experimental Ophthalmology*, vol. 30, n°. 3, p. 179–182, 2002, _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1442-9071.2002.00520.x. [En ligne]. Disponible : https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1442-9071.2002.00520.x

[14] R. Klein, M. Knudtson, B. Klein, B. Zinman, R. Gardiner, S. Suissa, A. Sinaiko, S. Donnelly, P. Goodyer, T. Strand et others, "The relationship of retinal vessel diameter to changes in diabetic nephropathy structural variables in patients with type 1 diabetes," *Diabetologia*, vol. 53, n°. 8, p. 1638–1646, 2010, publisher : Springer.

[15] R. Klein, C. E. Myers, M. D. Knudtson, K. E. Lee, R. Gangnon, T. Y. Wong et B. E. K. Klein, "Relationship of Blood Pressure and Other Factors to Serial Retinal Arteriolar Diameter Measurements Over Time : The Beaver Dam Eye Study," *Archives of Ophthalmology*, vol. 130, n°. 8, p. 1019–1027, août 2012. [En ligne]. Disponible : https://doi.org/10.1001/archophthalmol.2012.560

[16] T. Y. Wong, F. M. A. Islam, R. Klein, B. E. K. Klein, M. F. Cotch, C. Castro, A. R. Sharrett et E. Shahar, "Retinal Vascular Caliber, Cardiovascular Risk Factors, and Inflammation : The Multi-Ethnic Study of Atherosclerosis (MESA)," *Investigative Ophthalmology & Visual Science*, vol. 47, n°. 6, p. 2341–2350, juin 2006. [En ligne]. Disponible : https://doi.org/10.1167/iovs.05-1539

[17] J. J. Wang, P. Mitchell, H. Leung, E. Rochtchina, T. Y. Wong et R. Klein, "Hypertensive Retinal Vessel Wall Signs in a General Older Population," *Hypertension*, vol. 42, n°. 4, p. 534–541, oct. 2003, publisher : American Heart Association. [En ligne]. Disponible : https://www.ahajournals.org/doi/10.1161/01.HYP.0000090122.38230.41

[18] W. Smith, J. J. Wang, T. Y. Wong, E. Rochtchina, R. Klein, S. R. Leeder et P. Mitchell, “Retinal Arteriolar Narrowing Is Associated With 5-Year Incident Severe Hypertension,” *Hypertension*, vol. 44, n°. 4, p. 442–447, oct. 2004, publisher : American Heart Association. [En ligne]. Disponible : https://www.ahajournals.org/doi/full/10.1161/01.HYP.0000140772.40322.ec

[19] N. Dervenis, A. L. Coleman, A. Harris, M. R. Wilson, F. Yu, E. Anastasopoulos, P. Founti, T. Pappas, V. Kilintzis et F. Topouzis, “Factors Associated With Retinal Vessel Diameters in an Elderly Population : the Thessaloniki Eye Study,” *Investigative Ophthalmology & Visual Science*, vol. 60, n°. 6, p. 2208–2217, mai 2019. [En ligne]. Disponible : https://doi.org/10.1167/iovs.18-26276

[20] C. Y.-L. Cheung, W. Hsu, M. L. Lee, J. J. Wang, P. Mitchell, Q. P. Lau, H. Hamzah, M. Ho et T. Y. Wong, “A New Method to Measure Peripheral Retinal Vascular Caliber over an Extended Area,” *Microcirculation*, vol. 17, n°. 7, p. 495–503, 2010, _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1549-8719.2010.00048.x. [En ligne]. Disponible : https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1549-8719.2010.00048.x

[21] A. A. Abdulsahib, M. A. Mahmoud, M. A. Mohammed, H. H. Rasheed, S. A. Mostafa et M. S. Maashi, “Comprehensive review of retinal blood vessel segmentation and classification techniques : intelligent solutions for green computing in medical images, current challenges, open issues, and knowledge gaps in fundus medical images,” *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 10, n°. 1, p. 20, mars 2021. [En ligne]. Disponible : https://doi.org/10.1007/s13721-021-00294-7

[22] A. Khandouzi, A. Ariafar, Z. Mashayekhpour, M. Pazira et Y. Baleghi, “Retinal Vessel Segmentation, a Review of Classic and Deep Methods,” *Annals of Biomedical Engineering*, vol. 50, n°. 10, p. 1292–1314, oct. 2022. [En ligne]. Disponible : https://doi.org/10.1007/s10439-022-03058-0

[23] K. S. Kumar et N. P. Singh, “Analysis of retinal blood vessel segmentation techniques : a systematic survey,” *Multimedia Tools and Applications*, vol. 82, n°. 5, p. 7679–7733, févr. 2023. [En ligne]. Disponible : https://doi.org/10.1007/s11042-022-13388-9

[24] Q. Qin et Y. Chen, “A review of retinal vessel segmentation for fundus image analysis,” *Engineering Applications of Artificial Intelligence*, vol. 128, p. 107454, févr. 2024. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S095219762301638X

[25] Z. Liu, M. S. Sunar, T. S. Tan et W. H. W. Hitam, “Deep learning for retinal vessel segmentation : a systematic review of techniques and applications,”

Medical & Biological Engineering & Computing, févr. 2025. [En ligne]. Disponible : https://doi.org/10.1007/s11517-025-03324-y

[26] Q. Chen, J. Peng, S. Zhao et W. Liu, "Automatic artery/vein classification methods for retinal blood vessel : A review," *Computerized Medical Imaging and Graphics*, vol. 113, p. 102355, avr. 2024. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0895611124000326

[27] H. Hao, M. B. Sasongko, T. Y. Wong, M. Z. C. Azemin, B. Aliahmad, L. Hodgson, R. Kawasaki, C. Y. Cheung, J. J. Wang et D. K. Kumar, "Does Retinal Vascular Geometry Vary with Cardiac Cycle ?" *Investigative Ophthalmology & Visual Science*, vol. 53, n°. 9, p. 5799–5805, août 2012. [En ligne]. Disponible : https://doi.org/10.1167/iovs.11-9326

[28] R. ill, "Retina Identification," dans *Biometrics : Personal Identification in Networked Society*, A. K. Jain, R. Bolle et S. Pankanti, édit. Boston, MA : Springer US, 1996, p. 123–141. [En ligne]. Disponible : https://doi.org/10.1007/0-306-47044-6_6

[29] E. Midena, G. Marchione, S. Di Giorgio, G. Rotondi, E. Longhin, L. Frizziero, E. Pilotto, R. Parrozzani et G. Midena, "Ultra-wide-field fundus photography compared to ophthalmoscopy in diagnosing and classifying major retinal diseases," *Scientific Reports*, vol. 12, n°. 1, p. 19287, nov. 2022, publisher : Nature Publishing Group. [En ligne]. Disponible : https://www.nature.com/articles/s41598-022-23170-4

[30] M. J. Cree, E. Gamble et D. Cornforth, "Colour normalisation to reduce inter-patient and intra-patient variability in microaneurysm detection in colour retinal images," 2005, publisher : The University of Queensland.

[31] M. Niemeijer, J. Staal, B. Ginneken, M. Loog et M. Abramoff, "DRIVE : digital retinal images for vessel extraction," *Methods for evaluating segmentation and indexing techniques dedicated to retinal ophthalmology*, 2004.

[32] A. Hoover, V. Kouznetsova et M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical Imaging*, vol. 19, n°. 3, p. 203–210, mars 2000, conference Name : IEEE Transactions on Medical Imaging. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/845178

[33] C. G. Owen, A. R. Rudnicka, R. Mullen, S. A. Barman, D. Monekosso, P. H. Whincup, J. Ng et C. Paterson, "Measuring Retinal Vessel Tortuosity in 10-Year-Old Children : Validation of the Computer-Assisted Image Analysis of the Retina (CAIAR) Program," *Investigative Ophthalmology & Visual Science*, vol. 50, n°. 5, p. 2004–2010, mai 2009. [En ligne]. Disponible : https://doi.org/10.1167/iovs.08-3018

[34] J. O. *et al.*, “Retinal vessel segmentation by improved matched filtering : evaluation on a new high-resolution fundus image database,” *IET Image Processing*, vol. 7, n°. 4, p. 373–383, 2013, publisher : Wiley Online Library.

[35] J. Zhang, B. Dashtbozorg, E. Bekkers, J. P. W. Pluim, R. Duits et B. M. ter Haar Romeny, “Robust Retinal Vessel Segmentation via Locally Adaptive Derivative Frames in Orientation Scores,” *IEEE Transactions on Medical Imaging*, vol. 35, n°. 12, p. 2631–2644, déc. 2016, conference Name : IEEE Transactions on Medical Imaging. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/7530915

[36] K. Jin, X. Huang, J. Zhou, Y. Li, Y. Yan, Y. Sun, Q. Zhang, Y. Wang et J. Ye, “Fives : A fundus image dataset for artificial Intelligence based vessel segmentation,” *Scientific Data*, vol. 9, n°. 1, p. 475, 2022, publisher : Nature Publishing Group UK London.

[37] T. A. Qureshi, M. Habib, A. Hunter et B. Al-Diri, “A manually-labeled, artery/vein classified benchmark for the DRIVE dataset,” dans *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, juin 2013, p. 485–488, iSSN : 1063-7125. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/6627847

[38] J. I. O. *et al.*, “LES-AV dataset,” févr. 2020. [En ligne]. Disponible : https://figshare.com/articles/dataset/LES-AV_dataset/11857698

[39] R. Hemelings, B. Elen, I. Stalmans, K. Van Keer, P. De Boever et M. B. Blaschko, “Artery–vein segmentation in fundus images using a fully convolutional network,” *Computerized Medical Imaging and Graphics*, vol. 76, p. 101636, 2019, publisher : Elsevier.

[40] X. Lyu, L. Cheng et S. Zhang, “The RETA Benchmark for Retinal Vascular Tree Analysis,” *Scientific Data*, vol. 9, n°. 1, p. 397, juill. 2022, publisher : Nature Publishing Group. [En ligne]. Disponible : https://www.nature.com/articles/s41597-022-01507-y

[41] J. Van Eijgen, J. Fhima, M.-I. Billen Moulin-Romsée, J. A. Behar, E. Christinaki et I. Stalmans, “Leuven-Haifa High-Resolution Fundus Image Dataset for Retinal Blood Vessel Segmentation and Glaucoma Diagnosis,” *Scientific Data*, vol. 11, n°. 1, p. 257, févr. 2024, publisher : Nature Publishing Group. [En ligne]. Disponible : https://www.nature.com/articles/s41597-024-03086-6

[42] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson et M. Goldbaum, “Detection of blood vessels in retinal images using two-dimensional matched filters,” *IEEE Transactions on Medical Imaging*, vol. 8, n°. 3, p. 263–269, sept. 1989, conference Name : IEEE Transactions on Medical Imaging. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/34715/?arnumber=34715

[43] M. Al-Rawi, M. Qutaishat et M. Arrar, “An improved matched filter for blood vessel detection of digital retinal images,” *Computers in Biology and*

Medicine, vol. 37, nº. 2, p. 262–267, févr. 2007. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0010482506000424

[44] X. Meng, Y. Yin, G. Yang, Z. Han et X. Yan, "A framework for retinal vasculature segmentation based on matched filters." *Biomedical engineering online*, vol. 14, p. 94, oct. 2015, country : England tex.citation-subset : IM tex.completed : 2016-06-20 tex.issn-linking : 1475-925X tex.nlm-id : 101147518 tex.owner : NLM tex.pii : 10.1186/s12938-015-0089-2 tex.pmc : PMC4619384 tex.pubmodel : Electronic tex.pubstatus : epublish tex.revised : 2017-02-20.

[45] Y. Wang, G. Ji, P. Lin et E. Trucco, "Retinal vessel segmentation using multiwavelet kernels and multiscale hierarchical decomposition," *Pattern Recognition*, vol. 46, nº. 8, p. 2117–2133, août 2013. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0031320313000241

[46] G. Azzopardi, N. Strisciuglio, M. Vento et N. Petkov, "Trainable COSFIRE filters for vessel delineation with application to retinal images," *Medical Image Analysis*, vol. 19, nº. 1, p. 46–57, janv. 2015. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1361841514001364

[47] O. Chutatape, L. Zheng et S. Krishnan, "Retinal blood vessel detection and tracking by matched Gaussian and Kalman filters," dans *Proceedings of the 20th annual international conference of the IEEE engineering in medicine and biology society. Vol.20 biomedical engineering towards the year 2000 and beyond (cat. no.98ch36286)*. IEEE, 1998.

[48] Y. Tolias et S. Panas, "A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering," *IEEE Transactions on Medical Imaging*, vol. 17, nº. 2, p. 263–273, avr. 1998, conference Name : IEEE Transactions on Medical Imaging. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/700738/?arnumber=700738&tag=1

[49] M. Lalonde, L. Gagnon et M.-C. Boucher, "Non-recursive paired tracking for vessel extraction from retinal images," dans *Vision interface*, 2000, p. 61–68.

[50] M. J. Cree, D. Cornforth et H. Jelinek, "Vessel segmentation and tracking using a two-dimensional model," *IVC New Zealand*, p. 345–350, 2005, publisher : Citeseer.

[51] M. Adel, M. Rasigni, T. Gaidon, C. Fossati et S. Bourennane, "Statistical-based linear vessel structure detection in medical images," dans *2009 16th IEEE International Conference on Image Processing (ICIP)*, nov. 2009, p. 649–652, iSSN : 2381-8549. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/5413396

[52] B. Nayebifar et H. Abrishami Moghaddam, "A novel method for retinal vessel tracking using particle filters," *Computers in Biology and Medicine*, vol. 43, nº. 5, p. 541–548,

juin 2013. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S001048251300036X

[53] M. Martinez-Perez, A. Hughes, A. Stanton, S. Thom, A. Bharath et K. Parker, "Segmentation of retinal blood vessels based on the second directional derivative and region growing," dans *Proceedings 1999 International Conference on Image Processing (Cat. 99CH36348)*, vol. 2, oct. 1999, p. 173–176 vol.2. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/822878/?arnumber=822878&tag=1

[54] A. F. Frangi, W. J. Niessen, K. L. Vincken et M. A. Viergever, "Multiscale vessel enhancement filtering," dans *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*, W. M. Wells, A. Colchester et S. Delp, édit. Berlin, Heidelberg : Springer, 1998, p. 130–137.

[55] B. Zhang, L. Zhang, L. Zhang et F. Karray, "Retinal vessel extraction by matched filter with first-order derivative of Gaussian," *Computers in Biology and Medicine*, vol. 40, nº. 4, p. 438–445, avr. 2010. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0010482510000302

[56] M. Alhussein, K. Aurangzeb et S. I. Haider, "An Unsupervised Retinal Vessel Segmentation Using Hessian and Intensity Based Approach," *IEEE Access*, vol. 8, p. 165 056–165 070, 2020, conference Name : IEEE Access. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/9189768

[57] L. Gang, O. Chutatape et S. Krishnan, "Detection and measurement of retinal vessels in fundus images using amplitude modified second-order Gaussian filter," *IEEE Transactions on Biomedical Engineering*, vol. 49, nº. 2, p. 168–172, févr. 2002, conference Name : IEEE Transactions on Biomedical Engineering. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/979356

[58] M. Sofka et C. V. Stewart, "Retinal Vessel Centerline Extraction Using Multiscale Matched Filters, Confidence and Edge Measures," *IEEE Transactions on Medical Imaging*, vol. 25, nº. 12, p. 1531–1546, déc. 2006, conference Name : IEEE Transactions on Medical Imaging. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/4016172

[59] W. Cai et A. C. S. Chung, "Multi-resolution Vessel Segmentation Using Normalized Cuts in Retinal Images," dans *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006*, R. Larsen, M. Nielsen et J. Sporring, édit. Berlin, Heidelberg : Springer, 2006, p. 928–936.

[60] F. Zana et J.-C. Klein, "Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation," *IEEE Transactions on Image Processing*, vol. 10,

n°. 7, p. 1010–1019, juill. 2001, publisher : Institute of Electrical and Electronics Engineers (IEEE).

[61] F. Rossant, M. Badellino, A. Chavillon, I. Bloch et M. Paques, "A Morphological Approach for Vessel Segmentation in Eye Fundus Images, with Quantitative Evaluation," *Journal of Medical Imaging and Health Informatics*, vol. 1, n°. 1, p. 42–49, mars 2011.

[62] U. T. Nguyen, A. Bhuiyan, L. A. Park et K. Ramamohanarao, "An effective retinal blood vessel segmentation method using multi-scale line detection," *Pattern Recognition*, vol. 46, n°. 3, p. 703 – 715, 2013. [En ligne]. Disponible : http://www.sciencedirect.com/science/article/pii/S003132031200355X

[63] A. Budai, R. Bock, A. Maier, J. Hornegger, G. Michelson et others, "Robust vessel segmentation in fundus images," *International Journal of Biomedical Imaging*, vol. 2013, 2013, publisher : Hindawi.

[64] Y. Zhao, L. Rada, K. Chen, S. P. Harding et Y. Zheng, "Automated Vessel Segmentation Using Infinite Perimeter Active Contour Model with Hybrid Region Information with Application to Retinal Images," *IEEE Transactions on Medical Imaging*, vol. 34, n°. 9, p. 1797–1807, sept. 2015, conference Name : IEEE Transactions on Medical Imaging. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/7055281

[65] K. N. K. Noronha, K. Navya, "Support system for the automated detection of hypertensive retinopathy using fundus images," *International Conference on Electronic Design and Signal Processing (ICEDSP)*, p. 7–11, 2012.

[66] M. Niemeijer, J. Staal, B. van Ginneken, M. Loog et M. D. Abramoff, "Comparative study of retinal vessel segmentation methods on a new publicly available database," dans *Medical imaging 2004 : Image processing*, J. M. Fitzpatrick et M. Sonka, édit. SPIE, mai 2004.

[67] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever et B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging*, vol. 23, n°. 4, p. 501–509, 2004, publisher : IEEE.

[68] J. Soares, J. Leandro, R. Cesar, H. Jelinek et M. Cree, "Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification," *IEEE Transactions on Medical Imaging*, vol. 25, n°. 9, p. 1214–1222, sept. 2006, conference Name : IEEE Transactions on Medical Imaging. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/1677727

[69] E. Ricci et R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification." *IEEE transactions on medical imaging*, vol. 26, n°. 10, p. 1357–1365, oct. 2007, country : United States tex.citation-subset : IM

tex.completed : 2007-11-26 tex.issn-linking : 0278-0062 tex.nlm-id : 8310780 tex.owner : NLM tex.pubmodel : Print tex.pubstatus : ppublish tex.revised : 2007-10-22.

[70] X. You, Q. Peng, Y. Yuan, Y.-m. Cheung et J. Lei, "Segmentation of retinal blood vessels using the radial projection and semi-supervised approach," *Pattern Recognition*, vol. 44, nº. 10-11, p. 2314–2324, oct. 2011, publisher : Elsevier BV.

[71] C. A. Lupascu, D. Tegolo et E. Trucco, "FABC : retinal vessel segmentation using AdaBoost." *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 14, nº. 5, p. 1267–1274, sept. 2010, country : United States tex.citation-subset : IM tex.completed : 2010-12-27 tex.issn-linking : 1089-7771 tex.nlm-id : 9712259 tex.owner : NLM tex.pubmodel : Print-Electronic tex.pubstatus : ppublish tex.revised : 2010-09-03.

[72] E. Grisan et A. Ruggeri, "A divide et impera strategy for automatic classification of retinal vessels into arteries and veins," dans *Proceedings of the 25th annual international conference of the IEEE engineering in medicine and biology society (IEEE cat. no.03ch37439)*, vol. 1, 2003, p. 890–893 Vol.1.

[73] S. G. Vázquez, N. Barreira, M. G. Penedo, M. Ortega et A. Pose-Reino, "Improvements in retinal vessel clustering techniques : towards the automatic computation of the arterio venous ratio," *Computing*, vol. 90, nº. 3, p. 197–217, nov. 2010. [En ligne]. Disponible : https://doi.org/10.1007/s00607-010-0114-z

[74] M. Saez, S. González-Vázquez, M. González-Penedo, M. A. Barceló, M. Pena-Seijo, G. Coll de Tuero et A. Pose-Reino, "Development of an automated system to classify retinal vessels into arteries and veins," *Computer Methods and Programs in Biomedicine*, vol. 108, nº. 1, p. 367–376, oct. 2012. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0169260712000600

[75] A. Zamperini, A. Giachetti, E. Trucco et K. S. Chin, "Effective features for artery-vein classification in digital fundus images," dans *2012 25th IEEE International Symposium on Computer-Based Medical Systems (CBMS)*, juin 2012, p. 1–6, iSSN : 1063-7125. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/6266336

[76] D. Relan, T. MacGillivray, L. Ballerini et E. Trucco, "Retinal vessel classification : Sorting arteries and veins," dans *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, juill. 2013, p. 7396–7399, iSSN : 1558-4615. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/6611267

[77] D. Relan et R. Relan, "Unsupervised sorting of retinal vessels using locally consistent Gaussian mixtures," *Computer Methods and Programs in Biomedicine*, vol. 199, p.

105894, févr. 2021. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0169260720317272

[78] M. Niemeijer, B. van Ginneken et M. D. Abràmoff, "Automatic classification of retinal vessels into arteries and veins," dans *Medical imaging 2009 : Computer-aided diagnosis*, vol. 7260. SPIE, 2009, p. 422–429.

[79] X. Xu, W. Ding, M. D. Abràmoff et R. Cao, "An improved arteriovenous classification method for the early diagnostics of various diseases in retinal image," *Computer Methods and Programs in Biomedicine*, vol. 141, p. 3–9, avr. 2017. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0169260716309002

[80] Q. Mirsharif, F. Tajeripour et H. Pourreza, "Automated characterization of blood vessels as arteries and veins in retinal images," *Computerized Medical Imaging and Graphics*, vol. 37, n°. 7, p. 607–617, oct. 2013. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0895611113001237

[81] F. Huang, B. Dashtbozorg, T. Tan et B. M. ter Haar Romeny, "Retinal artery/vein classification using genetic-search feature selection," *Computer Methods and Programs in Biomedicine*, vol. 161, p. 197–207, juill. 2018. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0169260717312415

[82] F. Huang, B. Dashtbozorg et B. M. Romeny, "Artery/vein classification using reflection features in retina fundus images," *Machine Vision and Applications*, vol. 29, n°. 1, p. 23–34, 2018, publisher : Springer.

[83] C. Kondermann, D. Kondermann et M. Yan, "Blood vessel classification into arteries and veins in retinal images," dans *Medical Imaging 2007 : Image Processing*, vol. 6512. SPIE, mars 2007, p. 1401–1409. [En ligne]. Disponible : https://www.spiedigitallibrary.org/conference-proceedings-of-spie/6512/651247/Blood-vessel-classification-into-arteries-and-veins-in-retinal-images/10.1117/12.708469.full

[84] V. Vijayakumar, D. D. Koozekanani, R. White, J. Kohler, S. Roychowdhury et K. K. Parhi, "Artery/vein classification of retinal blood vessels using feature selection," dans *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, août 2016, p. 1320–1323, iSSN : 1558-4615. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/7590950

[85] S. Irshad, Y. , Xiaoxia, et Y. Zhang, "A new approach for retinal vessel differentiation using binary particle swarm optimization," *Computer Methods in Biomechanics and Biomedical Engineering : Imaging & Visualization*, vol. 9, n°. 5, p. 510–522, sept. 2021, publisher : Taylor & Francis _eprint : https ://doi.org/10.1080/21681163.2020.1870001. [En ligne]. Disponible : https://doi.org/10.1080/21681163.2020.1870001

[86] H. Li, W. Hsu, M. Lee et H. Wang, “A piecewise Gaussian model for profiling and differentiating retinal vessels,” dans *Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429)*, vol. 1, sept. 2003, p. I–1069, iSSN : 1522-4880. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/1247151

[87] S. G. Vázquez, B. Cancela, N. Barreira, M. G. Penedo, M. Rodrıguez-Blanco, M. P. Seijo, G. C. de Tuero, M. A. Barceló et M. Saez, “Improving retinal artery and vein classification by means of a minimal path approach,” *Machine Vision and Applications*, vol. 24, n°. 5, p. 919–930, juill. 2012, publisher : Springer Nature.

[88] K. Rothaus, X. Jiang et P. Rhiem, “Separation of the retinal vascular graph in arteries and veins based upon structural knowledge,” *Image and Vision Computing*, vol. 27, n°. 7, p. 864–875, juin 2009, publisher : Elsevier BV.

[89] K.-S. Lin, C.-L. Tsai, C.-H. Tsai, M. Sofka, S.-J. Chen et W.-Y. Lin, “Retinal Vascular Tree Reconstruction With Anatomical Realism,” *IEEE Transactions on Biomedical Engineering*, vol. 59, n°. 12, p. 3337–3347, déc. 2012, conference Name : IEEE Transactions on Biomedical Engineering. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/6280637

[90] B. Dashtbozorg, A. M. Mendonça et A. Campilho, “An Automatic Graph-Based Approach for Artery/Vein Classification in Retinal Images,” *IEEE Transactions on Image Processing*, vol. 23, n°. 3, p. 1073–1083, mars 2014, conference Name : IEEE Transactions on Image Processing. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/6517259

[91] E. Pellegrini, G. Robertson, T. MacGillivray, J. van Hemert, G. Houston et E. Trucco, “A graph cut approach to artery/vein classification in ultra-widefield scanning laser ophthalmoscopy,” *IEEE Transactions on Medical Imaging*, vol. 37, n°. 2, p. 516–526, févr. 2018, publisher : Institute of Electrical and Electronics Engineers (IEEE).

[92] Y. Zhao, J. Xie, H. Zhang, Y. Zheng, Y. Zhao, H. Qi, Y. Zhao, P. Su, J. Liu et Y. Liu, “Retinal Vascular Network Topology Reconstruction and Artery/Vein Classification via Dominant Set Clustering,” *IEEE Transactions on Medical Imaging*, vol. 39, n°. 2, p. 341–356, févr. 2020, conference Name : IEEE Transactions on Medical Imaging. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/8754802

[93] J. De, H. Li et L. Cheng, “Tracing retinal vessel trees by transductive inference,” *BMC Bioinformatics*, vol. 15, n°. 1, p. 20, janv. 2014. [En ligne]. Disponible : https://doi.org/10.1186/1471-2105-15-20

[94] J. De, L. Cheng, X. Zhang, F. Lin, H. Li, K. H. Ong, W. Yu, Y. Yu et S. Ahmed, “A Graph-Theoretical Approach for Tracing Filamentary Structures in Neuronal and Retinal Images,” *IEEE Transactions on Medical Imaging*, vol. 35, n°. 1, p. 257–272,

janv. 2016, conference Name : IEEE Transactions on Medical Imaging. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/7219463

[95] R. E. *et al.*, "Retinal artery-vein classification via topology estimation," *IEEE transactions on medical imaging*, vol. 34, n°. 12, p. 2518–2534, 2015, publisher : IEEE.

[96] Q. P. Lau, M. L. Lee, W. Hsu et T. Y. Wong, "Simultaneously Identifying All True Vessels From Segmented Retinal Images," *IEEE Transactions on Biomedical Engineering*, vol. 60, n°. 7, p. 1851–1858, juill. 2013, conference Name : IEEE Transactions on Biomedical Engineering. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/6423262

[97] I. Goodfellow, Y. Bengio et A. Courville, *Deep learning*. MIT press Cambridge, 2016, vol. 1, n°. 2.

[98] Y. Lecun, L. Bottou, Y. Bengio et P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, n°. 11, p. 2278–2324, nov. 1998. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/726791

[99] A. Krizhevsky, I. Sutskever et G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," dans *Advances in Neural Information Processing Systems*, vol. 25. Curran Associates, Inc., 2012. [En ligne]. Disponible : https://proceedings.neurips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html

[100] M. Melinscak, P. Prentasic et S. Loncaric, "Retinal Vessel Segmentation using Deep Neural Networks :," dans *Proceedings of the 10th International Conference on Computer Vision Theory and Applications*. Berlin, Germany : SCITEPRESS - Science and and Technology Publications, 2015, p. 577–582. [En ligne]. Disponible : http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0005313005770582

[101] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang et T. Wang, "A Cross-Modality Learning Approach for Vessel Segmentation in Retinal Images," *IEEE Transactions on Medical Imaging*, vol. 35, n°. 1, p. 109–118, janv. 2016. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/7161344

[102] P. Liskowski et K. Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE Transactions on Medical Imaging*, vol. 35, n°. 11, p. 2369–2380, nov. 2016.

[103] F. Girard et F. Cheriet, "Artery/vein classification in fundus images using CNN and likelihood score propagation," dans *2017 IEEE global conference on signal and information processing (GlobalSIP)*, 2017, p. 720–724.

[104] R. Welikala, P. Foster, P. Whincup, A. Rudnicka, C. Owen, D. Strachan et S. Barman, "Automated arteriole and venule classification using deep learning for retinal images

from the UK Biobank cohort," *Computers in Biology and Medicine*, vol. 90, p. 23–32, nov. 2017, publisher : Elsevier BV.

[105] J. Long, E. Shelhamer et T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," 2015, p. 3431–3440. [En ligne]. Disponible : https://openaccess.thecvf.com/content_cvpr_2015/html/Long_Fully_Convolutional_Networks_2015_CVPR_paper.html

[106] H. Fu, Y. Xu, S. Lin, D. W. Kee Wong et J. Liu, "DeepVessel : Retinal vessel segmentation via deep learning and conditional random field," dans *Medical image computing and computer-assisted intervention – MICCAI 2016*, S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal et W. Wells, édit. Cham : Springer International Publishing, 2016, p. 132–139.

[107] O. Ronneberger, P.Fischer et T. Brox, "U-net : Convolutional networks for biomedical image segmentation," dans *Medical image computing and computer-assisted intervention (MICCAI)*, ser. LNCS, vol. 9351. Springer, 2015, p. 234–241. [En ligne]. Disponible : http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a

[108] R. Biswas, A. Vasan et S. S. Roy, "Dilated Deep Neural Network for Segmentation of Retinal Blood Vessels in Fundus Images," *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol. 44, n°. 1, p. 505–518, mars 2020. [En ligne]. Disponible : https://doi.org/10.1007/s40998-019-00213-7

[109] S. Hussain, F. Guo, W. Li et Z. Shen, "DilUnet : A U-net based architecture for blood vessels segmentation," *Computer Methods and Programs in Biomedicine*, vol. 218, p. 106732, mai 2022. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0169260722001183

[110] D. Yang, G. Liu, M. Ren, B. Xu et J. Wang, "A Multi-Scale Feature Fusion Method Based on U-Net for Retinal Vessel Segmentation," *Entropy*, vol. 22, n°. 8, p. 811, août 2020, number : 8 Publisher : Multidisciplinary Digital Publishing Institute. [En ligne]. Disponible : https://www.mdpi.com/1099-4300/22/8/811

[111] X. Yang, L. Liu et T. Li, "MR-UNet : An UNet model using multi-scale and residual convolutions for retinal vessel segmentation," *International Journal of Imaging Systems and Technology*, vol. 32, n°. 5, p. 1588–1603, 2022, _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1002/ima.22728. [En ligne]. Disponible : https://onlinelibrary.wiley.com/doi/abs/10.1002/ima.22728

[112] W. Liu, H. Yang, T. Tian, Z. Cao, X. Pan, W. Xu, Y. Jin et F. Gao, "Full-Resolution Network and Dual-Threshold Iteration for Retinal Vessel and Coronary Angiograph Segmentation," *IEEE Journal of Biomedical and Health*

Informatics, vol. 26, nº. 9, p. 4623–4634, sept. 2022. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/9815506

[113] J. Gervelmeyer, S. Müller, Z. Huang et P. Berens, "Fundus Image Toolbox : A Python package for fundus image processing," *Journal of Open Source Software*, vol. 10, nº. 108, p. 7101, avr. 2025. [En ligne]. Disponible : https://joss.theoj.org/papers/10.21105/joss.07101

[114] G. Lepetit-Aimon, R. Duval et F. Cheriet, "Large receptive field fully convolutional network for semantic segmentation of retinal vasculature in fundus images," dans *Computational pathology and ophthalmic medical image analysis : First international workshop, COMPAY 2018, and 5th international workshop, OMIA 2018, held in conjunction with MICCAI 2018, granada, spain, september 16-20, 2018, proceedings 5*. Springer, 2018, p. 201–209.

[115] S. Zhang, H. Fu, Y. Yan, Y. Zhang, Q. Wu, M. Yang, M. Tan et Y. Xu, "Attention guided network for retinal image segmentation," dans *Medical image computing and computer assisted intervention–MICCAI 2019 : 22nd international conference, shenzhen, china, october 13–17, 2019, proceedings, part I 22*. Springer, 2019, p. 797–805.

[116] S. Woo, J. Park, J.-Y. Lee et I. S. Kweon, "CBAM : Convolutional Block Attention Module," juill. 2018, arXiv :1807.06521 [cs]. [En ligne]. Disponible : http://arxiv.org/abs/1807.06521

[117] L. Mou, Y. Zhao, L. Chen, J. Cheng, Z. Gu, H. Hao, H. Qi, Y. Zheng, A. Frangi et J. Liu, "CS-Net : Channel and Spatial Attention Network for Curvilinear Structure Segmentation," dans *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap et A. Khan, édit. Cham : Springer International Publishing, 2019, p. 721–730.

[118] X. Li, Y. Jiang, M. Li et S. Yin, "Lightweight Attention Convolutional Neural Network for Retinal Vessel Image Segmentation," *IEEE Transactions on Industrial Informatics*, vol. 17, nº. 3, p. 1958–1967, mars 2021. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/9091247

[119] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Kaiser et I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[120] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit et N. Houlsby, "An Image is Worth 16x16 Words : Transformers for Image Recognition at Scale," juin 2021, arXiv :2010.11929 [cs]. [En ligne]. Disponible : http://arxiv.org/abs/2010.11929

[121] H. Zhang, X. Zhong, Z. Li, Y. Chen, Z. Zhu, J. Lv, C. Li, Y. Zhou et G. Li, "TiM-Net : Transformer in M-Net for Retinal Vessel Segmentation," *Journal of Healthcare Engineering*, vol. 2022, n°. 1, p. 9016401, 2022, _eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1155/2022/9016401. [En ligne]. Disponible : https://onlinelibrary.wiley.com/doi/abs/10.1155/2022/9016401

[122] H. Yu, J.-h. Shim, J. Kwak, J. W. Song et S.-J. Kang, "Vision Transformer-Based Retina Vessel Segmentation with Deep Adaptive Gamma Correction," dans *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, mai 2022, p. 1456–1460, iSSN : 2379-190X. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/9747597

[123] J. Yi et C. Chen, "Multi-Task Segmentation and Classification Network for Artery/Vein Classification in Retina Fundus," *Entropy*, vol. 25, n°. 8, p. 1148, août 2023, number : 8 Publisher : Multidisciplinary Digital Publishing Institute. [En ligne]. Disponible : https://www.mdpi.com/1099-4300/25/8/1148

[124] L. Y. Li, A. A. Isaksen, B. Lebiecka-Johansen, K. Funck, V. Thambawita, S. Byberg, T. H. Andersen, O. Norgaard et A. Hulman, "Prediction of cardiovascular markers and diseases using retinal fundus images and deep learning : a systematic scoping review," *European Heart Journal - Digital Health*, vol. 5, n°. 6, p. 660–669, nov. 2024. [En ligne]. Disponible : https://doi.org/10.1093/ehjdh/ztae068

[125] X. Shen, J. Xu, H. Jia, P. Fan, F. Dong, B. Yu et S. Ren, "Self-attentional microvessel segmentation via squeeze-excitation transformer Unet," *Computerized Medical Imaging and Graphics*, vol. 97, p. 102055, avr. 2022. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0895611122000283

[126] S. Mishra, Y. X. Wang, C. C. Wei, D. Z. Chen et X. S. Hu, "VTG-Net : A CNN Based Vessel Topology Graph Network for Retinal Artery/Vein Classification," *Frontiers in Medicine*, vol. 8, nov. 2021, publisher : Frontiers. [En ligne]. Disponible : https://www.frontiersin.orghttps://www.frontiersin.org/journals/medicine/articles/10.3389/fmed.2021.750396/full

[127] Y. Li, Y. Zhang, J.-Y. Liu, K. Wang, K. Zhang, G.-S. Zhang, X.-F. Liao et G. Yang, "Global Transformer and Dual Local Attention Network via Deep-Shallow Hierarchical Feature Fusion for Retinal Vessel Segmentation," *IEEE Transactions on Cybernetics*, vol. 53, n°. 9, p. 5826–5839, sept. 2023, conference Name : IEEE Transactions on Cybernetics. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/9863763

[128] A. Galdran, A. Anjos, J. Dolz, H. Chakor, H. Lombaert et I. B. Ayed, "State-of-the-art retinal vessel segmentation with minimalistic models," *Scientific Reports*, vol. 12,

n°. 1, p. 6174, avr. 2022, publisher : Nature Publishing Group. [En ligne]. Disponible : https://www.nature.com/articles/s41598-022-09675-y

[129] P. Costa, A. Galdran, M. I. Meyer, M. Niemeijer, M. Abràmoff, A. M. Mendonça et A. Campilho, "End-to-End Adversarial Retinal Image Synthesis," *IEEE Transactions on Medical Imaging*, vol. 37, n°. 3, p. 781–791, mars 2018. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/8055572

[130] A. Lahiri, V. Jain, A. Mondal et P. K. Biswas, "Retinal Vessel Segmentation Under Extreme Low Annotation : A Gan Based Semi-Supervised Approach," dans *2020 IEEE International Conference on Image Processing (ICIP)*, oct. 2020, p. 418–422, iSSN : 2381-8549. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/9190882

[131] P. Andreini, G. Ciano, S. Bonechi, C. Graziani, V. Lachi, A. Mecocci, A. Sodi, F. Scarselli et M. Bianchini, "A Two-Stage GAN for High-Resolution Retinal Image Generation and Segmentation," *Electronics*, vol. 11, n°. 1, p. 60, janv. 2022, number : 1 Publisher : Multidisciplinary Digital Publishing Institute. [En ligne]. Disponible : https://www.mdpi.com/2079-9292/11/1/60

[132] S. Go, Y. Ji, S. J. Park et S. Lee, "Generation of Structurally Realistic Retinal Fundus Images with Diffusion Models," dans *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Seattle, WA, USA : IEEE, juin 2024, p. 2335–2344. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/10678363/

[133] C. Wu, Y. Zou et Z. Yang, "U-GAN : Generative Adversarial Networks with U-Net for Retinal Vessel Segmentation," dans *2019 14th International Conference on Computer Science & Education (ICCSE)*, août 2019, p. 642–646, iSSN : 2473-9464. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/8845397

[134] S. A. Kamran, K. F. Hossain, A. Tavakkoli, S. L. Zuckerbrod, K. M. Sanders et S. A. Baker, "RV-GAN : segmenting retinal vascular structure in fundus photographs using a novel multi-scale generative adversarial network," dans *Medical image computing and computer assisted intervention–MICCAI 2021 : 24th international conference, Strasbourg, France, September 27–October 1, 2021, proceedings, part VIII 24*. Springer, 2021, p. 34–44.

[135] S. Mukherjee et S. T. Acton, "Oriented filters for vessel contrast enhancement with local directional evidence," dans *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, avr. 2015, p. 503–506, iSSN : 1945-8452. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/7163921

[136] S. Shit, J. C. Paetzold, A. Sekuboyina, I. Ezhov, A. Unger, A. Zhylka, J. P. W. Pluim, U. Bauer et B. H. Menze, "clDice - a Novel Topology-Preserving Loss Function for

Tubular Structure Segmentation," dans *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, TN, USA : IEEE, juin 2021, p. 16 555–16 564. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/9578225/

[137] M. E. Gegundez-Arias, A. Aquino, J. M. Bravo et D. Marin, "A Function for Quality Evaluation of Retinal Vessel Segmentations," *IEEE Transactions on Medical Imaging*, vol. 31, n°. 2, p. 231–239, févr. 2012. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/6019055

[138] R. J. Araújo, J. S. Cardoso et H. P. Oliveira, "Topological Similarity Index and Loss Function for Blood Vessel Segmentation," juill. 2021, arXiv :2107.14531 [eess]. [En ligne]. Disponible : http://arxiv.org/abs/2107.14531

[139] A. H. Berger, L. Lux, A. Weers, M. Menten, D. Rueckert et J. C. Paetzold, "Pitfalls of topology-aware image segmentation," déc. 2024, arXiv :2412.14619 [cs]. [En ligne]. Disponible : http://arxiv.org/abs/2412.14619

[140] N. Stucki, J. C. Paetzold, S. Shit, B. Menze et U. Bauer, "Topologically Faithful Image Segmentation via Induced Matching of Persistence Barcodes," dans *Proceedings of the 40th International Conference on Machine Learning*. PMLR, juill. 2023, p. 32 698–32 727, iSSN : 2640-3498. [En ligne]. Disponible : https://proceedings.mlr.press/v202/stucki23a.html

[141] F. Milletari, N. Navab et S.-A. Ahmadi, "V-Net : Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," dans *2016 Fourth International Conference on 3D Vision (3DV)*, oct. 2016, p. 565–571. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/7785132/

[142] T. A. Soomro, A. J. Afifi, J. Gao, O. Hellwich, M. Paul et L. Zheng, "Strided U-Net Model : Retinal Vessels Segmentation using Dice Loss," dans *2018 Digital Image Computing : Techniques and Applications (DICTA)*, déc. 2018, p. 1–8. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/8615770

[143] D. Wang, A. Haytham, J. Pottenburgh, O. Saeedi et Y. Tao, "Hard Attention Net for Automatic Retinal Vessel Segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, n°. 12, p. 3384–3396, déc. 2020. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/9119750

[144] X. Jian, "Research on Arteriovenous Vessel Segmentation in Fundus Images Based on Topology," dans *2024 7th International Conference on Artificial Intelligence and Big Data (ICAIBD)*, mai 2024, p. 405–408, iSSN : 2769-3554. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/10604545

[145] M. Mao, J. Lin et S. Liu, "Topological Coherence Preserving Retinal Vessel Segmentation," dans *2023 13th International Conference on Information Technology*

in Medicine and Education (ITME), nov. 2023, p. 146–150, iSSN : 2474-3828. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/10505477

[146] X. Hu, F. Li, D. Samaras et C. Chen, "Topology-preserving deep image segmentation," *Advances in neural information processing systems*, vol. 32, 2019.

[147] X. Xiao, S. Lian, Z. Luo et S. Li, "Weighted Res-UNet for High-Quality Retina Vessel Segmentation," dans *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*, oct. 2018, p. 327–331, iSSN : 2474-3828. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/8589312

[148] H. Li, W. Hsu, M. L. Lee et T. Y. Wong, "Automatic grading of retinal vessel caliber," *IEEE Transactions on Biomedical Engineering*, vol. 52, n°. 7, p. 1352–1355, juill. 2005. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/1440616

[149] D. Shi, Z. Lin, W. Wang, Z. Tan, X. Shang, X. Zhang, W. Meng, Z. Ge et M. He, "A Deep Learning System for Fully Automated Retinal Vessel Measurement in High Throughput Image Analysis," *Frontiers in Cardiovascular Medicine*, vol. 9, mars 2022, publisher : Frontiers. [En ligne]. Disponible : https://www.frontiersin.org/journals/cardiovascular-medicine/articles/10.3389/fcvm.2022.823436/full

[150] H. G. Scheie, "Evaluation of ophthalmoscopic changes of hypertension and arteriolar sclerosis," *A.M.A. Archives of Ophthalmology*, vol. 49, n°. 2, p. 117–138, févr. 1953.

[151] W. Lotmar, A. Freiburghaus et D. Bracher, "Measurement of vessel tortuosity on fundus photographs," *Albrecht von Graefes Archiv für klinische und experimentelle Ophthalmologie*, vol. 211, n°. 1, p. 49–57, mars 1979. [En ligne]. Disponible : https://doi.org/10.1007/BF00414653

[152] W. E. Hart, M. Goldbaum, B. Côté, P. Kube et M. R. Nelson, "Measurement and classification of retinal vascular tortuosity," *International Journal of Medical Informatics*, vol. 53, n°. 2, p. 239–252, févr. 1999. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1386505698001634

[153] E. Grisan, M. Foracchia et A. Ruggeri, "A Novel Method for the Automatic Grading of Retinal Vessel Tortuosity," *IEEE Transactions on Medical Imaging*, vol. 27, n°. 3, p. 310–319, mars 2008, conference Name : IEEE Transactions on Medical Imaging. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/4359043

[154] E. Trucco, H. Azegrouz et B. Dhillon, "Modeling the Tortuosity of Retinal Vessels : Does Caliber Play a Role ?" *IEEE Transactions on Biomedical Engineering*, vol. 57, n°. 9, p. 2239–2247, sept. 2010, conference Name : IEEE Transactions on Biomedical Engineering. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/5473044

[155] M. A. Mainster, “The fractal properties of retinal vessels : Embryological and clinical implications,” *Eye*, vol. 4, n°. 1, p. 235–241, janv. 1990, publisher : Nature Publishing Group. [En ligne]. Disponible : https://www.nature.com/articles/eye199033

[156] L. S. Liebovitch et T. Toth, “A fast algorithm to determine fractal dimensions by box counting,” *Physics Letters A*, vol. 141, n°. 8, p. 386–390, nov. 1989. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/0375960189908542

[157] P. Zhu, F. Huang, F. Lin, Q. Li, Y. Yuan, Z. Gao et F. Chen, “The Relationship of Retinal Vessel Diameters and Fractal Dimensions with Blood Pressure and Cardiovascular Risk Factors,” *PLOS ONE*, vol. 9, n°. 9, p. e106551, sept. 2014, publisher : Public Library of Science. [En ligne]. Disponible : https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0106551

[158] C. D. Murray, “The Physiological Principle of Minimum Work,” *Proceedings of the National Academy of Sciences*, vol. 12, n°. 3, p. 207–214, mars 1926, publisher : Proceedings of the National Academy of Sciences. [En ligne]. Disponible : https://www.pnas.org/doi/abs/10.1073/pnas.12.3.207

[159] T. Luo, T. J. Gast, T. J. Vermeer et S. A. Burns, “Retinal Vascular Branching in Healthy and Diabetic Subjects,” *Investigative Ophthalmology & Visual Science*, vol. 58, n°. 5, p. 2685, mai 2017. [En ligne]. Disponible : https://pmc.ncbi.nlm.nih.gov/articles/PMC6024667/

[160] M. E. Martínez-Pérez, A. D. Hughes, A. V. Stanton, S. A. Thom, N. Chapman, A. A. Bharath et K. H. Parker, “Geometrical and Morphological Analysis of Vascular Branches from Fundus Retinal Images,” dans *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000*, S. L. Delp, A. M. DiGoia et B. Jaramaz, édit. Berlin, Heidelberg : Springer, 2000, p. 756–765.

[161] S. Garg, J. Sivaswamy et S. Chandra, “Unsupervised curvature-based retinal vessel segmentation,” dans *2007 4th IEEE International Symposium on Biomedical Imaging : From Nano to Macro*, avr. 2007, p. 344–347, iSSN : 1945-8452. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/4193293

[162] A. Perez-Rovira, T. MacGillivray, E. Trucco, K. S. Chin, K. Zutis, C. Lupascu, D. Tegolo, A. Giachetti, P. J. Wilson, A. Doney et B. Dhillon, “VAMPIRE : Vessel assessment and measurement platform for images of the REtina,” dans *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, août 2011, p. 3391–3394, iSSN : 1558-4615. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/6090918

[163] M. M. Fraz, R. A. Welikala, A. R. Rudnicka, C. G. Owen, D. P. Strachan et S. A. Barman, “QUARTZ : Quantitative Analysis of Retinal Vessel Topology and size – An

automated system for quantification of retinal vessels morphology," *Expert Systems with Applications*, vol. 42, n°. 20, p. 7221–7234, nov. 2015. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0957417415003504

[164] Y. Zhou, S. K. Wagner, M. A. Chia, A. Zhao, P. Woodward-Court, M. Xu, R. Struyven, D. C. Alexander et P. A. Keane, "AutoMorph : Automated Retinal Vascular Morphology Quantification Via a Deep Learning Pipeline," *Translational Vision Science & Technology*, vol. 11, n°. 7, p. 12, juill. 2022. [En ligne]. Disponible : https://doi.org/10.1167/tvst.11.7.12

[165] J. Fhima, J. Van Eijgen, I. Stalmans, Y. Men, M. Freiman et J. A. Behar, "PVBM : A Python Vasculature Biomarker Toolbox Based On Retinal Blood Vessel Segmentation." Springer, juill. 2022, p. 296–312, arXiv :2208.00392 [cs]. [En ligne]. Disponible : http://arxiv.org/abs/2208.00392

[166] D. Xu, M. L. Lee et W. Hsu, "Propagation Mechanism for Deep and Wide Neural Networks," dans *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, juin 2019, p. 9212–9220, iSSN : 2575-7075. [En ligne]. Disponible : https://ieeexplore.ieee.org/document/8953303

[167] J. Fhima, J. V. Eijgen, M.-I. B. Moulin-Romsée, H. Brackenier, H. Kulenovic, V. Debeuf, M. Vangilbergen, M. Freiman, I. Stalmans et J. A. Behar, "LUNet : deep learning for the segmentation of arterioles and venules in high resolution fundus images," *Physiological Measurement*, vol. 45, n°. 5, p. 055002, mai 2024, publisher : IOP Publishing. [En ligne]. Disponible : https://dx.doi.org/10.1088/1361-6579/ad3d28

[168] T. Mautuit, P. Cunnac, C. Y. Cheung, T. Y. Wong, S. Hogg, E. Trucco, V. Daien, T. J. MacGillivray, J. Labarère et C. Chiquet, "Concordance between SIVA, IVAN, and VAMPIRE Software Tools for Semi-Automated Analysis of Retinal Vessel Caliber," *Diagnostics*, vol. 12, n°. 6, p. 1317, juin 2022, number : 6 Publisher : Multidisciplinary Digital Publishing Institute. [En ligne]. Disponible : https://www.mdpi.com/2075-4418/12/6/1317

[169] C. Danielescu, M. G. Dabija, A. H. Nedelcu, V. V. Lupu, A. Lupu, I. Ioniuc, G.-E. Gîlcă-Blanariu, V.-C. Donica, M.-L. Anton et O. Musat, "Automated Retinal Vessel Analysis Based on Fundus Photographs as a Predictor for Non-Ophthalmic Diseases—Evolution and Perspectives," *Journal of Personalized Medicine*, vol. 14, n°. 1, p. 45, janv. 2024, number : 1 Publisher : Multidisciplinary Digital Publishing Institute. [En ligne]. Disponible : https://www.mdpi.com/2075-4426/14/1/45

[170] Z. Zhu, D. Shi, P. Guankai, Z. Tan, X. Shang, W. Hu, H. Liao, X. Zhang, Y. Huang, H. Yu, W. Meng, W. Wang, Z. Ge, X. Yang et M. He, "Retinal age gap as a predictive biomarker for mortality risk," *British Journal of Ophthalmology*, vol. 107, n°. 4, p.

547–554, avr. 2023, publisher : BMJ Publishing Group Ltd Section : Clinical science. [En ligne]. Disponible : https://bjo.bmj.com/content/107/4/547

[171] Z. Decencière, E., "Feedback on a publicly distributed image database : the Messidor database." *Image Analysis and Stereology*, p. 231–234, 2014.

[172] Z. L. Teo, Y.-C. Tham, M. Yu, M. L. Chee, T. H. Rim, N. Cheung, M. M. Bikbov, Y. X. Wang, Y. Tang, Y. Lu, I. Y. Wong, D. S. W. Ting, G. S. W. Tan, J. B. Jonas, C. Sabanayagam, T. Y. Wong et C.-Y. Cheng, "Global prevalence of diabetic retinopathy and projection of burden through 2045 : Systematic review and meta-analysis," *Ophthalmology*, vol. 128, nº. 11, p. 1580–1591, 2021. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0161642021003213

[173] P. Hooper, M. C. Boucher, A. Cruess, K. G. Dawson, W. Delpero, M. Greve, V. Kozousek, W.-C. Lam et D. A. Maberley, "Canadian Ophthalmological Society evidence-based clinical practice guidelines for the management of diabetic retinopathy," *Canadian Journal of Ophthalmology*, vol. 47, nº. 2, p. S1–S30, avr. 2012, publisher : Elsevier BV.

[174] P. Lanzetta, a. V. Sarao, P. H. Scanlon, J. Barratt, M. Porta, F. Bandello et A. Loewenstein, "Fundamental principles of an effective diabetic retinopathy screening program," *Acta Diabetologica*, vol. 57, nº. 7, p. 785–798, mars 2020, publisher : Springer Science and Business Media LLC.

[175] A. Y. Y. Cheng, "Canadian Diabetes Association 2013 clinical practice guidelines for the prevention and management of diabetes in Canada," *Canadian Journal of Diabetes*, vol. 37, nº. Suppl 1, p. S1–S3, 2013.

[176] P. Hooper, M. C. Boucher, A. Cruess, K. G. Dawson, W. Delpero, M. Greve, V. Kozousek, W.-C. Lam et D. A. Maberley, "Excerpt from the Canadian Ophthalmological Society evidence-based clinical practice guidelines for the management of diabetic retinopathy," *Canadian Journal of Ophthalmology*, vol. 52, p. S45–S74, nov. 2017, publisher : Elsevier BV.

[177] O. Egunsola, L. E. Dowsett, R. Diaz, M. H. Brent, V. Rac et F. M. Clement, "Diabetic retinopathy screening : a systematic review of qualitative literature," *Canadian Journal of Diabetes*, vol. 45, nº. 8, p. 725–733.e12, déc. 2021, publisher : Elsevier BV.

[178] D. Avidor, A. Loewenstein, M. Waisbourd et A. Nutman, "Cost-effectiveness of diabetic retinopathy screening programs using telemedicine : a systematic review," *Cost Effectiveness and Resource Allocation*, vol. 18, nº. 1, avr. 2020, publisher : Springer Science and Business Media LLC.

[179] R. Raman, S. Srinivasan, S. Virmani, S. Sivaprasad, C. Rao et R. Rajalakshmi, "Fundus photograph-based deep learning algorithms in detecting diabetic retinopathy," *Eye*

(London, England), vol. 33, n°. 1, p. 97–109, nov. 2018, publisher : Springer Science and Business Media LLC.

[180] E. Dugas, J. Jorge et W. Cukierski, "Diabetic retinopathy detection," 2015. [En ligne]. Disponible : https://kaggle.com/competitions/diabetic-retinopathy-detection

[181] V. Gulshan, R. P. Rajan, K. Widner, D. Wu, P. Wubbels, T. Rhodes, K. Whitehouse, M. Coram, G. Corrado, K. Ramasamy, R. Raman, L. Peng et D. R. Webster, "Performance of a deep-learning algorithm vs manual grading for detecting diabetic retinopathy in India," *JAMA Ophthalmology*, vol. 137, n°. 9, p. 987, sept. 2019, publisher : American Medical Association (AMA).

[182] D. Xiao, A. Bhuiyan, S. Frost, J. Vignarajan, M.-L. Tay-Kearney et Y. Kanagasingam, "Major automatic diabetic retinopathy screening systems and related core algorithms : a review," *Machine Vision and Applications*, vol. 30, p. 423–446, 2019, publisher : Springer.

[183] R. Raman, D. Dasgupta, K. Ramasamy, R. George, V. Mohan et D. Ting, "Using artificial intelligence for diabetic retinopathy screening : Policy implications," *Indian Journal of Ophthalmology*, vol. 69, n°. 11, p. 2993, 2021, publisher : Medknow.

[184] P. J. Phillips, C. A. Hahn, P. C. Fontana, D. A. Broniatowski et M. A. Przybocki, "Four principles of explainable artificial intelligence," National Institute of Standards and Technology, Gaithersburg, Maryland, Rapport technique NISTIR 8312, 2020.

[185] C. P. Wilkinson, F. L. Ferris III, R. E. Klein, P. P. Lee, C. D. Agardh, M. Davis, D. Dills, A. Kampik, R. Pararajasegaram, J. T. Verdaguer et others, "Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales," *Ophthalmology*, vol. 110, n°. 9, p. 1677–1682, 2003, publisher : Elsevier.

[186] S. Zachariah, W. Wykes et D. Yorston, "Grading diabetic retinopathy (DR) using the Scottish grading protocol," *Community eye health / International Centre for Eye Health*, vol. 28, n°. 92, p. 72–73, 2015, country : England tex.completed : 2016-07-15 tex.issn-linking : 0953-6833 tex.nlm-id : 8912615 tex.owner : PIP tex.pmc : PMC4944099 tex.pubmodel : Print tex.pubstate : ppublish tex.revised : 2018-11-13.

[187] M. Boucher, J. Qian, M. Brent, D. Wong, T. Sheidow, R. Duval, A. Kherani, R. Dookeran, D. Maberley, A. Samad et V. Chaudhary, "Evidence-based Canadian guidelines for tele-retina screening for diabetic retinopathy : recommendations from the Canadian Retina Research Network (CR2N) Tele-Retina Steering Committee," *Canadian Journal of Ophthalmology*, vol. 55, n°. 1, p. 14–24, févr. 2020, publisher : Elsevier BV.

[188] J. I. Orlando, H. Fu, J. B. Breda, K. Van Keer, D. R. Bathula, A. Diaz-Pinto, R. Fang, P.-A. Heng, J. Kim, J. Lee et others, "Refuge challenge : A unified framework for

evaluating automated methods for glaucoma assessment from fundus photographs," *Medical Image Analysis*, vol. 59, p. 101570, 2020, publisher : Elsevier.

[189] O. Kovalyk, J. Morales-Sánchez, R. Verdú-Monedero, I. Sellés-Navarro, A. Palazón-Cabanes et J.-L. Sancho-Gómez, "PAPILA : Dataset with fundus images and clinical data of both eyes of the same patient for glaucoma assessment," *Scientific Data*, vol. 9, n°. 1, p. 291, 2022, publisher : Nature Publishing Group UK London.

[190] J. H. Kumar, C. S. Seelamantula, J. Gagan, Y. S. Kamath, N. I. Kuzhuppilly, U. Vivekanand, P. Gupta et S. Patil, "Chákṣu : A glaucoma specific fundus image database," *Scientific Data*, vol. 10, n°. 1, p. 70, 2023, publisher : Nature Publishing Group UK London.

[191] L. Lin, M. Li, Y. Huang, P. Cheng, H. Xia, K. Wang, J. Yuan et X. Tang, "The SUSTech-SYSU dataset for automated exudate detection and diabetic retinopathy grading," *Scientific Data*, vol. 7, n°. 1, nov. 2020, publisher : Springer Science and Business Media LLC.

[192] E. Decencière, G. Cazuguel, X. Zhang, G. Thibault, J.-C. Klein, F. Meyer, B. Marcotegui, G. Quellec, M. Lamard, R. Danno, D. Elie, P. Massin, Z. Viktor, A. Erginay, B. Laÿ et A. Chabouis, "TeleOphta : Machine learning and image processing methods for teleophthalmology," *IRBM*, vol. 34, n°. 2, p. 196 – 203, 2013. [En ligne]. Disponible : http://www.sciencedirect.com/science/article/pii/S1959031813000237

[193] Y. Zhou, B. Wang, L. Huang, S. Cui et L. Shao, "A benchmark for studying diabetic retinopathy : Segmentation, grading, and transferability," *IEEE Transactions on Medical Imaging*, vol. 40, n°. 3, p. 818–828, mars 2021, publisher : Institute of Electrical and Electronics Engineers (IEEE).

[194] Q. Wei, X. Li, W. Yu, X. Zhang, Y. Zhang, B. Hu, B. Mo, D. Gong, N. Chen, D. Ding et others, "Learn to segment retinal lesions and beyond," dans *2020 25th international conference on pattern recognition (ICPR)*. IEEE, 2021, p. 7403–7410.

[195] P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabuddhe et F. Meriaudeau, "Indian diabetic retinopathy image dataset (IDRiD) : a database for diabetic retinopathy screening research," *Data*, vol. 3, n°. 3, p. 25, 2018, publisher : MDPI.

[196] T. Li, Y. Gao, K. Wang, S. Guo, H. Liu et H. Kang, "Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening," *Information Sciences*, vol. 501, p. 511–522, 2019. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S0020025519305377

[197] M. D. Abràmoff, J. C. Folk, D. P. Han, J. D. Walker, D. F. Williams, S. R. Russell, P. Massin, B. Cochener, P. Gain, L. Tang, M. La-

mard, D. C. Moga, G. Quellec et M. Niemeijer, “Automated analysis of retinal images for detection of referable diabetic retinopathy,” *JAMA Ophthalmology*, vol. 131, nº. 3, p. 351–357, mars 2013, tex.eprint : https ://jamanetwork.com/journals/jamaophthalmology/articlepdf/1668203/ecs120076_351_357.pdf. [En ligne]. Disponible : https://doi.org/10.1001/jamaophthalmol.2013.1743

[198] M. B. Sasongko, T. Y. Wong, T. T. Nguyen, C. Y. Cheung, J. E. Shaw et J. J. Wang, “Retinal vascular tortuosity in persons with diabetes and diabetic retinopathy,” *Diabetologia*, vol. 54, nº. 9, p. 2409–2416, sept. 2011.

[199] C. Playout, R. Duval et F. Cheriet, “A novel weakly supervised multitask architecture for retinal lesions segmentation on fundus images,” *IEEE Transactions on Medical Imaging*, vol. 38, nº. 10, p. 2434–2444, oct. 2019, publisher : Institute of Electrical and Electronics Engineers (IEEE).

[200] T. Rädsch, A. Reinke, V. Weru, M. D. Tizabi, N. Schreck, A. E. Kavur, B. Pekdemir, T. Roß, A. Kopp-Schneider et L. Maier-Hein, “Labelling instructions matter in biomedical image analysis,” *Nature Machine Intelligence*, vol. 5, nº. 3, p. 273–283, 2023, publisher : Nature Publishing Group UK London.

[201] K. Sechidis, G. Tsoumakas et I. Vlahavas, “On the stratification of multi-label data,” dans *Machine learning and knowledge discovery in databases*. Berlin, Heidelberg : Springer, 2011, p. 145–158.

[202] G. Lepetit-Aimon, C. Playout, M. C. Boucher, R. Duval, M. H. Brent et F. Cheriet, “MAPLES-DR : MESSIDOR Anatomical and Pathological Labels for Explainable Screening of Diabetic Retinopathy,” *Scientific Data*, vol. 11, nº. 1, p. 914, août 2024, publisher : Nature Publishing Group. [En ligne]. Disponible : https://www.nature.com/articles/s41597-024-03739-6

[203] Z. Legault, C. Playout, F. Girard et F. Cheriet, “Graph-based representation of retinal lesions for an interpretable diagnosis of diabetic retinopathy,” dans *Medical Imaging 2025 : Computer-Aided Diagnosis*, vol. 13407. SPIE, avr. 2025, p. 45–54. [En ligne]. Disponible : https://www.spiedigitallibrary.org/conference-proceedings-of-spie/13407/134070A/Graph-based-representation-of-retinal-lesions-for-an-interpretable-diagnosis/10.1117/12.3046239.full

[204] I. Goodfellow, Y. Bengio et A. Courville, “Convolutional Networks,” dans *Deep Learning*. MIT press Cambridge, 2016, vol. 1, nº. 2. [En ligne]. Disponible : https://www.deeplearningbook.org/contents/convnets.html

[205] G. Cheng, J. Han, P. Zhou et D. Xu, “Learning rotation-invariant and fisher discriminative convolutional neural networks for object detection,” *IEEE Transactions on*

Image Processing, vol. 28, n°. 1, p. 265–278, 2018, publisher : IEEE.

[206] D. Laptev, N. Savinov, J. M. Buhmann et M. Pollefeys, "Ti-pooling : transformation-invariant pooling for feature learning in convolutional neural networks," dans *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, p. 289–297.

[207] D. Marcos, M. Volpi et D. Tuia, "Learning rotation invariant convolutional filters for texture classification," dans *2016 23rd international conference on pattern recognition (ICPR)*. IEEE, 2016, p. 2012–2017.

[208] T. Cohen et M. Welling, "Group equivariant convolutional networks," dans *International conference on machine learning*. PMLR, 2016, p. 2990–2999.

[209] E. J. Bekkers, M. W. Lafarge, M. Veta, K. A. Eppenhof, J. P. Pluim et R. Duits, "Roto-translation covariant convolutional networks for medical image analysis," dans *International conference on medical image computing and computer-assisted intervention*. Springer, 2018, p. 440–448.

[210] W. T. Freeman et W. H. A. *et al.*, "The design and use of steerable filters," *IEEE Transactions on Pattern analysis and machine intelligence*, vol. 13, n°. 9, p. 891–906, 1991.

[211] T. S. Cohen et M. Welling, "Steerable cnns," *arXiv preprint arXiv :1612.08498*, 2016.

[212] M. Weiler, F. A. Hamprecht et M. Storath, "Learning steerable filters for rotation equivariant CNNs," dans *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, p. 849–858.

[213] S. Graham, D. Epstein et N. Rajpoot, "Dense steerable filter CNNs for exploiting rotational symmetry in histology images," *IEEE Transactions on Medical Imaging*, vol. 39, n°. 12, p. 4124–4136, 2020, publisher : IEEE.

[214] M. Weiler et G. Cesa, "General e (2)-equivariant steerable cnns," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[215] H. Naderi, L. Goli et S. Kasaei, "Scale equivariant CNNs with scale steerable filters," dans *2020 international conference on machine vision and image processing (MVIP)*. IEEE, 2020, p. 1–5.

[216] B. Li, Q. Wang et G. H. Lee, "FILTRA : Rethinking steerable CNN by filter transform," dans *International conference on machine learning*. PMLR, 2021, p. 6515–6522.

[217] K. He, X. Zhang, S. Ren et J. Sun, "Delving deep into rectifiers : Surpassing human-level performance on imagenet classification," dans *Proceedings of the IEEE international conference on computer vision*, 2015, p. 1026–1034.

[218] J. Bergstra, R. Bardenet, Y. Bengio et B. Kégl, “Algorithms for hyper-parameter optimization,” *Advances in neural information processing systems*, vol. 24, 2011.

[219] R. Estrada, M. J. Allingham, P. S. Mettu, S. W. Cousins, C. Tomasi et S. Farsiu, “Retinal artery-vein classification via topology estimation,” *IEEE Transactions on Medical Imaging*, vol. 34, nº. 12, p. 2518–2534, déc. 2015, publisher : Institute of Electrical and Electronics Engineers (IEEE).

[220] T. C. Lee, R. L. Kashyap et C. N. Chu, “Building Skeleton Models via 3-D Medial Surface Axis Thinning Algorithms,” *CVGIP : Graphical Models and Image Processing*, vol. 56, nº. 6, p. 462–478, nov. 1994. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S104996528471042X

[221] P. J. Schneider, “An algorithm for automatically fitting digitized curves,” dans *Graphics gems*, 1990, p. 612–626.

[222] P. S. Raina, C. Wolfson, S. A. Kirkland, L. E. Griffith, M. Oremus, C. Patterson, H. Tuokko, M. Penning, C. M. Balion, D. Hogan, A. Wister, H. Payette, H. Shannon et K. Brazil, “The Canadian Longitudinal Study on Aging (CLSA),” *Canadian Journal on Aging / La Revue canadienne du vieillissement*, vol. 28, nº. 3, p. 221–229, sept. 2009. [En ligne]. Disponible : https://www.cambridge.org/core/journals/canadian-journal-on-aging-la-revue-canadienne-du-vieillissement/article/canadian-longitudinal-study-on-aging-clsa/EEF88C12B71EC7FED5CF26F005430F20

[223] C. Y.-l. Cheung, W. T. Tay, M. K. Ikram, Y. T. Ong, D. A. De Silva, K. Y. Chow et T. Y. Wong, “Retinal Microvascular Changes and Risk of Stroke,” *Stroke*, vol. 44, nº. 9, p. 2402–2408, sept. 2013, publisher : American Heart Association. [En ligne]. Disponible : https://www.ahajournals.org/doi/full/10.1161/STROKEAHA.113.001738

[224] E. Sandoval-Garcia, S. McLachlan, A. H. Price, T. J. MacGillivray, M. W. J. Strachan, J. F. Wilson et J. F. Price, “Retinal arteriolar tortuosity and fractal dimension are associated with long-term cardiovascular outcomes in people with type 2 diabetes,” *Diabetologia*, vol. 64, nº. 10, p. 2215–2227, oct. 2021. [En ligne]. Disponible : https://doi.org/10.1007/s00125-021-05499-z

[225] “GROBID,” 2008, tex.archiveprefix : swh tex.eprint : 1 :dir :dab86b296e3c3216e2241968f0d63b68e8209d3c. [En ligne]. Disponible : https://github.com/kermitt2/grobid

[226] J. Odstrcilik, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, T. Kubena, P. Cernosek, O. Svoboda et E. Angelopoulou, “Retinal vessel segmentation by improved matched filtering : evaluation on a new high-resolution fundus image database,” *IET Image Processing*, vol. 7, nº. 4, p. 373–383, 2013,

_eprint : https ://onlinelibrary.wiley.com/doi/pdf/10.1049/iet-ipr.2012.0455. [En ligne]. Disponible : https://onlinelibrary.wiley.com/doi/abs/10.1049/iet-ipr.2012.0455

[227] L. Maier-Hein, A. Reinke, P. Godau, M. D. Tizabi, F. Buettner, E. Christodoulou, B. Glocker, F. Isensee, J. Kleesiek, M. Kozubek, M. Reyes, M. A. Riegler, M. Wiesenfarth, A. E. Kavur, C. H. Sudre, M. Baumgartner, M. Eisenmann, D. Heckmann-Nötzel, T. Rädsch, L. Acion, M. Antonelli, T. Arbel, S. Bakas, A. Benis, M. B. Blaschko, M. J. Cardoso, V. Cheplygina, B. A. Cimini, G. S. Collins, K. Farahani, L. Ferrer, A. Galdran, B. van Ginneken, R. Haase, D. A. Hashimoto, M. M. Hoffman, M. Huisman, P. Jannin, C. E. Kahn, D. Kainmueller, B. Kainz, A. Karargyris, A. Karthikesalingam, F. Kofler, A. Kopp-Schneider, A. Kreshuk, T. Kurc, B. A. Landman, G. Litjens, A. Madani, K. Maier-Hein, A. L. Martel, P. Mattson, E. Meijering, B. Menze, K. G. M. Moons, H. Müller, B. Nichyporuk, F. Nickel, J. Petersen, N. Rajpoot, N. Rieke, J. Saez-Rodriguez, C. I. Sánchez, S. Shetty, M. van Smeden, R. M. Summers, A. A. Taha, A. Tiulpin, S. A. Tsaftaris, B. Van Calster, G. Varoquaux et P. F. Jäger, "Metrics reloaded : recommendations for image analysis validation," *Nature Methods*, vol. 21, n°. 2, p. 195–212, févr. 2024, number : 2 Publisher : Nature Publishing Group. [En ligne]. Disponible : https://www.nature.com/articles/s41592-023-02151-z

[228] G. Lepetit-Aimon, M.-C. Boucher, R. Duval et F. Cheriet, "Steered Convolutional Neurons to Better Learn the Classification of Retinal Vessels," dans *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, avr. 2023, p. 1–5, iSSN : 1945-8452. [En ligne]. Disponible : https://ieeexplore.ieee.org/abstract/document/10230840

[229] G. Lepetit-Aimon, L. Mameri, R. Duval et F. Cheriet, "Enhanced Graph Representation for Reliable Retinal Vessel Analysis : CLSA Population-Based Study," dans *2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI)*, avr. 2025, p. 1–5, iSSN : 1945-8452. [En ligne]. Disponible : https: //ieeexplore.ieee.org/abstract/document/10981176

[230] A. Théberge, C. Desrosiers, A. Boré, M. Descoteaux et P.-M. Jodoin, "What matters in reinforcement learning for tractography," *Medical Image Analysis*, vol. 93, p. 103085, avr. 2024. [En ligne]. Disponible : https://www.sciencedirect.com/science/article/pii/S1361841524000100

ANNEXE A VARIANCE DU SIGNAL D'INFÉRENCE DANS UN STEERED CNN

L'équation simplifiée 5.12 définit la $n^{\text{ième}}$ sortie de la $l^{\text{ième}}$ couche convolutive d'un steered CNN par :

$$y_n^{(l)}(i,j) = \sum_m \sum_{k,r} \omega_{n,m,k,r} \cdot \mathrm{z}_{m,k,r}^{(l)}(i,j)$$

où $\omega_{n,m,k,r}$ est un paramètre du modèle, et où $z_{m,k,r}$ est la corrélation croisée entre la $m^{\text{ième}}$ entrée de la couche et le filtre élémentaire $\mathbf{\Psi}_{k,r}$:

$$\mathrm{z}_{m,k,r}^{(l)}(i,j) = \sum_{p,q} x_m^{(l)}(i+p, j+q) \cdot \Psi_{k,r}(p,q)$$

avec :

$$\mathrm{Var}\Big[\mathrm{z}_{m,k,r}\Big] = \mathrm{Var}\Big[x_m^{(l)}\Big] \cdot \sum_{p,q} \Psi_{k,r}^2(p,q) = \mathrm{Var}\Big[x_m^{(l)}\Big] \cdot \Big|\mathbf{\Psi}_{k,r}\Big|^2$$

$$\mathrm{E}\Big[\mathrm{z}_{m,k,r}\Big] = \mathrm{E}\Big[x_m^{(l)}\Big] \cdot \sum_{p,q} \Psi_{k,r}(p,q) = \mathrm{E}\Big[x_m^{(l)}\Big] \cdot \sum \mathbf{\Psi}_{k,r}$$

Notons qu'en normalisant chaque filtre élémentaire de sorte que $\Big|\mathbf{\Psi}_{k,r}\Big|_2 = 1$, on assure que la variance de $z_{m,k,r}^{(l)}$ soit la même que celle de $x_m^{(l)}$.

En supposant que la sortie correspondante de la couche convolutive précédente $y_m^{(l-1)}$ suive une distribution gaussienne de moyenne nulle, alors l'espérance et la variance de $x_m^{(l)} = \mathrm{ReLU}\Big(y_m^{(l-1)}\Big)$ sont celles d'une demi-gaussienne :

$$\mathrm{E}\Big[x_m^{(l)}\Big] = \sqrt{\frac{1}{\pi}\mathrm{Var}\Big[y_m^{(l-1)}\Big]} \qquad \text{et} \qquad \mathrm{Var}\Big[x_m^{(l)}\Big] = \frac{1}{2}\mathrm{Var}\Big[y_m^{(l-1)}\Big] \cdot \Big(1 - \frac{2}{\pi}\Big)$$

Par ailleurs, $\omega_{n,m,k,r}$ et $z_{m,k,r}^{(l)}(i,j)$ sont indépendantes et $\mathrm{E}\Big(\omega_{n,m,k,r}\Big) = 0$ on a donc :

$$Var[y_n^{(l)}] = m \cdot \mathrm{N}_{k,r} \cdot \mathrm{Var}\Big[\omega_{n,m,k,r}\Big]\Big(\mathrm{Var}\Big[\mathrm{z}_{m,k,r}^{(l)}\Big] + \mathrm{E}\Big[\mathrm{z}_{m,k,r}^{(l)}\Big]^2\Big)$$

où $\mathrm{N}_{k,r}$ est le nombre de filtres élémentaires $\mathbf{\Psi}_{k,r}$.

La combinaison de toutes ces équations donne :

$$\mathrm{Var}\Big[y_n^{(l)}\Big] = m \cdot \mathrm{N}_{k,r} \cdot \mathrm{Var}\Big[\omega_{n,m,k,r}\Big] \cdot \frac{1}{2}\mathrm{Var}\Big[y_m^{(l-1)}\Big]\left(\Big(1 - \frac{2}{\pi}\Big)\Big|\mathbf{\Psi}_{k,r}\Big|_2^2 + \frac{2}{\pi}\Big(\sum \mathbf{\Psi}_{k,r}\Big)^2\right)$$

$$\frac{\mathrm{Var}\Big[y_n^{(l)}\Big]}{\mathrm{Var}\Big[y_n^{(l-1)}\Big]} = \frac{m \cdot \mathrm{N}_{k,r}}{2} \cdot \mathrm{Var}\Big[\omega_{n,m,k,r}\Big] \cdot \left(\Big(1 - \frac{2}{\pi}\Big)\Big|\mathbf{\Psi}_{k,r}\Big|_2^2 + \frac{2}{\pi}\Big(\sum \mathbf{\Psi}_{k,r}\Big)^2\right)$$

En imposant que $\mathrm{Var}\Big[y_n^{(l)}\Big] = \mathrm{Var}\Big[y_n^{(l-1)}\Big]$ on a alors :

$$\mathrm{Var}\Big[\omega_{n,m,k,r}\Big] = \frac{2}{m \cdot \mathrm{N}_{k,r}} \cdot \frac{1}{1 + \frac{2}{\pi}(1 - (\sum \mathbf{\Psi}_{k,r})^2)}$$

Or $\sum \mathbf{\Psi}_{k,r} = 0 \quad \forall k, r \in (\mathbb{N}^*, \mathbb{N}^*)$, on peut donc simplifier la formule ci-dessus lorsque $k > 0$. La variance de $\omega_{n,m,k,r}$ permettant de conserver une variance constante pendant l'inférence est donc :

$$Var\Big[\omega_{n,m,k,r}\Big] = \begin{cases} \dfrac{2}{m \cdot \mathrm{N}_{k,r}} \cdot \dfrac{1}{1 + \frac{2}{\pi}\Big(1 - \Big(\sum_{p,q} \mathbf{\Psi}_{k,r}(p,q)\Big)^2\Big)} & \text{si } k = 0 \\ \dfrac{2}{m \cdot \mathrm{N}_{k,r}} \cdot \dfrac{\pi}{\pi + 2} & \text{si } k > 0 \end{cases}$$