الجمهورية الجزائرية الديمقراطية الشعبية
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
جامعة أبي بكر بلقايد – تلمسان –
Université Aboubakr Belkaïd – Tlemcen –
Faculté de TECHNOLOGIE

MEMOIRE

Présenté pour l'obtention du **diplôme** de **MASTER**

En : Génie biomédical

Spécialité : Imagerie médicale

Par : Rahmoun Souhila

Sujet

Collecte d'une base de données en ophtalmologie : Détection automatique de l'œdème maculaire diabétique

Soutenu publiquement, le 19/06/2025, devant le jury composé de :

M Chikh Mohamed Amine	Professeur	Université de Tlemcen	Président
Mme Feroui Amel	Professeur	Université de Tlemcen	Examinateur
M.Lazouni Mohamed El Amine	MCA	Université de Tlemcen	Encadrant
Mlle Elaouaber Zineb Aziza	MCA	Université de Tlemcen	Co-Encadrante

Année universitaire :2024 /2025

Remerciements

Avant toute personne, je rends grâce à Dieu, le tout-puissant, pour m'avoir accordé la santé, la patience et la force nécessaires pour mener à bien ce travail. Sans sa guidance et sa protection, rien n'aurait été possible.

Je tiens à exprimer toute ma reconnaissance et ma profonde gratitude à **Mr Lazouni Mohamed El Amine**, mon encadrant, pour sa confiance, ses conseils éclairés, et sa présence bienveillante durant chaque étape de ce mémoire. Merci de m'avoir inspirée et encouragée à donner le meilleur de moi-même.

Un remerciement tout particulier à **Mlle Elaouaber Zineb Aziza** ma co-encadrante, dont le soutien, la patience et l'écoute ont été inestimables. Sa précieuse aide m'a apporté une grande sérénité dans ce travail exigeant. Merci pour ta présence constante, ton engagement sincère et ton amitié qui m'a portée jusqu'au bout.

Je remercie également de tout cœur **Mme Lyla Ryma Lazouni**, pour son accompagnement attentif, sa disponibilité et sa grande gentillesse.

Je tiens également à remercier les membres du jury **M Chikh Mohamed Amine** et **Mme Feroui Amel** pour avoir accepté d'évaluer ce travail, et pour leurs remarques constructives et enrichissantes.

Mes remerciements les plus sincères vont à l'équipe de **la Clinique d'Ophtalmologie Lazouni**, composée aussi bien de médecins que de personnel soignant. Vous n'êtes pas seulement des collègues, vous êtes pour moi ma seconde famille. Merci pour votre accueil, votre bienveillance, votre solidarité, et surtout pour le soutien moral inestimable que vous m'avez offert tout au long de cette aventure.

Et parce que rien de tout cela n'aurait été possible sans eux, je souhaite exprimer un remerciement exceptionnel à **ma famille**, et tout particulièrement à **mes parents**., pour leur amour inconditionnel, leur soutien moral et leurs encouragements constants. Un grand merci également à **mes amis**, pour leur présence et leur soutien tout au long de ce parcours.

Enfin, j'exprime ma reconnaissance à toute personne ayant contribué, de près ou de loin, à l'achèvement de ce travail.

Dédicace

Je dédie ce modeste travail aux personnes les plus chers au monde.

À mes chers parents

Pour votre amour inconditionnel, votre patience infinie et votre foi constante en moi. Votre soutien est la fondation sur laquelle j'ai construit ce parcours.

À mes sœurs et mon frère,

Pour votre présence, vos mots rassurants et votre complicité, qui ont tant compté dans les moments d'effort et de doute.

À mes petits neveux et nièces,

Ces petits cœurs pleins de lumière, qui m'apportent une joie sincère.

À tous mes cher(e)s ami(e)s,

Ils vont trouver ici le témoignage d'une fidélité et amitié infinie.

Résumé

L'œdème maculaire diabétique (OMD) est l'une des principales causes de perte de vision chez les personnes atteintes de diabète. Cette pathologie, souvent silencieuse dans ses premiers stades, nécessite un dépistage précoce pour éviter des complications graves et irréversibles. Face à l'augmentation du nombre de patients et au besoin d'un diagnostic rapide, les outils automatisés basés sur l'intelligence artificielle représentent une solution prometteuse.

La première étape de notre travail a consisté à collecter et annoter une base de données rétinienne locale, constituée d'images issues de deux modalités complémentaires : la tomographie en cohérence optique (OCT) et la rétinographie à grand champ. Ces images ont été acquises dans un environnement clinique réel, auprès de patients présentant des cas normaux ou pathologiques, ce qui renforce la pertinence de notre approche.

Sur cette base de données, nous avons ensuite conçu et testé un système automatique capable de classer les images comme normales ou pathologiques, en s'appuyant sur des techniques avancées de traitement d'images et d'apprentissage profond. Plusieurs architectures de réseaux de neurones convolutifs ont été explorées (CNN, VGG16, ResNet50, MobileNetV2, InceptionV3, EfficientNetB0), avec des phases de prétraitement, d'augmentation des données, et de fine-tuning adaptées aux spécificités de chaque modalité d'imagerie.

Nous avons évalué la performance de chaque modèle sur les bases locales, ainsi que sur des bases publiques, en utilisant des métriques standards telles que la précision, la sensibilité, la spécificité, le F1-score et l'AUC. Les résultats obtenus sont très prometteurs : nos modèles ont démontré une capacité élevée à détecter automatiquement les signes de l'OMD, avec des performances comparables, voire supérieures, à ceux de la littérature.

Ce travail confirme le rôle essentiel que peuvent jouer les outils d'intelligence artificielle dans le diagnostic assisté, en particulier pour le dépistage précoce de pathologies oculaires chroniques comme l'OMD.

Mots clés : Œdème maculaire diabétique, Image rétinienne, OCT, Rétinographie, Deep learning, Diagnostic automatique, Intelligence artificielle.

Abstract

Diabetic macular edema (DME) is one of the leading causes of vision loss among individuals with diabetes. This condition, often asymptomatic in its early stages, requires early detection to prevent serious and irreversible visual impairment. In response to the growing number of patients and the need for timely diagnosis, automated tools based on artificial intelligence (AI) are emerging as a promising solution.

The first step of our work involved collecting and annotating a local retinal image dataset, composed of two complementary imaging modalities: optical coherence tomography (OCT) and ultra-widefield color fundus photography. These images were acquired in a real clinical setting from both healthy and pathological cases, enhancing the clinical relevance and diversity of our dataset.

Using this database, we designed and tested an automated system capable of classifying retinal images as either normal or pathological, leveraging advanced image processing techniques and deep learning algorithms. Several convolutional neural network (CNN) architectures were explored, including CNN from scratch, VGG16, ResNet50, MobileNetV2, InceptionV3, and EfficientNetB0. Each model was trained using appropriate preprocessing steps, data augmentation strategies, and fine-tuning procedures adapted to each imaging modality.

The models were evaluated on both the local and public datasets using standard performance metrics, including accuracy, sensitivity, specificity, F1-score, and AUC. The results are highly promising: our models demonstrated strong capabilities in detecting DME, achieving performance levels comparable to or exceeding those reported in the current literature.

This work highlights the important role that AI-based tools can play in computer-aided diagnosis, particularly for the early screening of chronic retinal diseases such as DME.

Keywords: Diabetic macular edema, Annotated dataset, OCT, Fundus photography, Deep learning, Artificial intelligence, Image classification.

ملخص

تُعدّ الوذمة البقعية الناتجة عن السكري من أبرز الأسباب المؤدية إلى فقدان البصر لدى المصابين بداء السكري. وتتميّز هذه الحالة بأنها غالبًا ما تكون بدون أعراض في مراحلها الأولى، مما يجعل الكشف المبكر أمرًا بالغ الأهمية لتفادي مضاعفات بصرية خطيرة ودائمة. ومع تزايد أعداد المرضى والحاجة إلى تشخيص دقيق وسريع، برزت أنظمة الذكاء الاصطناعي كحل واعد لتقديم دعم فعال للتشخيص.

في هذا العمل، تم إنشاء قاعدة بيانات محلية تحتوي على صور شبكية للعين تم جمعها من مرضى حقيقيين، وتضمنت نوعين من الصور :صور مقطعية للعين باستخدام تقنية التصوير البصري المقطعي، وصور لقاع العين واسعة النطاق .وقد شملت البيانات حالات طبيعية ومرضية، مما يعزز مصداقية النموذج المُقترح.

استُخدمت هذه القاعدة لتطوير واختبار نظام آلي قادر على تصنيف الصور إلى سليمة أو مصابة بالمرض، بالاعتماد على تقنيات متقدمة من معالجة الصور والتعلم العميق .تم اعتماد وتجربة عدة نماذج من الشبكات العصبية الالتفافية، مع تنفيذ مراحل من التحسين المسبق للصور، وتكبير البيانات، وضبط دقيق للنماذج بما يتناسب مع خصائص كل نوع من الصور.

تم تقييم أداء النماذج المطورة على القاعدة المحلية، وكذلك على قواعد بيانات عالمية، باستخدام مؤشرات معيارية مثل معدل الدقة، الحساسية، النوعية، معامل F1 ، والمساحة تحت منحنى الأداء .وقد أظهرت النتائج دقة عالية في الكشف التلقائي عن علامات المرض، بأداء يضاهي أو يفوق النتائج المنشورة في الدراسات السابقة.

تُبرز هذه الدراسة الإمكانات الكبيرة التي توفّرها تقنيات الذكاء الاصطناعي في دعم القرارات الطبية، خصوصًا في مجال الكشف المبكر عن أمراض الشبكية المزمنة مثل الوذمة البقعية السكرية.

الكلمات المفتاحية :الوذمة البقعية الناتجة عن السكري، شبكية العين، التصوير المقطعي البصري، تصوير قاع العين، التعلم العميق، التشخيص التلقائي، الذكاء الاصطناعي.

Table des matières

Liste des figures

Liste des tableaux

Liste des acronymes

IA: Intelligence Artificielle

ML: Machine Learning

DL: Deep Learning

CNN: Convolutional Neural Network

CAD: Computer-Aided Detection

TL : Transfer Learning

OMD : Œdème Maculaire Diabétique

RD : Rétinopathie Diabétique

RDNP : Rétinopathie Diabétique Non Proliférante

RDP : Rétinopathie Diabétique Proliférante

OCT : Optical Coherence Tomography

FA : Fluorescein Angiography

EPR : Épithélium Pigmentaire Rétinien

BHR : Barrière Hémato-Rétinienne

DMLA : Dégénérescence Maculaire Liée à l'Âge

VEGF: Vascular Endothelial Growth Factor

CMT: Central Macular Thicknes

AUC : Area Under Curve

VGG16: Visual Geometry Group 16

IDE: Integrated Development Environment

Introduction générale

1. Contexte général

La vision est l'un des sens les plus précieux de l'être humain. Toute altération de cette fonction peut avoir des conséquences profondes sur la qualité de vie, l'autonomie et l'équilibre psychologique d'un individu. C'est pourquoi la préservation de la santé visuelle représente un enjeu majeur en santé publique. Dans ce contexte, le diabète présente l'une des pathologies chroniques les plus menaçantes pour la vision. Par son impact progressif et souvent silencieux sur les vaisseaux sanguins de la rétine, il est responsable des complications oculaires redoutables, notamment l'Œdème Maculaire Diabétique (OMD) qui constitue la première cause de cécité évitable chez les adultes en âge de travailler.

L'OMD se manifeste par une accumulation de liquide dans la macula, due à une perméabilité vasculaire rétinienne accrue. Il peut apparaître à n'importe quel stade de la rétinopathie diabétique (RD), nécessitant une prise en charge rapide afin d'éviter une dégradation irréversible de l'acuité visuelle. Ainsi, la détection précoce de l'OMD et le suivi attentif de son évolution sont cruciaux pour prévenir les complications graves.

La tomographie par cohérence optique (OCT) et la rétinographie jouent aujourd'hui un rôle fondamental dans la détection et le suivi de cette pathologie. L'OCT permet une visualisation en haute résolution des structures rétiniennes quant à la rétinographie, elle fournit une vue d'ensemble des altérations vasculaires rétiniennes.

Les techniques de traitement d'image, et plus récemment les approches d'intelligence artificielle telles que le l'apprentissage profond, ont commencé à s'imposer comme des solutions prometteuses. Grace à ces avancées et en particulier dans le domaine de l'imagerie médicale, plusieurs systèmes d'aide à la détection médicale ont été développés. Ces systèmes permettent d'aider les médecins dans leurs décisions en analysant automatiquement les images issues des examens médicaux. Ils facilitent ainsi la détection des anomalies et améliorent la rapidité et la précision des diagnostics.

Ce mémoire vise à contribuer à l'amélioration de la détection de l'OMD à travers l'analyse d'images rétiniennes (OCT et rétinographies), en collectant une base de données locale issue d'un environnement clinique réel, en appliquant des méthodes avancées de traitement d'images, de classification, et en comparant différentes approches de détection automatique.

2. Problématique

L'OMD est une complication fréquente et invalidante de la RD, constituant aujourd'hui l'une des premières causes de baisse d'acuité visuelle chez les personnes atteintes de diabète, et peut évoluer vers une perte de vision irréversible en l'absence de diagnostic et de traitement précoces.

Selon la Fédération Internationale du Diabète, plus de 500 millions de personnes sont atteintes de diabète dans le monde, et près d'un tiers d'entre elles développeront une forme de RD au cours de leur vie, avec un risque accru de progression vers un OMD [1]. Cette situation soulève une véritable problématique clinique : le nombre de patients diabétiques ne cesse d'augmenter, les ophtalmologistes doivent traiter un volume croissant d'images dans un temps de consultation limité, et le diagnostic repose encore largement sur l'expertise humaine du praticien. Face à cette prévalence croissante, la détection précoce de l'OMD représente un enjeu crucial de santé publique, notamment dans les pays où l'accès régulier à un ophtalmologiste est limité.

Actuellement, les images médicales posent plusieurs défis. Elles sont souvent affectées par des variations de qualité, à la présence d'artefacts ou de bruit, et à des contrastes faibles qui rendent difficile l'identification visuelle précise des zones atteintes. Ces contraintes peuvent affecter la fiabilité du diagnostic, notamment lorsque les images proviennent de contextes cliniques réels et non contrôlés.

Les méthodes de traitement d'image, ont pris de l'ampleur. Elles permettent non seulement de détecter automatiquement les signes d'un OMD sur les images OCT ou rétinographiques, mais aussi d'améliorer la précision et la rapidité des diagnostics. Toutefois, le développement de ces systèmes soulève également certains défis : la disponibilité limitée de bases de données annotées, la diversité des appareils d'imagerie utilisés, la variabilité inter-patients, et la nécessité de modèles robustes capables de généraliser sur des données hétérogènes.

Dans ce cadre, la collecte de bases de données annotées issues d'environnements cliniques réels, combinée à des approches rigoureuses de prétraitement des images, de classification et d'évaluation, s'avère essentielle pour développer des outils de détection efficaces et adaptables.

3. Objectifs et motivations

L'objectif principal de ce travail est la collecte d'une base de données multimodale annotée ainsi que le développement d'un système automatique performant pour la détection précoce de l'œdème maculaire diabétique (OMD) à partir d'images médicales rétiniennes, notamment les images OCT et les rétinographies couleur.

La détection précoce de cette maladie est essentielle pour prévenir la perte irréversible de vision chez les patients diabétiques. Le développement des systèmes d'aide au diagnostic représente donc une approche prometteuse pour améliorer le dépistage et la prise en charge clinique.

4. Contributions

Ce mémoire apporte plusieurs contributions majeures à l'état de l'art dans la détection automatique de l'OMD, à savoir :

- Collecte et annotation d'une base de données multimodale issue d'une clinique d'ophtalmologie, composée d'images OCT et de rétinographies, représentant des conditions réelles d'acquisition avec une diversité de cas normaux et pathologiques.
- Développement d'un système automatique basé sur des réseaux de neurones convolutifs (CNN) permettant la classification binaire (normal vs OMD) à partir d'images OCT et de rétinographies qu'elles soient de bonne ou de mauvaise qualité.
- Construction des nouveaux modèles d'apprentissage profond plus efficaces et précis adaptées aux spécificités des images rétiniennes.
- Évaluation rigoureuse du système sur la base locale ainsi que sur une base de données publique reconnue, et validation avec des experts en ophtalmologie.

Toutes ces contributions ont considérablement amélioré les résultats existants dans la littérature, notamment pour la détection automatique et précise des lésions maculaires caractéristiques de l'OMD à partir des images OCT et des rétinographies.

5. Plan de mémoire

Ce mémoire est organisé en quatre chapitres principaux, chacun structurant une étape essentielle de notre démarche de détection automatique de l'œdème maculaire diabétique (OMD) à partir d'images médicales. Voici un aperçu du contenu de chaque chapitre :

Chapitre 1 : Contexte médical

Ce chapitre introduit les bases anatomiques et pathologiques nécessaires à la compréhension du sujet. Il détaille l'anatomie de l'œil, en insistant sur les structures concernées par l'OMD telles que la rétine et la macula. L'œdème maculaire diabétique (OMD) y est défini avec ses mécanismes physiopathologiques, ses classifications et ses manifestations cliniques. Ce chapitre se termine par

une présentation des principales techniques d'imagerie utilisées pour son diagnostic, notamment l'OCT et la rétinographie, ainsi que les traitements actuels.

Chapitre 2 : État de l'art

Nous présentons dans ce chapitre une synthèse des travaux récents portant sur la détection de l'OMD à l'aide de méthodes d'intelligence artificielle. Les approches sont classées selon les modalités d'images (OCT, rétinographie), et selon les techniques employées (traitement d'images, apprentissage automatique, deep learning). Les performances rapportées dans la littérature sont discutées afin d'identifier les points forts et les limites des méthodes existantes.

Chapitre 3 : Bases de données

Ce chapitre est dédié à la présentation des données utilisées dans ce travail. Nous y décrivons en détail la base de données locale collectée en milieu clinique, en précisant les modalités d'acquisition. Nous présentons également la base publique exploitée pour la complémentarité des tests. Enfin, nous détaillons l'environnement technique utilisé pour le développement des modèles : langages de programmation, bibliothèques logicielles, et outils d'expérimentation.

Chapitre 4 : Méthodes de détection et résultats expérimentaux

Ce dernier chapitre présente l'approche proposée pour la détection automatique de l'OMD. Nous y exposons les différentes architectures de Deep Learning mises en œuvre (CNN, VGG16, ResNet50, etc.), les stratégies d'entraînement, l'augmentation des données, ainsi que les protocoles de validation. Les résultats obtenus sont analysés de manière détaillée en s'appuyant sur des métriques classiques (accuracy, sensibilité, spécificité, AUC, F1-score), et comparés aux approches existantes pour évaluer la performance du système.

Le mémoire se termine par une conclusion générale, récapitulant les apports de ce travail et ouvrant la voie à plusieurs perspectives d'amélioration et de recherche future.

I. Chapitre I : Aspect Médical

1. Introduction

L'œil est un organe sensible, complexe, et délicat dont sa principale fonction est la vision (perception des images et des couleurs), il assure leurs transformations en signaux nerveux afin d'être interprétés par le cerveau sous forme d'images [2]. Grâce à ses structures anatomiques parfaitement organisées, l'œil permet une vision nette, précise et adaptée aux conditions d'éclairage variées.

Cependant, comme tout organe biologique, il peut être le siège de multiples pathologies. Certaines d'entre elles, notamment d'origine métabolique comme le diabète, affectent les structures internes de l'œil, en particulier la rétine. Parmi ces complications, la rétinopathie diabétique (RD) et son évolution vers l'œdème maculaire diabétique (OMD) constituent des causes majeures de baisse de vision, pouvant aller jusqu'à la cécité si elles ne sont pas détectées et traitées à temps.

Dans ce premier chapitre nous allons tout d'abord présenter les notions de base de l'anatomie de l'œil nécessaires à la compréhension de ce mémoire, en mettant la particularité sur la rétine qui est l'objet nécessaire de ce travail, ainsi que les pathologies les plus courantes qui peuvent touchés cette partie de l'œil. L'objectif principale est d'abord de comprendre la structure et le fonctionnement de cet organe afin d'appréhender les différentes affections pouvant altérer la vision et nécessitant une prise en charge spécifique.

2. Anatomie de l'œil

L'œil est l'organe responsable de la vision, il est de forme sphérique, sa longueur est d'environ 24 mm, son poids est de 7 g et son volume est de 6.5cm3 [3]. Chacune de ses structures joue un rôle essentiel dans la vision depuis la capture de la lumière jusqu'à la transmission des images au cerveau via le nerf optique.

Il est divisé en deux segments : le segment antérieur et le segment postérieur [3] . L'œil est constitué de plusieurs parties structurels, parmi eux nous nous intéressons sur : la rétine, la macula, la fovéa et les vaisseaux rétiniens [2] qui sont illustrés dans la figureI.1 qui suit.

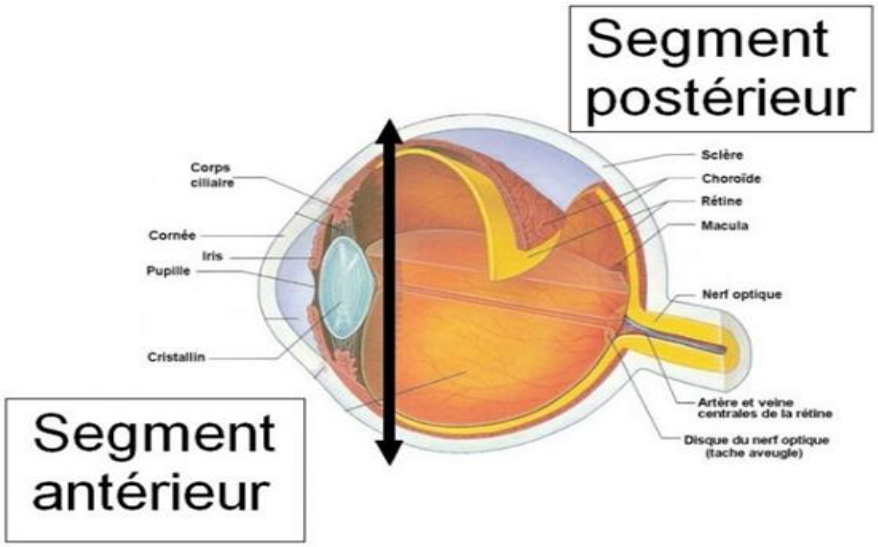


Figure I.1: Anatomie de l'œil humain

2.1 La rétine

La rétine est une fine membrane neurosensorielle, située dans la partie interne du globe oculaire (fond d'œil), sa structure est complexe, à la fois transparente et extrêmement sensible à la lumière. Ce tissu neurosensoriel joue un rôle essentiel dans le processus visuel, il capte la lumière sous forme de signal et le converti en signal électrique via le nerf optique. Cette fonction est réalisée grâce à des cellules appelés photorécepteurs. Ces derniers sont divisés en deux : les cônes et les batônnets [4] comme la montre la figure I.2 qui suit.

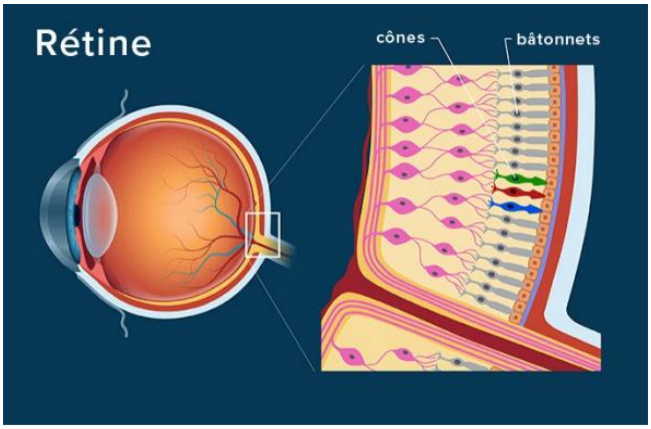


Figure I.2: Schéma de la rétine

Les cônes sont responsables de la vision centrale, ils permettent une perception nette des détails fins ainsi qu'une excellente distinction des couleurs. Les cônes sont particulièrement concentrés dans une zone centrale de la rétine appelée macula, où la vision est la plus nette.

En périphérie, la rétine est principalement composée de bâtonnets, des photorécepteurs particulièrement réactifs aux faibles niveaux de lumière. Ces bâtonnets assurent la vision périphérique et nocturne (vision de nuit), permettant de percevoir les mouvements et les formes dans des conditions de faible éclairage.

La répartition de ces photorécepteurs au sein de la rétine est hétérogène : les bâtonnets dominent en périphérie rétinienne, tandis que les cônes sont massivement présents dans la région centrale, en particulier au niveau de la fovéa, située au cœur de la macula. [5] [4].

Cette organisation offre à la rétine une capacité exceptionnelle à s'adapter aux variations d'intensité lumineuse et à offrir une vision claire et détaillée dans une large gamme de conditions.

2.2 La macula

La macula ou la tâche jaune, est une petite zone située au centre de la rétine ou du fond de l'œil, le nom de tache jaune vient du fait que cette zone possède une coloration jaunâtre par rapport au reste de la rétine [6]. Sa forme est ovale, riche en cônes et est responsable de la vision centrale de haute précision, qui permet de lire, reconnaître les visages ou percevoir les détails fins [7].

La figure I.3 qui suit illustre une image représentative de la macula.

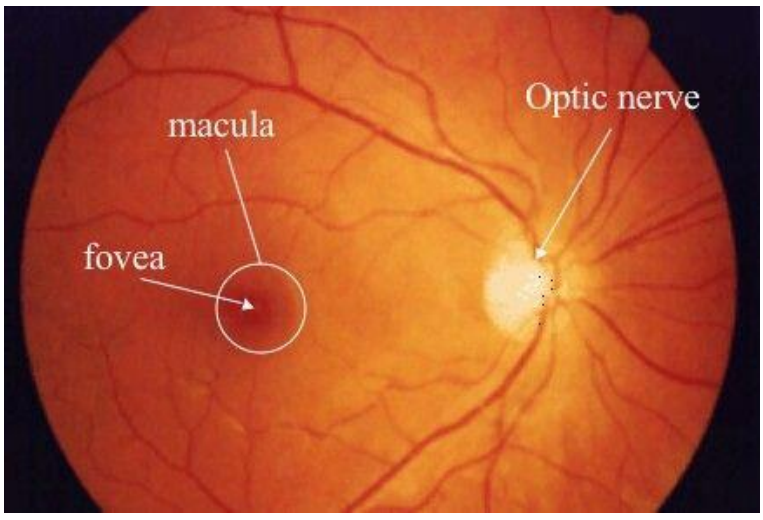


Figure I.3: Image du fond d'œil montrant la macula, la fovéa

2.3 La fovéa

Au centre de la macula se trouve la fovéa, c'est une sorte de dépression d'environs 1,5 mm de diamètre, où la densité des photorécepteurs (cônes) est maximale, sa structure particulière, dépourvue de vascularisation, limite les interférences optiques et permet une perception visuelle nette et précise [7]. Toute atteinte de cette zone entraîne une diminution significative de la vision centrale.

Dans cette petite région l'épaisseur de la rétine est minimale (environ 150 µm) et elle est maximale au niveau de la macula (environ 250 à 300 µm). Cette variabilité anatomique est un critère important dans l'évaluation des pathologies maculaires. Toute modification de la structure ou de l'épaisseur de la macula, comme dans le cas d'un Œdème Maculaire Diabétique (OMD), peut altérer la qualité de la vision centrale. Une augmentation de l'épaisseur rétinienne dans cette région entraîne une déformation des photorécepteurs et une perturbation de la transmission du signal visuel, ce qui se traduit par une vision floue ou déformée [7].

L'analyse de l'épaisseur rétinienne au niveau de la macula et de la fovéa est donc importante dans le diagnostic et le suivi de l'OMD. Une augmentation de l'épaisseur à cet endroit est généralement le premier signe d'un œdème débutant, et la mesure régulière de cette épaisseur permet d'évaluer la réponse au traitement et la progression de la pathologie

2.4 L'épithélium pigmentaire rétinien (EPR)

L'Epithélium Pigmentaire Rétinien (EPR) est une fine couche de cellules situé sous la rétine. Il nourrit les photorécepteurs (cônes et bâtonnets) en leur fournissant les nutriments nécessaires en éliminant les déchets métaboliques. Il participe également à l'absorption de la lumière parasite grâce à la mélanine présente dans ses cellules, ce qui améliore le contraste et la netteté de la vision. L'EPR joue également un rôle dans la régulation de l'équilibre des fluides et des ions dans la rétine. Il contrôle les échanges de nutriments et de déchets entre la rétine et la choroïde, tout en maintenant une barrière protectrice appelée : Barrière Hémato-Rétinienne (BHR). Cette barrière limite le passage de substances potentiellement nocives, protégeant ainsi la rétine des agressions extérieures [8].

Un dysfonctionnement de l'EPR peut entraîner une accumulation de liquide sous la rétine, une dégénérescence des photorécepteurs et une perte progressive de la vision [8]. Le bon

fonctionnement de l'EPR est donc essentiel à la santé rétinienne et à la qualité de la vision. La figure I.4 qui suit illustre une image représentative de l'EPR.

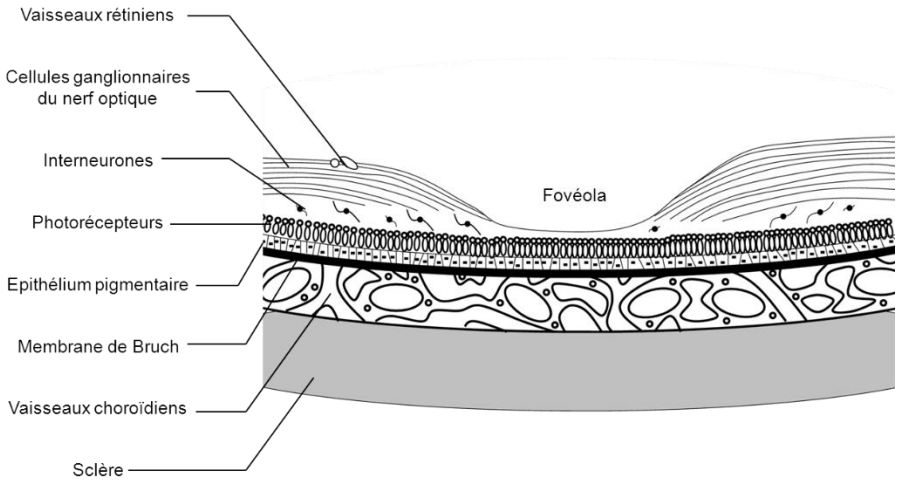


Figure I.4: Couches de la rétine : épithélium pigmentaire rétinien

3. Vascularisation de l'œil

La vascularisation de l'œil joue un rôle essentiel dans le bon fonctionnement de la rétine et le maintien de la vision. Les cellules de la rétine ont besoin d'un apport continu en oxygène et en nutriments pour rester en bonne santé. En même temps, elles produisent des déchets qu'il faut absolument éliminer. Sans cette circulation sanguine, la rétine ne peut pas remplir son rôle. Cette circulation repose en réalité sur deux systèmes complémentaires :

D'une part la circulation choroïdienne, elle est responsable de l'irrigation des couches externes de la rétine, en particulier les photorécepteurs, cependant ces cellules travaillent en continu et consomment beaucoup d'énergie. La circulation choroïdienne leur apporte donc l'oxygène dont elles ont besoin et permet aussi de se débarrasser des déchets, comme le dioxyde de carbone (CO_2) ou les résidus issus de l'activité cellulaire [9].

D'une part la circulation choroïdienne. Elle irrigue principalement les couches externes de la rétine, en particulier les photorécepteurs, ces cellules travaillent en continu et consomment beaucoup d'énergie. La circulation choroïdienne leur apporte donc l'oxygène dont elles ont besoin et permet aussi de se débarrasser des déchets, comme le dioxyde de carbone (CO_2) ou les résidus issus de l'activité cellulaire [9].

D’autre part, la circulation rétinienne prend en charge l’irrigation des couches internes de la rétine. Cette partie du réseau sanguin est constituée de fins vaisseaux, appelés capillaires rétiniens, qui permettent des échanges très rapides entre le sang et les cellules nerveuses de la rétine [10].

Ce système est régulé par un sorte de filtre naturel appelé la barrière hémato-rétinienne (BHR). Elle contrôle ce qui peut entrer ou sortir des vaisseaux sanguins vers les cellules rétiniennes. C’est une barrière protectrice qui empêche les substances nocives d’atteindre la rétine, tout en laissant passer les éléments nécessaires. Il faut savoir que les photorécepteurs, en particulier, consomment énormément d’énergie pour transformer la lumière en signaux électriques [11] [12].

Si la circulation sanguine est perturbée par exemple si les capillaires fuient ou se bouchent cela va directement impacter la vision.

Toute perturbation de cette circulation peut avoir des conséquences directes sur la fonction visuelle, en particulier dans le cas de l'œdème maculaire diabétique (OMD). C’est pourquoi une bonne évaluation de la vascularisation rétinienne est indispensable dans le diagnostic et le suivi des pathologies oculaires.

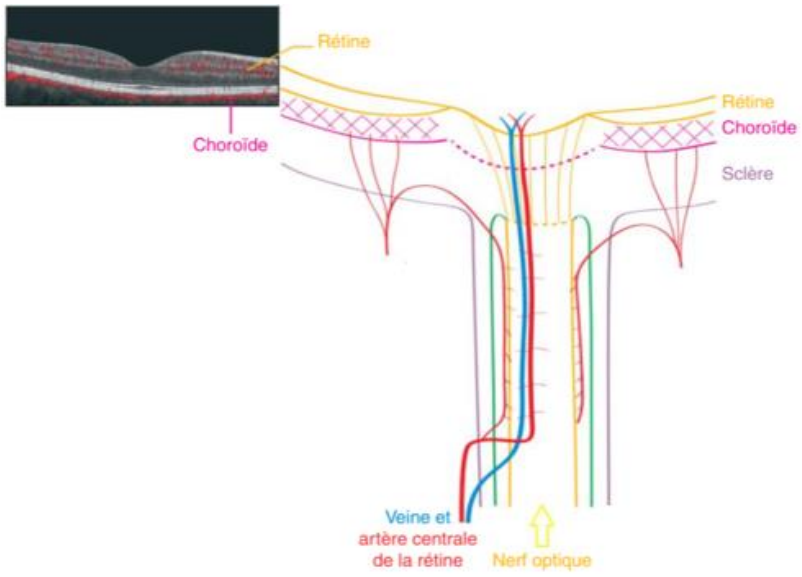


Figure I.5: La vascularisation rétinienne et choroïdienne

4. Les pathologies oculaires

Plusieurs pathologies peuvent affectés la rétine, et la macula lorsqu'elles n'ont pas été diagnostiquées à temps, parmi les principales on cite : la Rétinopathie Riabétique (RD) et l'Œdème Maculaire Diabétique (OMD).

4.1 La rétinopathie diabétique (RD)

4.1.1 Définition de la RD

La rétinopathie diabétique (RD) est l'une des complications les plus fréquentes du diabète sucré, et la cause principale de malvoyance chez les personnes diabétiques. Cette maladie touche la rétine et se développe quand le taux de sucre dans le sang reste trop élevé pendant longtemps.

Cela abîme les petits vaisseaux sanguins de la rétine, ce qui peut entraîner des fuites, des obstructions ou même la formation de nouveaux vaisseaux anormaux. Petit à petit, ces changements peuvent altérer la vision, voire la faire disparaître si rien n'est fait. La RD peut évoluer silencieusement (sans aucun symptôme) pendant plusieurs années avant de provoquer des symptômes visuels [13] [14].

Selon les dernières estimations de la Fédération Internationale du Diabète (IDF), 589 millions d'adultes âgés (entre 20 et 79 ans) vivent actuellement avec le diabète à travers le monde. Ce chiffre est en constante augmentation et pourrait atteindre 853 millions d'ici 2050 [15]. Cependant environ 1 personne diabétique sur 3 développe cette complication, et 1 sur 10 risques une forme sévère qui peut vraiment menacer la vision [16].

Quand les symptômes apparaissent, ils peuvent inclure : une vision floue ou qui change au fil des jours, des taches sombres dans le champ de vision, une perception altérée des couleurs, ou, dans les cas graves, une perte soudaine de la vision comme la montre la figure I.6 qui suit.

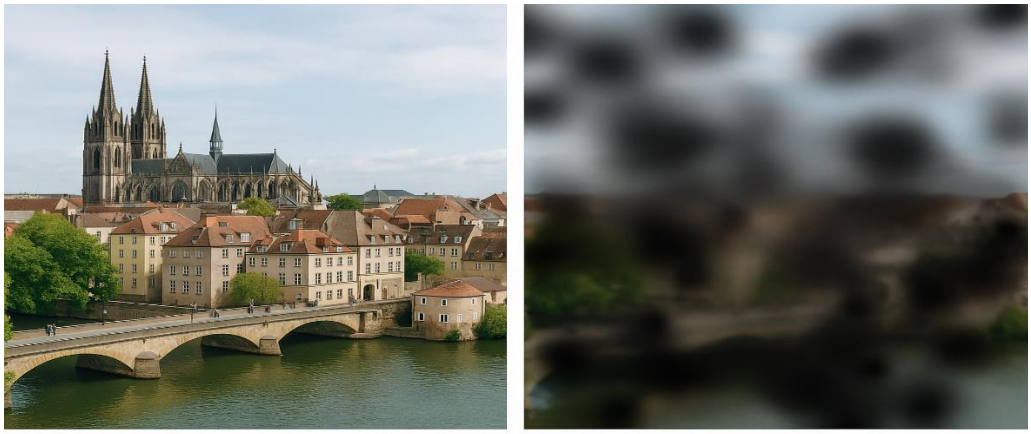


Figure I.6: Image représentative d'une vision normale et d'une personne atteinte de RD

4.1.2 Les lésions associées à la RD

La RD se manifeste par plusieurs types de lésions caractéristiques visibles à l'examen du fond d'œil [17]. On observe notamment :

- **Les microanévrismes** : Premiers signes visibles, qui correspondent à des dilatations localisées des capillaires rétiniens. Ils apparaissent comme des petits points rouges.
- **Les hémorragies rétiniennes** : Ce sont des accumulations du sang dans la rétine. Elles sont dues à des vaisseaux sanguins endommagés qui fuitent du sang. Elles peuvent apparaître sous forme de taches rondes ou de flammes.
- **Les exsudats** : Ce sont des dépôts jaunâtres, formés à partir des protéines et des lipides qui s'échappent des vaisseaux endommagés.
- **Les nodules cotonneux :** Il s'agit de petites zones blanchâtres et floues, causés par une interruption locale de l'apport sanguin. Cette ischémie entraîne l'accumulation de débris issus des fibres nerveuses, ce qui donne cet aspect cotonneux.
- **Les néovaisseaux** : Ce sont de nouveaux vaisseaux sanguins anormaux qui se forment à la surface de la rétine, ils apparaissent lorsque certaines zones de la rétine manquent d'oxygène, un phénomène qu'on appelle ischémie rétinienne. En réponse à ce manque, l'œil libère des substances comme le VEGF (Vascular Endothelial Growth Factor), qui stimulent la croissance de nouveaux vaisseaux pour essayer de compenser ce déficit en oxygène.

La figure I.7 qui suit nous montre les différentes lésions associées à une RD.

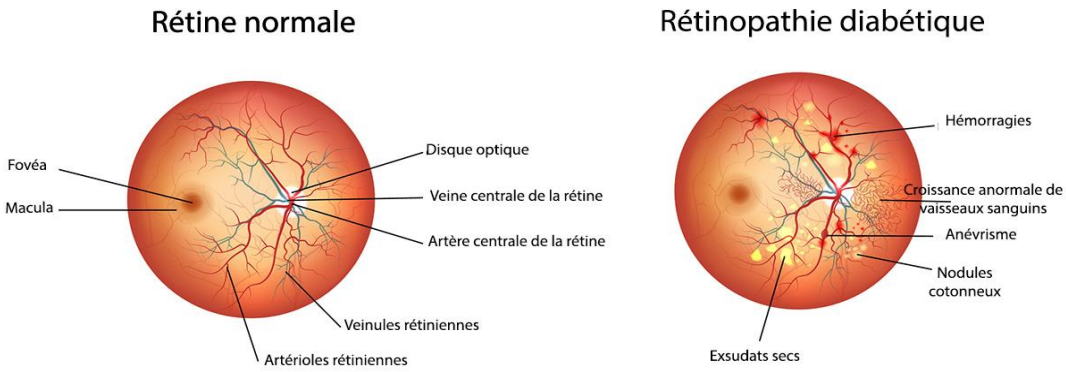


Figure I.7: Les différents lésions de la RD

4.1.3 Les types de la RD

La RD est classée en deux grandes catégories selon son degré de gravité et selon les signes cliniques, dans ce qui suit nous citons : la rétinopathie diabétique non proliférante et proliférante [18].

- **La rétinopathie diabétique non proliférante (RDNP)** : c'est un stade précoce de la maladie, caractérisé par les micro anévrismes, les hémorragies, et exsudats causés par la fuite des vaisseaux sanguins. Elle peut évoluer lentement et rester asymptomatique pendant longtemps [18].
- **La rétinopathie diabétique proliférante (RDP)** : stade avancé de la maladie où la rétine, en souffrance d'oxygène, produit des facteurs de croissance (VEGF) qui stimulent la formation de néo vaisseaux. Ces derniers sont fragiles et peuvent saigner ou provoquer un décollement de la rétine, entraînant une perte sévère de la vision [18].

La classification de la rétinopathie diabétique (RD) repose sur la gravité des altérations rétiniennes observées. Elle comprend plusieurs stades, allant de la rétinopathie diabétique non proliférante (RDNP), elle-même divisée en formes légère, modérée et sévère, jusqu'à la rétinopathie diabétique proliférante (RDP), que nous allons résumer dans le tableau suivant.

Tableau I-1:Les différents stades de la RD

Stade	Description
1 : RDNP minime	Microanévrismes seulement
2 : RDNP modérée	Microanévrismes - Exsudats secs - Nodules cotonneux – Hémorragies rétiniennes.
3 : RDNP sévère	Microanévrismes - Exsudats mous - Nodules cotonneux - Hémorragies rétiniennes. – Anomalie veineuse -Hémorragies intra-rétiniennes étendues.
4 : Rétinopathie diabétique	Néovaisseaux – hémorragies vitré/ pré-rétinienne.

Ces lésions évolutives peuvent, en l'absence de prise en charge, mener à des complications sévères, notamment l'œdème maculaire diabétique (OMD), qui est le principal facteur de perte visuelle chez les patients atteints de RD [19].

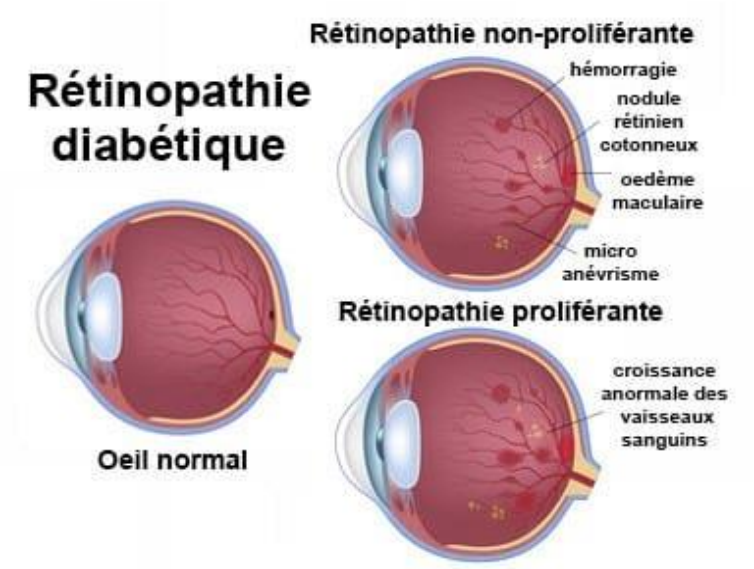


Figure I.8: Types de la RD et présence d'OMD

4.2 L'Œdème Maculaire Diabétique (OMD)

L'OMD est une complication fréquente de la rétinopathie diabétique, représentant une cause majeure de perte de vision chez les personnes atteintes de diabète. Contrairement à d'autres manifestations de la rétinopathie diabétique, l'OMD peut survenir à n'importe quel stade de la maladie, ce qui en fait une préoccupation majeure pour les patients diabétiques [20].

4.2.1 Définition de l'OMD

L'OMD est diagnostiqué et même défini par un gonflement ou un épaississement au niveau de la macula chez les personnes diabétiques, ce gonflement est dû à l'accumulation anormale d'un liquide riche en lipides qui est souvent fuité par des vaisseaux sanguins endommagés et plus précisément la barrière hémato-rétinienne (BHR) [19]. Comme nous avons dit précédemment cette barrière permet une vascularisation rétinienne efficace et contrôlée, cependant son altération entraine une fuite anormale de liquides entrainant ainsi un épaississement maculaire (voir la figure I.8).

En termes d'épidémiologie, environ 10 % des personnes atteintes de diabète développeront un OMD, ce qui représente plusieurs dizaines de millions de patients dans le monde, d'autres études ont montrés qu'une personne sur 3 présente des signes cliniques d'un OMD. Cette complication est aujourd'hui la principale cause de perte de vision chez les diabétiques, soulignant l'importance d'un dépistage et d'une prise en charge précoces [21] [22].

Cliniquement, l'OMD se manifeste par une baisse progressive de l'acuité visuelle, la déformation des images (métamorphopsies), une altération de la perception des couleurs, une vision floue, une diminution du contraste visuel et l'apparition d'une tache noire au centre de la vision. Ces symptômes peuvent considérablement affecter la qualité de vie des patients, rendant les activités quotidiennes telles que la lecture ou la reconnaissance des visages difficiles [23].

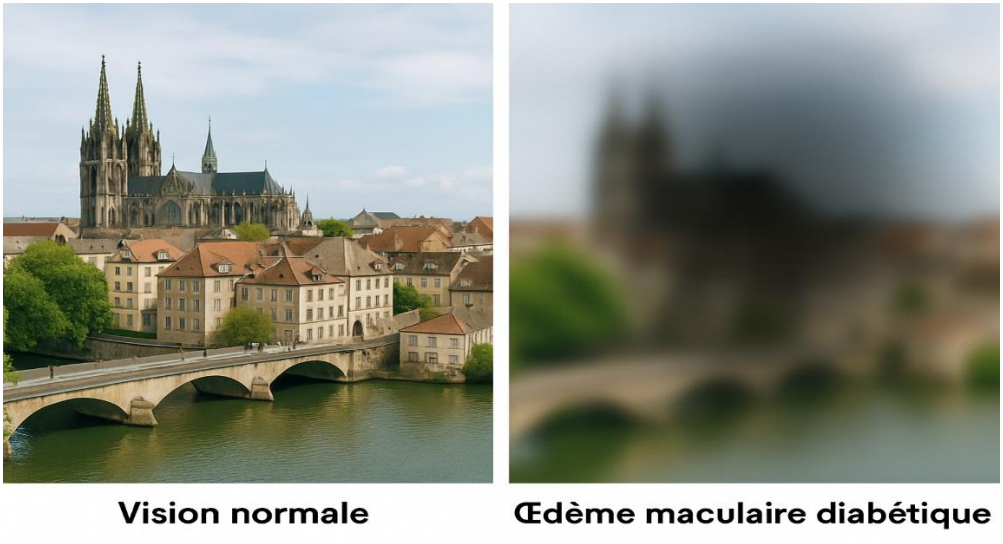


Figure I.9: Vision normale et vision d'un cas atteint d'OMD

4.2.2 Les lésions associées à l'OMD

L'OMD se manifeste par plusieurs caractéristiques que les cliniciens peuvent observer dans différents examens, parmi lesquelles nous citons :

- **Épaississement rétinien** : une augmentation de l'épaisseur de la rétine due à l'accumulation de liquide dans la macula.
- **Exsudats durs** : ce sont des dépôts lipidiques brillants provenant de la fuite de liquide plasmatique, ils sont souvent disposés en couronne autour de la macula.
- **Microanévrismes** : Des petites dilatations des capillaires rétiniens, ils sont souvent à l'origine des fuites.
- **Kystes intrarétiniens** : ce sont de petites cavités remplies de liquide qui se forment à l'intérieur des couches de la rétine, en particulier dans la région maculaire. Ils résultent de l'accumulation de fluide due à une rupture de la BHR.
- **Ischémie maculaire** : c'est la réduction ou l'interruption de la circulation sanguine au niveau de la macula, elle se manifeste par une obstruction des capillaires rétiniens.
- **Les nodules cotonneux** : également appelés *exsudats mous*, sont de petites zones blanchâtres et floues visibles sur la rétine lors d'un examen du fond d'œil. Ils apparaissent en raison d'une interruption de la circulation sanguine dans les petits vaisseaux rétiniens.

4.2.3 Types de l'OMD

L'OMD peut être classé en fonction de sa localisation par rapport au centre de la macula et de sa sévérité, ce qui permet une meilleure évaluation clinique et une prise en charge adaptée. La classification de l'OMD selon la localisation (focale et diffuse) a été proposée par le « Early Treatment Diabetic Retinopathy Study » (ETDRS) dans les années 1980 [24].

L'OMD focal est causé par des fuites localisées provenant de petits gonflements (appelés microanévrismes) dans les vaisseaux sanguins proches de la macula. Ces fuites provoquent une accumulation de liquide dans une zone précise de la rétine, accompagnée parfois de dépôts jaunes appelés exsudats durs [24].

L'OMD diffus, quant à lui, survient lorsque de nombreux petits vaisseaux sanguins dans la rétine deviennent fragiles et laissent passer du liquide de manière anormale. Cela permet au liquide

de s'échapper un peu partout dans la rétine, ce qui entraîne un gonflement plus étendu de la macula [24].

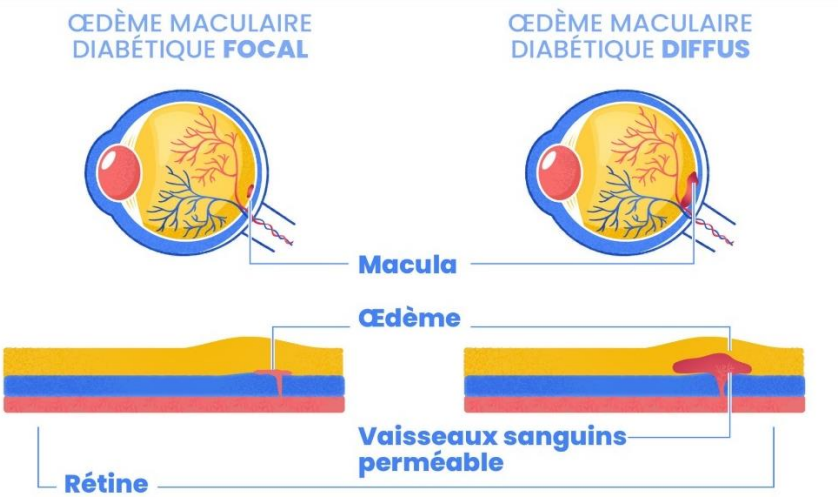


Figure I.10: Types de l''OMD (focal et diffus)

L'OMD est classifié aussi selon sa sévérité, basée sur la classification internationale de la société américaine d'ophtalmologie (American Academy of Ophthalmology), afin d'aider les cliniciens à déterminer le traitement approprié et à évaluer le pronostic visuel des patients atteints d'OMD [24]. La classification détaillée est illustrée dans le tableau I-2 ci-dessous.

Tableau I-2: Classification de l'œdème maculaire diabétique selon la sévérité

Type d'œdème maculaire diabétique	Description
OMD **minime**	Épaississement rétinien ou présence d'exsudats secs au pôle postérieur, mais distants du centre de la fovéa.
OMD **modéré**	Épaississement rétinien ou exsudats secs s'approchant du centre de la macula sans l'atteindre.
OMD **sévère**	Épaississement rétinien ou exsudats secs atteignant le centre de la macula.

5. Techniques d'imageries pour la détection de l'OMD

Afin de limiter l'impact de ces pathologies sur la santé visuelle, il est essentiel d'identifier et diagnostiquer ces pathologies à un stade précoce et adopter une stratégie thérapeutique adaptée.

Pour cela plusieurs modalités d'imagerie et d'analyse avancés ont été développés visant à repérer les anomalies structurelles et vasculaires de la rétine, notamment la rétinographie, l'angiographie et la tomographie par cohérence optique (OCT). Ces méthodes permettent de visualiser les structures de fond d'œil avec une précision, d'identifier les anomalies, d'effectuer un diagnostic qui soit précoce et précis et de suivre l'évolution des maladies rétiniennes, pour enfin parvenir à des stratégies thérapeutiques.

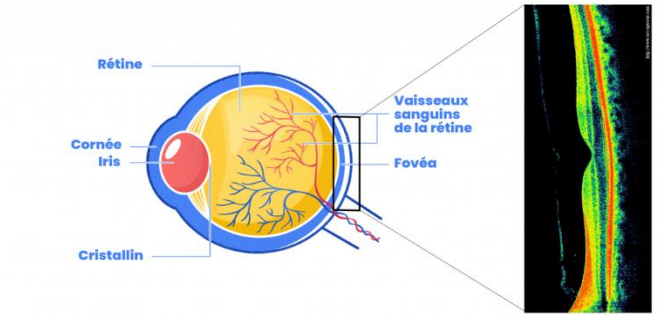
Par la suite nous allons présenter en détails les principales techniques de détection.

5.1 Tomographie par cohérence optique (OCT)

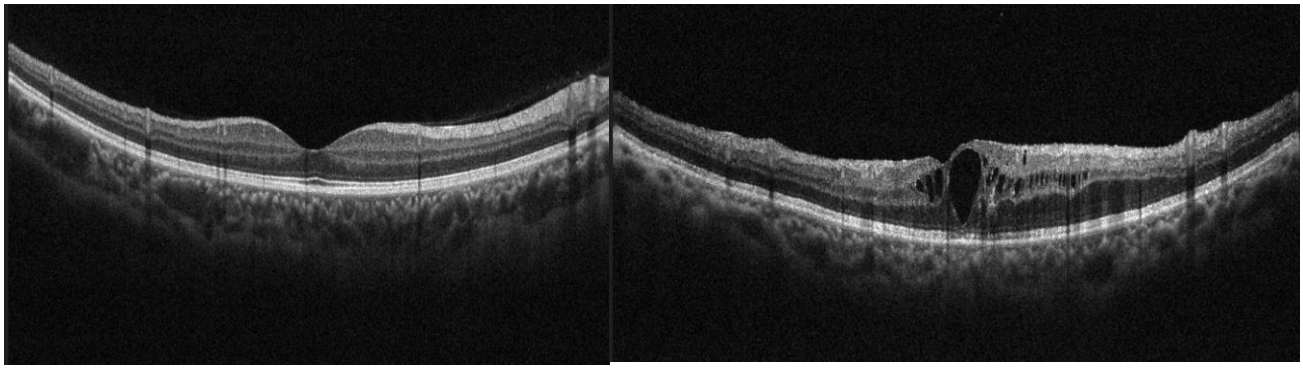
La tomographie en cohérence optique (OCT) est une technique d'imagerie moderne, non-invasive (sans contact), permettant la réalisation de coupes rétiniennes de haute résolution, afin de détecter la présence de liquide intra-rétinien ou sous-rétinien chez les personnes atteintes d'OMD. Son principe de fonctionnement est proche de celui de l'échographie en mode B, mais la formation de l'image dépend des propriétés optiques et non pas acoustiques des tissus traversés. Elle utilise une source de lumière proche de l'infrarouge qui traverse les tissus rétiniens, permettant de mesurer les réflexions et d'obtenir une image en coupe détaillée de la structure rétinienne [25]. Elle fournit de manière fiable et objective des images de la rétine. Elle permet également la mesure de l'épaisseur et du volume rétinien avec une résolution de 5 à 10 µ [26] .

L'OCT est particulièrement utile pour évaluer l'épaisseur rétinienne, et identifier les modifications structurelles de la macula. Elle permet de classer l'OMD en fonction du type d'accumulation de liquide, qu'il soit spongieux, kystique ou séreux. De plus, l'OCT offre une mesure quantitative précise de l'épaisseur rétinienne centrale, essentielle pour évaluer la gravité de l'œdème et suivre la réponse au traitement. Elle permet de poser certains diagnostics, de confirmer d'autres, de faire un suivi régulier de certaines pathologies surtout chroniques et d'apporter des renseignements pronostiques comme dans le cas de l'OMD [27].

(a)



(b)



(c) (d)

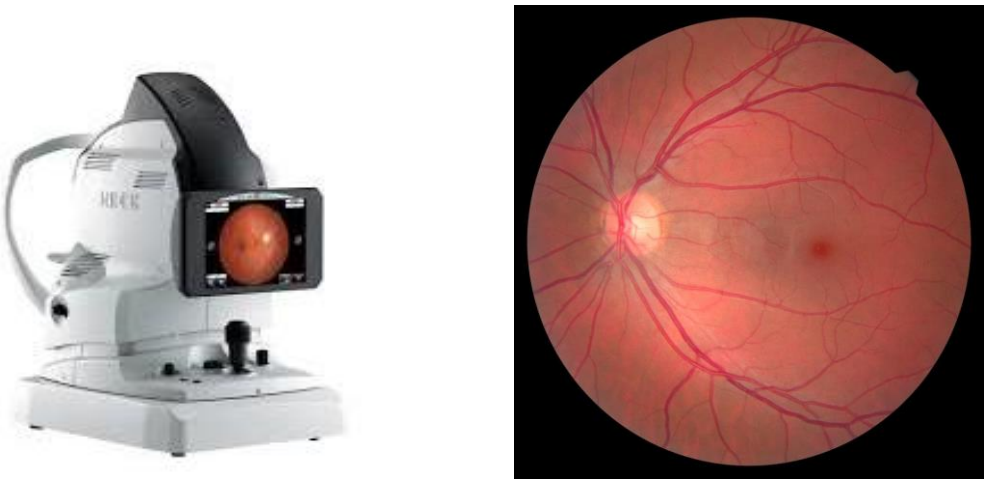
Figure I.11: Examen OCT : (a) Appareil OCT Solix ; (b) : Coupe de l'OCT ; (c) une rétine seine ; (d) : une rétine affectée par l'OMD

5.2 Rétinographie en couleur

La rétinographie en couleur est une technique d'imagerie simple et non invasive, elle permet de capturer des images détaillées du fond de l'œil dans une vue 2D élargie, en particulier de la rétine. Cet examen est indolore, rapide et peut souvent être réalisé sans dilatation pupillaire, bien qu'une dilatation puisse être nécessaire pour obtenir des images de meilleure qualité dans certains cas. Les images obtenues permettent une visualisation précise de la macula, de la papille optique et des vaisseaux rétinien. Elle est essentielle pour le dépistage, le diagnostic et le suivi de nombreuses pathologies rétiniennes, notamment la rétinopathie diabétique, et l'œdème maculaire diabétique. Il existe plusieurs types de cette modalité, chacune d'entre elles possède un champ de vision différent.

5.2.1 La rétinographie classique (à champ réduit)

La rétinographie classique est la forme la plus répandue de cet examen. Elle permet de capturer des images du fond de l'œil avec un champ de vision généralement limité à environ 45 à 60 degrés, ce qui correspond principalement à la région centrale de la rétine, sans capturer les périphéries de la rétine.



(a) (b)

Figure I.12: Rétinographie. (a) Rétinographe ; (b) Image du fond d'œil à champs réduit

La rétinographie classique a un champ de vision assez réduit, ce qui signifie qu'elle ne permet pas de voir toute la périphérie de la rétine. Cela peut poser un problème, car certaines lésions importantes peuvent se trouver en dehors de la zone visible. Pour mieux observer l'ensemble de la rétine, on utilise aujourd'hui la rétinographie grand champ, qui permet de voir une zone beaucoup plus large du fond de l'œil. La rétinographie grand champ est une évolution technologique qui permet de capturer des images couvrant une surface beaucoup plus étendue de la rétine. Cette capacité à visualiser une plus grande portion de la rétine en un seul cliché est particulièrement bénéfique pour détecter des anomalies situées en périphérie de la rétine, telles que celles observées dans la rétinopathie diabétique.

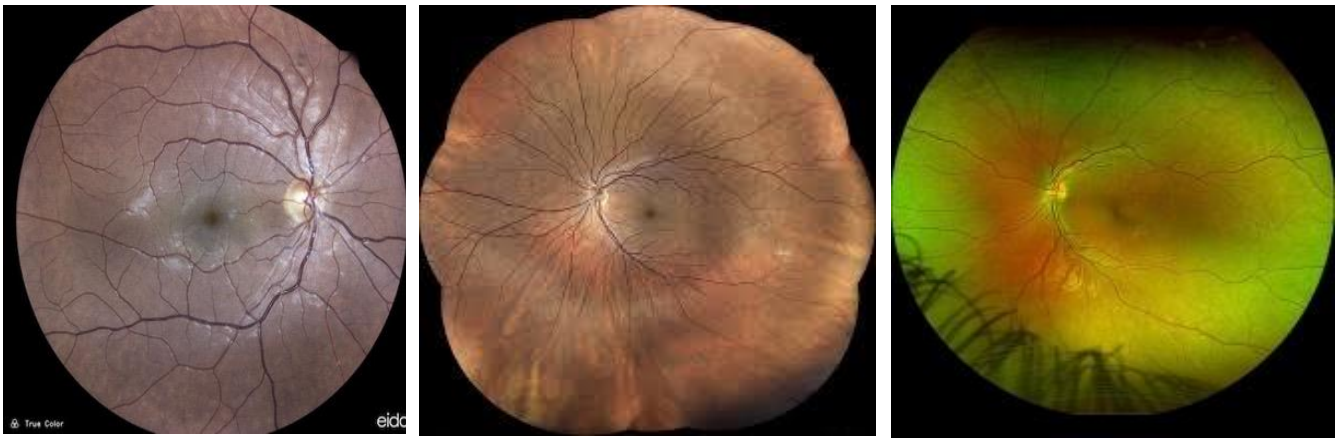
5.2.2 La rétinographie grand champs

La rétinographie grand champs comme le « Eidon », offre un champ de vision plus large allant de 120 à 200 degrés, en une seule prise on obtient un champ de 120 degrés (Figure I.13 (a)), mais avec la fonction mosaïque en capturant 9 quadrants (un montage de 9 images de périphérie

différents et d'angle différents) on obtient une image de champs de vision 200 degrés (ultra grand champs) sois 80% de la surface de la rétine (Figure I.13 (b)) [28] .

L'imagerie ultra grand champ, quant à elle, repousse encore les limites en capturant jusqu'à 200 degrés de la rétine en une seule prise (Figure I.13 (c)) [29]. Offrant ainsi une vue plus détaillée et plus précise y compris les zones périphériques, qui ne sont pas visibles avec une rétinographie classique [30] .Cette technologie avancée, comme celle proposée par « Optos », utilise des lasers à plusieurs longueurs d'onde pour obtenir des images en haute résolution. Elle est particulièrement utile pour visualiser des anomalies vasculaires, telles que des fuites, des obstructions ou des néovaisseaux, avec une précision inégalée.

Les images sont acquises en quelques secondes et sans nécessité de dilatation chimiques des pupilles, ce qui en fait un outil pratique pour une utilisation clinique quotidienne [29].



(a) (b) (c)

Figure I.13: Les différents prises de la rétinographie ; (a) : Image 120 degrés prise par Eidon ; (b) : Image 200 degrés de 9 quadrants prise par Eidon ; (c) : Image 200 degrés prise par Optos

La rétinographie est largement utilisée pour suivre l'évolution des pathologies rétiniennes sur de longues périodes, notamment l'OMD en comparant les images prises à différents moments. Elle offre également des avantages significatifs en matière de coût par rapport à d'autres techniques d'imagerie rétinienne. Grâce à sa portabilité, son rapport coût-efficacité, la qualité des images obtenues et la rapidité d'acquisition, elle constitue une solution particulièrement adaptée à une utilisation clinique [31].

Certes, il ne s'agit pas d'une modalité permettant de renseigner sur l'épaisseur rétinienne ou sur la vascularisation, mais bien utile pour diagnostiquer des anomalies maculaires structurelles telles que les exsudats maculaires ou les hémorragies ou de micro anévrismes, qui sont des marqueurs caractéristiques de la rétinopathie diabétique, ce qui permet d'améliorer la précision du diagnostic [29]. Elle est souvent utilisée en complément de l'OCT et de l'angiographie pour obtenir une vue d'ensemble de la rétine.

Figure I.14: Image rétinienne d'un patient atteint d'OMD

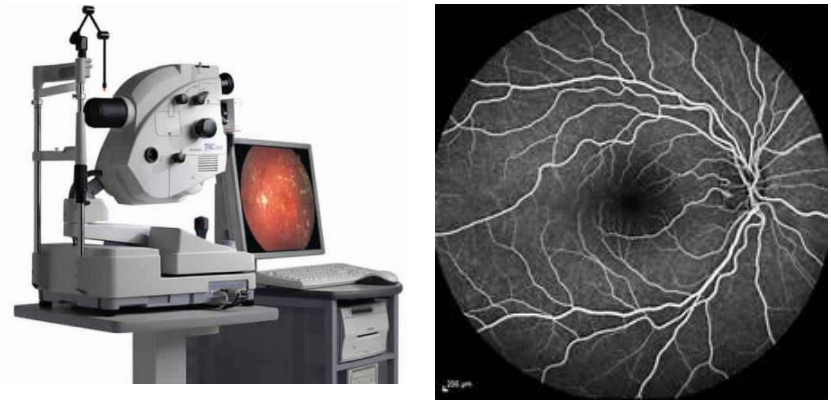
5.3 Angiographie la fluorescéine (FA)

Les altérations de la structure vasculaire rétinienne constituent l'un des principaux indicateurs du diagnostic de nombreuses affections oculaires, en particulier l'œdème maculaire diabétique. L'évaluation précise de cette micro vascularisation est donc essentielle, et peut être réalisée grâce à l'angiographie à la fluorescéine. Cette méthode d'imagerie consiste à capturer une série d'images du fond d'œil avant et après l'injection intraveineuse d'un colorant fluorescent, la fluorescéine sodique [32].

Contrairement aux deux modalités citées précédemment l'angiographie a la fluorescéine (FA) est une technique d'imagerie invasive qui permet d'explorer la vascularisation de la rétine et de visualisé la circulation sanguine. Elle repose sur l'injection d'un colorant fluorescent par voie intraveineuse , ce colorant fluorescent appelé aussi « fluorescéine » circule dans les vaisseaux sanguins de la rétine en émettant une lumière visible lorsqu'il est excité par une source lumineuse

bleue , cette dernière sera captée par une caméra spécifique équipée de filtres permettant non seulement de visualiser la vascularisation rétinienne, mais aussi d'identifier des zones de fuite vasculaire, telles que celles présentes dans l'OMD [33] .

Ces images seront capturées en temps réel offrant ainsi des informations structurelles et fonctionnelles. Cette technique joue un rôle clé dans l'évaluation de la gravité de l'OMD. Elle aide à choisir la stratégie thérapeutique la plus adaptée et permet de suivre l'évolution de la maladie avec précision au fil du temps.



(a) (b)

Figure I.15: Angiographie a la fluorescéine. (a) Angiographe ; (b) Image agiographique

6. Traitements d'OMD

Les anomalies rétiniennes, qu'elles soient d'origine vasculaire, inflammatoire ou dégénérative, représentent une cause majeure de perte de vision dans le monde. Leur impact sur la santé visuelle peut être particulièrement grave lorsqu'elles ne sont pas diagnostiquées ou prises en charge à temps. Heureusement, les avancées médicales et technologiques ont permis de développer une large gamme de traitements visant à ralentir, stabiliser ou parfois même améliorer les atteintes rétiniennes. Le choix du traitement dépend de plusieurs facteurs, notamment la nature de la pathologie, sa sévérité, son évolution, ainsi que le profil du patient. Parmi les approches thérapeutiques les plus couramment utilisées figurent les injections intravitréennes d'agents pharmacologiques (anti-VEGF ou corticoïdes), les traitements par laser, et, dans certains cas, les interventions chirurgicales. Ces traitements permettent non seulement de maîtriser les symptômes, mais aussi de préserver la vision fonctionnelle, voire de la restaurer partiellement lorsque cela est possible.

6.1 Interventions chirurgicales

Dans le traitement de l'œdème maculaire diabétique (OMD), deux principales interventions chirurgicales sont utilisées : la vitrectomie et la chirurgie au laser.

- **La vitrectomie :** Cette intervention est souvent utilisée en cas de décollement de la rétine ou d'hémorragie intraoculaire. Elle est souvent indiquée en cas d'OMD persistant. Elle consiste à retirer le vitré endommagé afin de le remplacer par une solution saline [34]. Elle est souvent réalisée en complément des injections d'anti-VEGF pour maximiser les résultats fonctionnels et anatomiques.
- **Chirurgie au laser :** La chirurgie au laser est une technique médicale qui utilise un faisceau de lumière amplifiée (laser), utilisé pour coaguler les vaisseaux sanguins anormaux ou sceller les déchirures rétiniennes. Dans le cadre de l'OMD elle est utilisée pour sceller les micro anévrismes et réduire l'exsudation. Elle est indiquée lorsque l'œdème est localisé et non diffus [35].

6.2 Traitements pharmaceutiques

Les traitements médicamenteux constituent la première ligne de traitement de l'OMD, en particulier les injections intravitréennes d'anti-VEGF ainsi que les corticoïdes. Ces médicaments permettent de contrôler la perméabilité vasculaire et de réduire l'inflammation rétinienne.

Les molécules d'anti-VEGF (facteur de croissance endothélial vasculaire) sont devenues le traitement de référence pour l'OMD. Le VEGF est une protéine qui stimule la formation de nouveaux vaisseaux sanguins anormaux, contribuant à la fuite de liquide dans la macula. Les anti-VEGF agissent en bloquant cette protéine, réduisant ainsi l'œdème et améliorant l'acuité visuelle. Ils sont administrés par injections intravitréennes et ont révolutionné le traitement de la DMLA humide et de l'OMD. Parmi les anti-VEGF les plus souvent utilisés sont le Ranibizumab [36],l'Aflibercept [37] et le Brolucizumab [38].Chaque molécule est efficace dans la diminution de l'épaisseur maculaire centrale (CMT) et l'amélioration de l'acuité visuelle.

Les corticoïdes, sont des médicaments administrés sous forme d'injections ou d'implants, ils sont utilisés pour leur action anti-inflammatoire en réduisent l'inflammation et l'œdème. Ils sont souvent utilisés en complément des anti-VEGF, pour un traitement fiable et efficace [39].

6.3 Traitements au laser

- **La photocoagulation au laser focal :** Cette technique de traitement consiste à appliquer de petites brûlures au laser sur la rétine, afin de réduire l'œdème maculaire, en fermant les microanévrismes et en limitant les fuites capillaires en réduisant leurs perméabilité, l'objectif de cette technique n'est pas d'améliorer l'acuité visuelle mais bien la stabilisée [35] .
- **Laser micropulsé :** Le laser micropulsé est une technique plus récente, utilisant des impulsions courtes à faible énergie pour stimuler les cellules rétiniennes sans provoquer de dommages thermiques. Il a montré une amélioration significative de l'acuité visuelle dans les formes légères à modérées d'OMD [40].
- **Laser sélectif :** ce type de laser est moins invasif et il cible spécifiquement les vaisseaux anormaux sans endommager les tissus sains [41].

7. Conclusion

Dans ce chapitre, nous avons exploré l'anatomie de l'œil en mettant en avant ses principales structures et leurs rôles essentiels dans la fonction visuelle. Ensuite, nous avons présenté les pathologies oculaires notamment la rétinopathie diabétique et l'œdème maculaire diabétique en insistant sur leurs causes, leurs manifestations cliniques et leurs conséquences. Enfin, nous avons présenté les méthodes de détection et les stratégies thérapeutiques permettant de diagnostiquer et de traiter ces affections, en tenant compte des avancées technologiques dans le domaine de l'ophtalmologie.

Enfin, nous avons abordé les principales méthodes de diagnostic, dont la tomographie par cohérence optique (OCT) et l'angiographie à la fluorescéine (FA), qui permettent une détection précoce et un suivi précis de ces pathologies. Ces techniques d'imagerie avancées jouent un rôle crucial dans l'élaboration des stratégies thérapeutiques visant à préserver la fonction visuelle.

L'ensemble des connaissances acquises dans ce chapitre constitue une base essentielle pour mieux comprendre les enjeux du diagnostic et du suivi de l'OMD, et ainsi optimiser sa prise en charge à travers les avancées technologiques et les approches thérapeutiques innovantes. Dans le chapitre suivant nous allons voir l'état de l'art en se basant sur les études récentes et sur la littérature.

II. Chapitre II : Etat de l’art

1. Introduction

Au cours de ces dernières années, le domaine de l'ophtalmologie a connu de grands progrès grâce à l'intégration de l'intelligence artificielle dans les systèmes d'aide au diagnostic. Ces outils sont désormais capables d'analyser avec précision les structures de l'œil, permettant de distinguer efficacement les cas sains des cas pathologiques. L'analyse d'images médicales joue ainsi un rôle central dans le dépistage précoce des maladies oculaires, le suivi de leur évolution, ainsi que dans l'évaluation de l'efficacité des traitements.

Parmi les complications associées au diabète, l'œdème maculaire diabétique (OMD) figure parmi les plus fréquentes et fait l'objet de nombreuses recherches. Les travaux dans ce domaine s'articulent généralement autour de deux axes principaux : la détection précise des zones affectées de la rétine et la classification des cas, afin d'aider les professionnels de santé dans leur prise de décision.

Dans ce chapitre, nous présenterons un état de l'art récent sur cette thématique, ainsi que les principaux fondements de l'intelligence artificielle en se focalisant sur l'apprentissage profond. Les modèles d'apprentissage par transfert les plus populaires seront ensuite décris. Enfin, nous terminerons par une description des différents métrique d'évaluation employées pour mesurer les performances des modèles.

2. Etat de l'art

L'œdème maculaire diabétique est une complication oculaire très fréquente, c'est pour cette raison que plusieurs chercheurs se sont intéressés à la recherche dans le développement de méthodes automatisées de détection de cette maladie. Dans cette partie, nous présenterons un état de l'art des méthodes de détection de l'OMD à partir des images rétiniennes, structurée selon le type de modalité utilisée : les méthodes basées sur les images OCT et celles basées sur la rétinographie.

2.1 Méthodes de détection basées sur l'OCT

En 2018, Kaymak et Serener ont mené une étude sur la détection automatisée de OMD à l'aide de l'architecture AlexNet, en s'appuyant sur un jeu de données déséquilibré composé de 26 315 images normales et 11 348 images OMD issues de la base de données publique OCT2017. Le

modèle, évalué sur 500 images (250 par classe), a atteint une précision de 99,8 %, une sensibilité de 100 % et une spécificité de 99,6 % [42].

Kermany et al. (2018) [42] ont utilisé l'architecture InceptionV3 pré-entrainée pour la détection automatique de l'OMD sur une base de données publique OCT2017 multi-classes contenant 84000 images. Le modèle a été évalué sur un jeu de test indépendant de 1 000 images (250 par classe) et a atteint une précision globale de 96,6 % dans une tâche de classification multi-classes (néovascularisation choroïdienne (CNV), OMD, drusen, normal). Pour le classifieur binaire (OMD, Normal), il a atteint une précision de 98,20 %, une sensibilité de 96,8 %, une spécificité de 99,6 %, et une AUC de 99,87 %, démontrant une excellente robustesse diagnostique et une capacité de généralisation.

Dans une étude menée par Tsuji et al. (2020) [43], des réseaux de capsules (CapsNet) ont été exploré pour la classification automatique des images OCT. Leur objectif était de différencier plusieurs pathologies rétiniennes, dont l'œdème maculaire diabétique. Pour ce faire, ils ont utilisé une la même base que Kermany et al. [43]. Le modèle CapsNet a atteint une précision spécifique de 99,2 % pour la détection de l'OMD.

Kamble et al. (2018) [44] ont développé un modèle basé sur Inception-ResNet-V2 pour l'analyse automatique de l'OMD à partir d'une base de données OCT locale composée au total d'environ 115 volumes, visant à entraîner et tester le modèle. Les résultats ont montré une précision élevée, atteignant jusqu'à 100 % sur certains ensembles de test.

Dans l'étude de Jaimes et al. (2025) [45], un modèle VGG16 modifié a été entraîné pour classifier les images de la base OCT2017 en quatre catégories : CNV, l'OMD, les druses et le cas sain. Après l'équilibrage de la base de données, 33 792 images ont été réparties équitablement entre les classes. Le classifieur a atteint une précision globale de 95,19 %, avec une précision moyenne de 95,29 et un F1-score moyen de 95,20 %. Pour la classe OMD spécifiquement, le modèle a obtenu une précision de 98,77 %, une sensibilité de 94,47 % et un F1-score de 96,57 %.

Li et al (2018**)** [46] ont développé une méthode entièrement automatisée pour détecter l'OMD et la dégénérescence maculaire liée à l'âge (DMLA), à partir des images OCT. Pour la classification binaire (normal vs OMD), l'approche repose sur l'affinage d'un modèle VGG-16 pré-entraîné.

Évaluée sur 500 images (250 par classe), la méthode a atteint une précision, une sensibilité et une spécificité de 98,8 %.

Dans le même contexte, Ajaz et al. (2021) [47] ont analysé les différentes méthodes utilisées pour la détection automatique de l'OMD en utilisant des images OCT. Ils ont comparé les méthodes classiques (SVM, forêts aléatoires) aux approches par Deep Learning (CNN). Ils soulignent la supériorité des modèles profonds en termes de précision, tout en notant le manque de bases de données publiques et standardisées, ce qui limite la généralisation des modèles.

Powroznik et al. (2024) [48] ont conçu un modèle hybride CNN-GRU pour la classification multiclasses des images OCT. Utilisant la base de données publique Labelled OCT Dataset. Le modèle a atteint une précision de 98,90 % pour la détection de l'OMD, surpassant plusieurs CNN standards.

Midena et al. (2023) [49] ont validé un algorithme d'IA destiné à l'analyse automatique des images OCT chez des patients atteints d'OMD. La méthode vise à détecter et quantifier plusieurs biomarqueurs tels que le Fluide Intra Rétinien (IRF), le fluide Sous-Rétinien (SRF), l'intégrité de la Membrane Limitante Externe (ELM), la Zone Ellipsoïde (EZ) et les Foyers Hyperréflectifs Rétiniens (HRF). L'étude a été menée sur 303 yeux montrant une forte concordance entre les résultats obtenus et les évaluations manuelles des experts. Cette quantification est rendue possible par des algorithmes basés sur les Réseaux Adversariaux Génératifs (GANs), offrant une précision de 94,7% à 95,7%. La précision de segmentation et d'analyse variait entre 94,7% et 100%.

Manikandan et al. (2023) [50] ont évalué l'efficacité des modèles de deep learning pour la détection de l'OMD à partir d'images OCT. Ils ont inclus 53 études regroupant un total de 1,4 million d'images OCT. Les modèles analysés étaient majoritairement basés sur des CNN montrant une aire sous la courbe ROC (AUC) moyenne de 0,9727, avec une sensibilité de 96 %.

Ces travaux positionnent l'OCT comme une modalité particulièrement fiable pour la détection automatisée de l'OMD.

2.2 Méthodes de détection basées sur la rétinographie

Li et al. (2022) [51] ont proposé un système de détection automatique de la RD et de l'OMD à partir d'images du fond d'œil, basé sur cinq modèles Inception améliorés. Leur méthode s'appuie

sur deux bases de données : une base locale composée de 7 916 images issues de 2 966 patients, et la base publique Messidor-2. Le processus de traitement inclut le recadrage des images, la suppression des bordures noires, des techniques d'augmentation (rotations, flips), ainsi qu'un fine-tuning à partir de poids pré entraînés sur ImageNet. Pour la base locale, le modèle atteint une précision de 97,2 % pour la classification NRDR/RDR (Non-Referable vs Referable DR) et 97,4 % pour NRDMO/RDMO (Non-Referable vs Referable DMO), avec une AUC de 0,994 pour l'OMD. Pour la base Messidor-2, le modèle conserve une excellente robustesse, avec une AUC de 0,948 pour l'OMD.

Dans leur revue systématique, Rodríguez-Miguel et al. (2021) ont passé en revue les avancées réalisées dans le domaine de la détection automatique de l'OMD, en particulier celles basées sur les techniques de deep learning appliquées à des images du fond d'œil, par exemple Singh et al. (2020), ont introduit une architecture HE-CNN (Hierarchical Ensemble of CNNs), obtenant une précision de 96,12 %. En outre, l'étude de Varadarajan et al. (2020) a montré qu'il était possible de prédire la présence d'un OMD centré uniquement à partir d'images rétinographiques, en utilisant un CNN entraîné sur des paires Retino/OCT, avec une AUC de 0,89 et une valeur prédictive positive de 61%. Ces approches ont pour avantage d'utiliser des images plus accessibles, notamment dans des contextes de dépistage de masse où les dispositifs OCT sont peu disponibles. Ainsi Khojasteh et al. (2019) ont comparé plusieurs modèles profonds, dont des CNN et un réseau ResNet-50 combiné à un SVM, atteignant une précision de 98 % et une sensibilité de 99 % sur les bases DIARETDB1 et e-Ophtha [52].

Guo et al. (2022) [53] ont proposé une méthode de détection automatique et de classification de l'OMD à partir d'images du fond d'œil, reposant sur un réseau de neurones profonds notamment YOLO (You Only Look Once) pour identifier en temps réel des exsudats durs. Afin d'optimiser la précision, l'approche intègre un prétraitement sur le canal vert, une fusion multi-échelle et une optimisation des ancres par K-means, ainsi qu'une augmentation par rotation, retournement et recadrage. Testé sur la base MESSIDOR, le modèle atteint une précision globale de 96 %, avec des taux de détection des exsudats variant selon leur probabilité : 100 % (forte), 79,1 % (moyenne) et 60,4 % (faible).

Fu et al. (2022) [54] ont développé une méthode de classification automatique de l'OMD basée sur l'architecture ResNet50 en utilisant des techniques de prétraitement, comme la correction

d'intensité et l'augmentation des données. Les résultats sont prometteurs, avec une précision de 97,06 % et une sensibilité de 88,64 % sur la base de données MESSIDOR.

Manikandan et al. (2023) [50] ont mené une méta-analyse portant sur l'efficacité des modèles d'apprentissage profond pour la détection de cette maladie à partir d’images du fond d’œil. L’étude, basée sur 53 travaux regroupant 472 328 images, a analysé plusieurs architectures CNN (InceptionV3, EfficientNet, VGG16, DenseNet) associées à des techniques de prétraitement. Les résultats montrent une sensibilité moyenne de 94 %, confirmant l’efficacité de ces modèles pour un dépistage fiable et non invasif.

Une étude a utilisé des modèles pré entraînés, tels que ResNet50, VGG16 et VGG19, pour évaluer la sévérité de la RD et le risque d'OMD en utilisant base de données IDRiD. Les images ont subi diverses techniques de prétraitement, notamment l'amélioration de la luminosité, de la couleur et du contraste, les perturbations de couleur et l'égalisation adaptative de l'histogramme limitée au contraste (CLAHE). ResNet50 a obtenu une précision de 60,2 % pour la classification de la sévérité de la RD et de 82,5 % pour l'évaluation du risque d'OMD [55].

Dans le même contexte, Yang et al. (2021) [56] ont montré que l’intégration de la rétinographie grand champ avec des modèles d’intelligence artificielle améliore la détection des anomalies rétiniennes telles que la RD et l’OMD, allant jusqu’à une sensibilité de 95 % avec les modèles comme ResNet, Inception-v3, VGG16.

Li et al. (2024) [57] ont participé au challenge MICCAI UWF4DR en développant un modèle d’apprentissage profond pour détecter automatiquement cette maladie à partir d’images rétiniennes ultra-grand champ (UWF). Ils ont utilisé plusieurs architectures CNN, dont EfficientNet-B0, ResNet-18 et une version améliorée ML-EfficientNet-B0, testées sur deux bases : la base UWF4DR et la base publique DeepDRiD. Le modèle ML-EfficientNet-B0 combiné à une stratégie Test-Time Augmentation (TTA) a obtenu des performances élevées, avec un AUROC de 0.9820.

Dans une étude récente, Bressler et al. (2025) [58] ont proposé un système autonome de dépistage de l’OMD en utilisant les réseaux de neurones convolutifs. Entraîné sur la base de données EyePACS contenant plus de 32 000 images issues de près de 16 000 patients, le modèle a été évalué à trois niveaux : image, œil et patient. Les résultats obtenus sont particulièrement solides, avec une AUC de 0,954 au niveau de l’image, 0,964 au niveau de l’œil, et 0,962 au niveau du

patient. Une validation externe sur la base IDRiD et d'autres ensembles d'images a permis de confirmer la robustesse et la capacité de généralisation de l'approche.

Varadarajan et al. (2018) [59] ont menés une étude pour diagnostiquer l'OMD en utilisant des CNN sur une base de données comprenant plus de 30 000 images couleur du fond d'œil. Pour la détection de l'OMD centré, le modèle a obtenu une AUC de 0,89. En ce qui concerne la détection du fluide intra rétinien, l'AUC a été de 0,81. Pour la détection du fluide sous-rétinien, le modèle a atteint une AUC de 0,88.

3. L'intelligence artificielle

L'intelligence artificielle (IA) est un domaine de l'informatique qui regroupe un ensemble de techniques ou d'algorithmes permettant à des machines d'imiter ou de simuler des comportements humains, comme apprendre, raisonner, percevoir, comprendre un langage ou même prendre des décisions. Elle repose sur des algorithmes capables d'analyser de grandes quantités de données, de s'adapter à de nouvelles situations et s'améliorer avec d'autres. Elle englobe l'apprentissage automatique (Machine Learning ou ML) et l'apprentissage profond (Deep Learning ou DL).

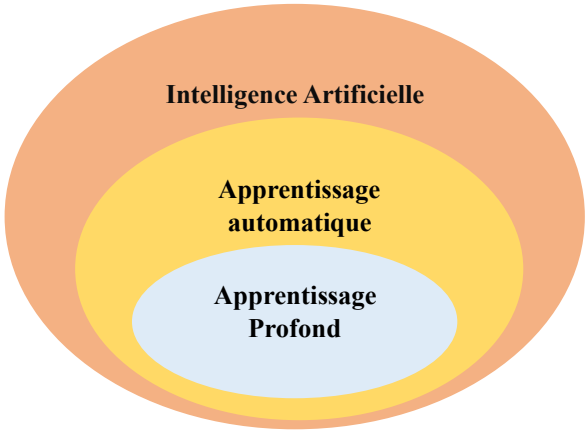


Figure II.1: La relation entre DL, ML et IA

4. L'apprentissage automatique (Machine Learning)

L'apprentissage automatique est une branche de l'intelligence artificielle [60] qui regroupe un ensemble de méthodes mathématiques permettant à un système informatique d'apprendre

automatiquement à effectuer des tâches spécifiques à partir d'un ensemble de données. Cet apprentissage repose généralement sur l'analyse d'un grand nombre de données pour reconnaître des motifs et prendre enfin des décisions.

L'apprentissage automatique se décline en trois grandes catégories : l'apprentissage supervisé, l'apprentissage non supervisé, et l'apprentissage semi-supervisé [61]. Chaque type est adapté à une nature particulière de données et intègre un ensemble d'algorithmes destinés à résoudre divers types de problèmes.

- **Apprentissage supervisé** : C'est la forme la plus courante et la plus accessible de l'apprentissage automatique. Elle s'applique lorsque les données sont étiquetées, c'est-à-dire que chaque exemple est associé à une sortie connue. L'algorithme apprend à associer les entrées à la sortie correspondante à partir d'un sous-ensemble de données d'entraînement. Une fois le modèle formé, il est testé sur de nouvelles données pour évaluer sa capacité à généraliser et à prédire correctement les classes.

- **Apprentissage non supervisé** : Ce type d'apprentissage est utilisé lorsque les données ne sont pas étiquetées. L'objectif est alors de découvrir des structures ou des regroupements sous-jacents dans les données selon des critères de similarité prédéfinis. Cette approche permet de regrouper automatiquement des échantillons en classes, souvent utilisée en segmentation d'images pour identifier des régions homogènes. Le succès de cette méthode dépend du choix judicieux du nombre de groupes et de la mesure de similarité utilisée.

- **Apprentissage semi-supervisé** : Ce type d'apprentissage se situe à mi-chemin entre l'apprentissage supervisé et non supervisé. Il est utilisé lorsque seule une partie des données est étiquetée, tandis que le reste ne l'est pas. L'algorithme exploite les exemples étiquetés pour guider l'analyse des exemples non étiquetés, ce qui permet d'améliorer la performance tout en réduisant le coût d'annotation manuelle. Cette méthode est particulièrement utile dans les domaines où l'obtention de données annotées est coûteuse ou difficile, comme en imagerie médicale. En combinant les avantages des deux approches précédentes, l'apprentissage semi-supervisé permet une meilleure généralisation du modèle tout en limitant les ressources nécessaires.

5. L'apprentissage profond (Deep Learning)

Dans les approches classiques d'apprentissage automatique, l'extraction des caractéristiques et la classification sont deux étapes distinctes. D'abord, on extrait manuellement les informations ou les caractéristiques jugées pertinentes à partir des données brutes. Ensuite, ces caractéristiques sont utilisées pour entraîner un modèle qui sera capable de faire des prédictions. Ce processus peut rapidement devenir lourd et compliqué, surtout lorsqu'on travaille avec de grandes quantités de données. Il nécessite non seulement beaucoup de temps, mais aussi une expertise importante pour choisir les bons attributs.

C'est justement pour surmonter ces limites que l'apprentissage profond (Deep Learning) s'est imposé. Grâce à ses architectures de réseaux de neurones profonds, il permet d'apprendre automatiquement les caractéristiques les plus utilesà partir des données brutes. Autrement dit, le système est capable de découvrir tout seul les bonnes représentations à utiliser, sans intervention humaine, à condition de disposer d'un volume suffisant de données étiquetées [62].

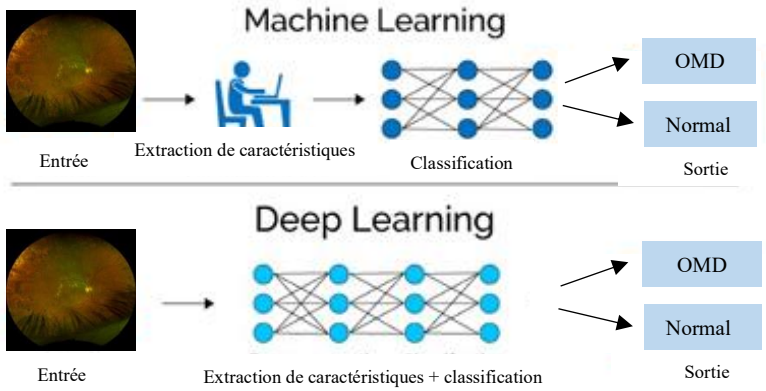


Figure II.2: La différence entre l'apprentissage automatique et l'apprentissage profond

Il existe plusieurs modèles de Deep Learning, Dans cette partie, nous allons nous focaliser sur des algorithmes les plus performants qui sont les réseaux de neurones convolutifs**.**

5.1 Les réseaux de neurones convolutifs

Les réseaux de neurones convolutifs (CNN) sont des modèles d'apprentissage profond spécialement conçus pour analyser les images. Le terme convolution indique que le réseau effectue plusieurs opérations mathématiques. Ils sont composés de plusieurs couches, dont les plus

importantes sont les couches de *convolution* et de *pooling*. La convolution applique des filtres pour extraire automatiquement des caractéristiques de l'image (comme les bords, les textures, ou les formes), tandis que le pooling réduit la taille des données pour simplifier le traitement.

Au fur et à mesure que l'image passe à travers les couches, le réseau apprend des informations de plus en plus complexes. Les premières couches du réseau analysent des éléments simples de l'image, tandis que les couches plus profondes combinent ces éléments pour construire des représentations plus complexes et abstraites.

Une fois les caractéristiques extraites par les couches de convolution, leur sortie est transformée en un vecteur, appelé vecteur de caractéristiques. Ce vecteur est ensuite transmis à une ou plusieurs couches entièrement connectées, chargées d'effectuer des tâches comme la classification ou la segmentation. La dernière couche, dite de classification, utilise une fonction d'activation pour donner la probabilité que l'image appartienne à chaque catégorie. L'ensemble du réseau est entraîné en ajustant les connexions entre les neurones grâce à un algorithme de rétropropagation, dans le but de réduire l'erreur entre les prédictions et les résultats attendus [63].

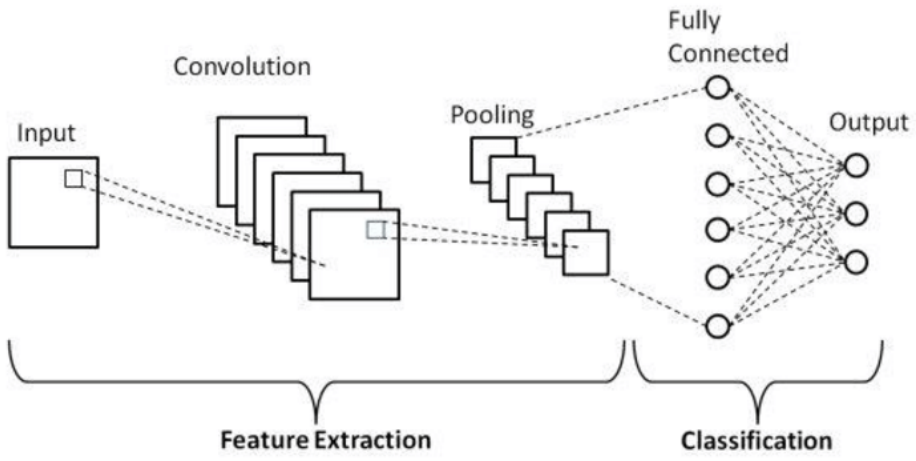


Figure II.3: Exemple d'architecture d'un CNN

Les CNN sont composées de plusieurs couches, chacune joue un rôle important dans le bon fonctionnement du réseau. Dans la partie qui suit, nous allons expliquer brièvement le rôle de chacune de ces couches.

5.2.1 Les couches d'un CNN

Les couches d'un CNN sont les éléments clés qui permettent au modèle de traiter efficacement les images et d'accomplir des tâches complexes. Chaque type de couche joue un rôle bien défini dans le processus global d'analyse et d'interprétation des données. On distingue principalement quatre types de couches fondamentales dans un CNN : la couche de convolution, la fonction d'activation, la couche de pooling et la couche entièrement connectée. Les paragraphes suivants détaillent le rôle et le fonctionnement de chacune de ces couches dans l'architecture d'un CNN.

1) Couche de convolution (CONV)

C'est la couche principale d'un CNN. Elle est utilisée pour extraire des caractéristiques à partir de l'image d'entrée en appliquant des filtres (ou noyaux) de petites tailles (généralement 3x3 ou 5x5). Ces filtres glissent sur l'image en effectuant des calculs (convolution) à chaque pas. Au départ, la fenêtre du filtre est placée en haut à gauche de l'image. Ensuite, elle se déplace progressivement vers la droite, selon un certain pas (nombre de cases). Lorsque la fenêtre atteint le bord de l'image, elle descend d'un pas et continue à se déplacer horizontalement, jusqu'à ce que le filtre ait parcouru toute l'image. À chaque déplacement, une convolution est effectuée, consistant à multiplier les valeurs des pixels de l'image par les valeurs du filtre, puis additionner les résultats pour générer une nouvelle valeur. Cette dernière est enregistrée dans une nouvelle image appelée carte de caractéristiques (Feature Map), qui représente des informations plus abstraites de l'image. Ces cartes de caractéristiques sont ensuite transmises à la couche suivante. L'objectif principal de cette opération est de détecter des éléments spécifiques dans l'image, tels que des lignes, des courbes ou des textures. Les filtres, qui sont des matrices de poids initialisées aléatoirement au départ, sont ensuite ajustés pendant l'entraînement pour mieux extraire les caractéristiques pertinentes et s'adapter aux données d'entraînement. La première couche de convolution capture des caractéristiques de bas niveau, tandis que les couches suivantes extraient des caractéristiques de plus en plus complexes [64] [65].

Pour chaque couche de convolution, plusieurs hyperparamètres doivent être définis au préalable : la taille et le nombre de filtres, le pas (ou "stride") et le remplissage avec des zéros (ou "zero padding"). La taille des filtres doit être inférieure à celle de l'image d'entrée. En règle générale, des filtres de petite taille répartis sur plusieurs couches de convolution donnent de meilleurs résultats que des filtres plus grands utilisés sur un nombre réduit de couches. Le pas de convolution indique le nombre de pixels par lesquels le filtre se déplace après chaque opération

(c'est-à-dire le décalage du filtre). Il permet d'éviter les chevauchements entre les différentes portions de l'image pendant le passage du filtre. En augmentant la taille du pas, la taille du volume de sortie diminue. Enfin, le remplissage à zéros sert à ajuster la taille des bords de l'image, de manière à conserver la même dimension après la convolution. Ces hyperparamètres doivent être soigneusement choisis, car certaines valeurs offrent de meilleures performances et un temps d'entraînement plus rapide que d'autres [64] [65].

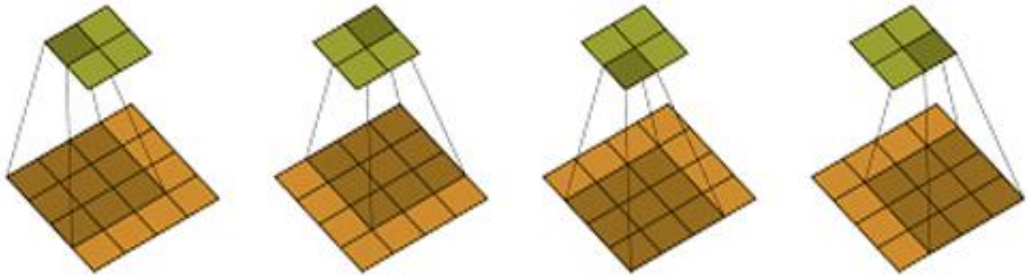


Figure II.4: Schéma du parcours de la fenêtre de filtre sur l'image

2) La fonction d'activation

Après chaque couche de convolution, une fonction d'activation est appliquée aux cartes de caractéristiques pour introduire de la non-linéarité dans le modèle. Cela permet de modéliser des relations complexes entre les entrées et les sorties. Parmi les fonctions d'activation les plus couramment utilisées sont la sigmoïde, la tangente hyperbolique (tanh), la ReLU, la Leaky ReLU et la Softmax [65].

- **Fonction sigmoïde :** cette fonction sert à transformer une valeur réelle en une valeur comprise entre 0 et 1, elle est utilisée pour normaliser la sortie d'un neurone et elle est de forme S, l'équation de cette fonction est définie par :

$$\text{Sigmoide}(x) = \frac{1}{1+e^{-x}} \qquad \textbf{(II-1)}$$

Où x est la valeur d'entrée. Pour des valeurs positives de *x*, la sortie tend vers 1, tandis qu'elle tend vers 0 pour des valeurs négatives.

- **Fonction tanh (tangente hyperbolique) :** Similaire à la sigmoïde, sauf que la fonction tanh possède une plage de sortie étendue de −1 à 1. Cette fonction est préférée dans le cas de centrage les données autour de zéro, ce qui peut accélérer l'apprentissage. Elle est définie par :

$$\tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \qquad \textbf{(II-2)}$$

- **Fonction ReLU (Rectified Linear Unit) :** c'est la fonction la plus simple et la plus utilisée dans les architectures de Deep Learning en raison de son efficacité, elle sert à remplacer les valeurs d'entrée négatives par zéro. Elle est définie par la fonction suivante :

$$\mathrm{ReLu}(x) = \max(0, x) \qquad \textbf{(II-3)}$$

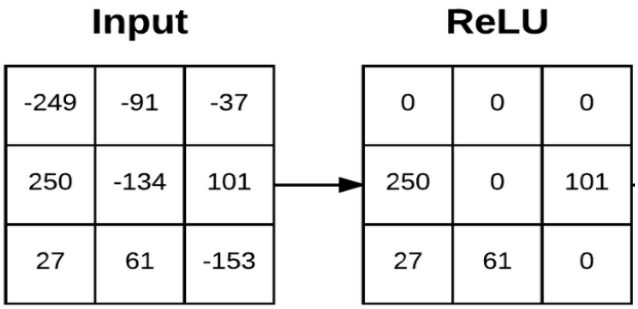


Figure II.5: Principe de fonctionnement de la fonction ReLU

Cependant cela peut entrainer une désactivation de neurones, ce qui réduit la capacité d'apprendre certaines caractéristiques. Pour résoudre ce problème, des variantes de la fonction ReLU ont été proposées tel que la fonction Leaky ReLU.

- **Fonction Leaky ReLU** : C'est une amélioration de la ReLU classique. Elle permet une petite pente pour les valeurs négatives afin d'éviter l'inactivation complète des neurones. Elle est définie par :

$$\text{Leaky ReLU}(x) = \begin{cases} 0.01x, & \text{si } x < 0 \\ x & \text{sinon} \end{cases} \qquad \textbf{(II-4)}$$

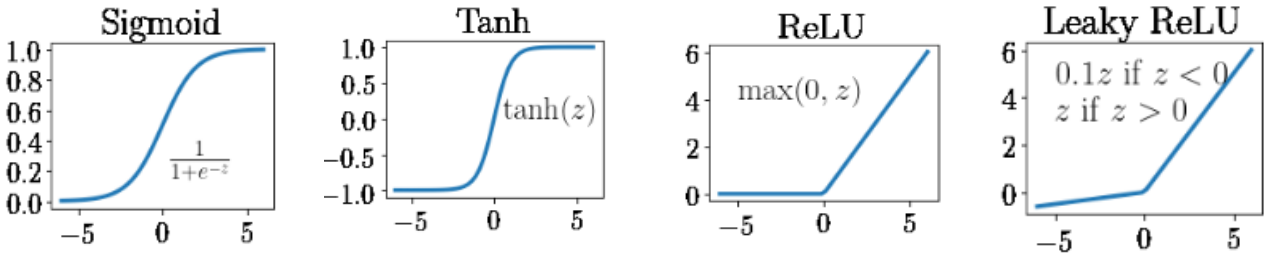


Figure II.6: Les courbes des fonctions d'activation

- **La fonction Softmax :** Elle principalement utilisée dans les couches de sortie des réseaux de neurones pour la classification multi-classes. Elle convertit un vecteur de valeurs (z) en une distribution de probabilités normalisées. Chaque probabilité est comprise entre 0 et 1, et la

somme de toutes les probabilités est égale à 1. La prédiction finale correspond à la classe ayant la probabilité la plus élevée. La fonction Softmax est définie par :

$$\text{Softmax}(z_i) = \frac{e^{zi}}{\sum_{j=1}^{k} e^{zj}} \qquad \textbf{(II-5)}$$

Où $\boldsymbol{z_i}$ est le score pour la classe *i, k* étant le nombre total de classes.

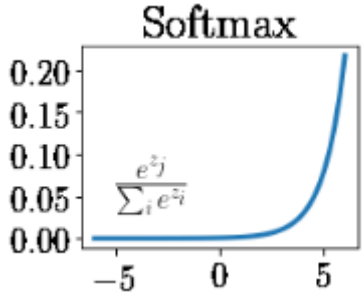


Figure II.7: La courbe de la fonction Softmax

3) La couche de pooling (POOL) :

Cette couche se trouve généralement entre deux couches de convolution. C'est une forme de sous-échantillonnage, elle permet de réduire la taille de la carte de caractéristiques en gardant ces informations les plus essentielles. Son rôle est d'examiner toutes les valeurs présentes dans une petite fenêtre et de réaliser un calcul statistique simple. Plusieurs méthodes existent, la plus utilisée c'est le Max Pooling qui consiste à conserver uniquement la valeur la plus élevée de la fenêtre et à ignorer les autres, et la deuxième méthode c'est l'Average Pooling, qui permet de conserver la valeur moyenne de chaque fenêtre.

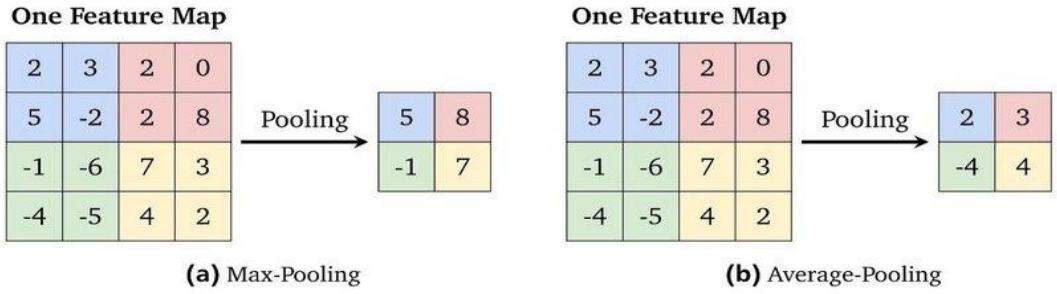


Figure II.8: Processus de max et Average Pooling

4) La couche entièrement connectée (Fully Connected Layer)

La couche entièrement connectée (Fully Connected layer) est généralement positionnée à la fin d'un CNN. Elle reçoit en entrée les caractéristiques extraites par les couches précédentes et les convertit en un vecteur de sortie destiné à des tâches spécifiques, telles que la classification ou la

régression. À l'image d'un réseau de neurones artificiel classique (ANN), chaque neurone de cette couche est connecté à l'ensemble des neurones de la couche précédente. Le nombre de neurones présents dépend de la nature de la tâche, notamment du nombre de classes à prédire. Cette couche est fréquemment suivie d'une fonction d'activation, pour une classification binaire, un seul neurone est utilisé, dans ce cas une fonction sigmoïde est employée. Elle convertit la sortie en une probabilité comprise entre 0 et 1, permettant ainsi de distinguer entre deux classes [64] [65].

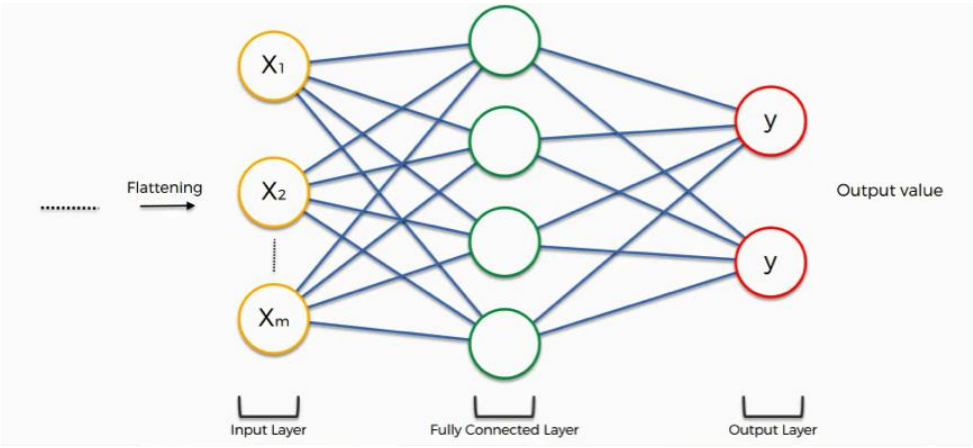


Figure II.9: Représentation de la couche entièrement connectée

5) La fonction de perte :

La fonction de perte, qu'on appelle aussi fonction d'erreur ou de coût, sert à mesurer la différence entre ce que le modèle prédit et la réponse attendue (étiquette). Elle donne un seul nombre qui représente l'erreur moyenne du modèle sur les données d'entraînement. Ce nombre est ensuite utilisé pour corriger et améliorer les performances du réseau, grâce à un algorithme d'optimisation. Selon le type de problème, différentes fonctions de perte peuvent être utilisées. Par exemple, pour les problèmes de régression, on utilise souvent la MSE (erreur quadratique moyenne). Pour les problèmes de classification, c'est généralement l'entropie croisée (Cross-Entropy Loss) qui est employée. Cette dernière fonctionne avec une fonction sigmoïde en sortie, ce qui permet de transformer les sorties du modèle en probabilités comprises entre 0 et 1, afin de choisir la classe la plus probable.

5.2.2 Algorithmes d'optimisation

Les algorithmes d'optimisation servent à ajuster les paramètres du modèle (comme les poids et les biais) pour réduire l'erreur et améliorer les prédictions. En gros, après chaque passage des données à travers le réseau, l'algorithme regarde combien l'erreur a diminué ou augmenté, puis il

ajuste les paramètres pour rendre les prochaines prédictions plus précises. Cela se fait grâce à une méthode appelée descente de gradient. L'idée est de calculer la pente de la fonction de coût (l'erreur) par rapport aux paramètres du modèle, et de mettre à jour les paramètres dans la direction où l'erreur est la plus forte, pour arriver au minimum global de l'erreur. Il y a aussi un paramètre important dans ce processus : le taux d'apprentissage. Ce taux détermine la taille du pas qu'on va faire à chaque ajustement. Si le pas est trop grand, on risque de passer à côté de la meilleure solution. S'il est trop petit, l'apprentissage devient trop lent [64].

Il existe différentes versions de la descente de gradient, comme le Gradient Descent classique qui ajuste les paramètres après avoir vu toutes les données, ou le Stochastic Gradient Descent (SGD) qui ajuste les paramètres après chaque donnée, rendant l'apprentissage plus rapide. Et des méthodes comme Adam ou RMSProp permettent d'optimiser l'apprentissage en ajustant le taux d'apprentissage automatiquement [64].

5.2.3 Les hyperparamètres en apprentissage profond

Les hyperparamètres sont des éléments essentiels dans l'apprentissage profond. Ils doivent être soigneusement définis avant le début de l'entraînement d'un modèle afin d'optimiser sa performance et d'obtenir les meilleurs résultats. Ces paramètres influencent la manière dont l'apprentissage progresse et la vitesse à laquelle le modèle converge [32]. Parmi les principaux hyperparamètres, nous citons :

- **Le taux d'apprentissage (learning rate) :** il détermine la taille du pas avec lequel les poids sont ajustés lors de la descente de gradient. Si ce taux est trop élevé, la convergence peut être rapide, mais le modèle risque d'être instable. À l'inverse, un taux trop bas peut entraîner une convergence trop lente.
- **Le nombre d'époques (epochs) :** il s'agit du nombre de fois où le modèle passe sur l'ensemble des données d'entraînement. Un nombre d'époques trop faible peut mener à un sous-apprentissage, tandis qu'un nombre trop élevé risque de provoquer un sur-apprentissage.
- **La taille du mini-lot (mini-batch size) :** c'est le nombre d'échantillons utilisés pour calculer une seule mise à jour des poids. Une taille de mini-lot trop petite peut rendre l'apprentissage instable, alors qu'une taille trop grande peut ralentir la convergence.
- **Le nombre de couches et la taille du réseau :** la structure du modèle, en particulier sa profondeur, joue un rôle crucial dans ses performances. Un modèle trop simple peut souffrir

de sous-apprentissage, tandis qu'un modèle trop complexe peut entraîner un sur-apprentissage.

- **Le choix de la fonction de coût** : la fonction de coût choisie a un impact sur la capacité du modèle à généraliser. Un mauvais choix peut limiter les performances du modèle.
- **Le choix de l'algorithme d'optimisation :** il détermine la méthode utilisée pour ajuster les poids du modèle au cours de l'entraînement.

5.2.4 Les types d'architectures de CNN pour la classification d'images

Durant ces dernières années plusieurs architectures de CNN ont été développés, et ont reconnu un très grand succès grâce à leurs performances dans divers domaines, notamment la classification des images médicales. Parmi ces architectures nous allons citer dans la partie qui suit : VGGNet (VGG16), ResNet50, EfficienNetB0, InceptionV3, MobileNetV2.

1) VGGNet :

VGGNet est une architecture de CNN qui a été développée en 2014 à l'université d'oxford par le Groupe de Géométrie Visuelle (Visual Geometry Group VGG) dans le but de faire la classification d'images. Ce modèle est connu par son architecture simple et puissante. Il se distingue par l'utilisation de petites convolutions (3x3), empilées en profondeur. Dans ses versions VGG-16 et VGG-19, le nombre de couches convolutionnelles atteint 16 et 19 respectivement. L'architecture comprend des couches convolutionnelles qui extraient les caractéristiques de l'image, suivies de couches de pooling (MaxPooling) qui réduisent la taille de l'image tout en conservant les informations importantes. Enfin, les couches fully connected (FC) se chargent de la classification des images en fonction des caractéristiques extraites par les couches précédentes. L'entrée du réseau est une image couleur (RVB) de taille 224x224x3.

Le VGG16 contient 16 couches en tout, 13 de convolutions et 3 entièrement connectées (FC), ces derniers sont distribués en 6 blocs, 5 blocs pour les convolutions et un seul pour les couches FC. Le premier bloc occupe 2 couches de convolutions avec des filtres de taille 3x3, un pas de 1 et une profondeur de 64 (nombres de filtres), le deuxième bloc est similaire au premier sauf que lui sa profondeur est de 128, le troisième bloc contient 3 couches de convolutions avec la même taille des filtres et de pas, mais avec une profondeur de 256, le quatrième et le cinquième bloc sont similaires au troisième, la différence réside dans la profondeur qui est de 512. Chaque couche de CONV est suivie par une fonction Relu et entre chaque bloc se retrouve une couche de MaxPooling de taille 2x2 et un pas de 2. Enfin le dernier bloc qui sera la sortie des blocs précédents, contient 3

couches FC avec un nombre de neurones de 4096, 4096 et 1000 respectivement, suivis par une fonction Softmax, ils sont utilisés et pour la classification [66]. Le schéma représentatif de VGG16 est illustré dans la figure suivante :

Figure II.10: Schéma de l'architecture d'un VGG16

2) ResNet50 :

ResNet-50 est une architecture de CNN qui a été développée en 2015 par Kaiming He et ses collègues de Microsoft Research dans le but d'améliorer la classification d'images sur des bases de données complexes comme ImageNet. Ce modèle est connu par sa capacité à être très profond tout en gardant une facilité d'entraînement grâce à l'introduction du concept innovant des connexions résiduelles. Contrairement aux architectures traditionnelles, ResNet ne cherche pas uniquement à empiler des couches, mais introduit des raccourcis (skip connections) qui permettent aux informations de passer directement d'une couche à une autre sans modification, ce qui aide à éviter le problème de dégradation de performance rencontré dans les réseaux très profonds. ResNet-50 contient 50 couches avec poids (convolutions et FC) réparties en plusieurs blocs organisés de manière structurée. L'architecture commence par une première couche de convolution avec un filtre de taille 7x7, un pas de 2 et une profondeur de 64, suivie immédiatement d'une couche de MaxPooling de taille 3x3 et un pas de 2 pour réduire rapidement la dimension de l'image d'entrée. Ensuite, le cœur du réseau est constitué de 4 ensembles de blocs résiduels, appelés bottleneck blocks, chacun d'eux étant composé de 3 couches de convolution successives avec des tailles de filtres 1x1, 3x3 et 1x1 [67].

Le premier ensemble contient 3 bottleneck blocks avec des profondeurs de 64, 64 et 256, respectivement. Le deuxième ensemble contient 4 bottleneck blocks avec des profondeurs de 128, 128 et 512. Le troisième ensemble, plus grand, contient 6 bottleneck blocks avec des profondeurs de 256, 256 et 1024. Enfin, le quatrième ensemble contient 3 bottleneck blocks avec des

profondeurs de 512, 512 et 2048. Chaque bloc utilise la fonction d'activation ReLU après chaque convolution, et à l'intérieur de chaque bottleneck, la première convolution 1x1 sert à réduire la dimensionnalité, la deuxième 3x3 traite l'information, et la troisième 1x1 restaure la dimension initiale. Le raccourci (connexion résiduelle) est ajouté à la sortie du bloc avant d'appliquer ReLU, ce qui permet au réseau de mieux propager les gradients pendant l'entraînement. À la fin de tous ces blocs, ResNet-50 utilise une couche de global average pooling pour réduire la taille des caractéristiques extraites, suivie d'une couche fully connected de 1000 neurones pour classer l'image parmi 1000 classes différentes grâce à une fonction d'activation Softmax. L'entrée du réseau est une image couleur (RVB) de taille 224x224x3, similaire aux autres architectures classiques comme VGGNet [67]. Le schéma représentatif de ResNet-50 est illustré dans la figure suivante :

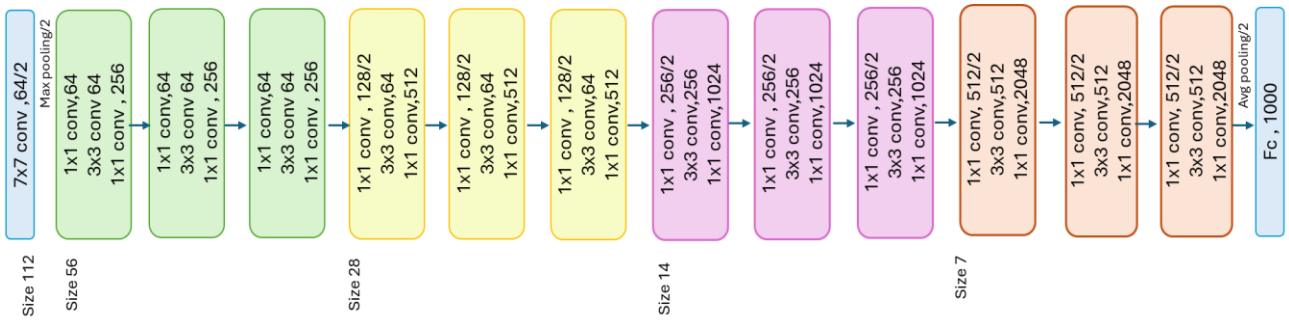


Figure II.11: Schéma de l'architecture d'un ResNet50

3) MobileNetV2 :

MobileNetV2 est une architecture de réseau de neurones convolutifs (CNN) développée par Google en 2018, L'architecture de MobileNetV2 commence par une couche de convolution standard avec 32 filtres de taille 3x3 et un pas de 2. Elle est suivie de 19 blocs résiduels inversés, chacun conçu pour optimiser le flux d'informations et réduire la complexité computationnelle [68]. Chaque bloc suit une séquence en trois étapes :

1. **Expansion** : une convolution 1x1 augmente la dimensionnalité des canaux d'entrée.
2. **Convolution en profondeur** : une convolution 3x3 est appliquée séparément sur chaque canal, réduisant ainsi le nombre de paramètres.
3. **Projection** : une autre convolution 1x1 réduit la dimensionnalité des canaux à leur taille d'origine.

Contrairement aux architectures résiduelles traditionnelles, les connexions résiduelles dans MobileNetV2 relient les couches de faible dimensionnalité, d'où le terme « inversé ».

Une particularité de MobileNetV2 est l'utilisation de la fonction d'activation ReLU6. De plus, les couches de projection utilisent des activations linéaires afin de préserver l'information dans les espaces de faible dimensionnalité, évitant ainsi la perte d'information que pourrait entraîner une activation non linéaire. L'entrée du réseau est une image couleur (RVB) de taille 224x224x3. Après les blocs résiduels inversés, une couche de convolution 1x1 avec 1280 filtres est appliquée, suivie d'une couche de pooling global pour réduire chaque carte de caractéristiques à une seule valeur. Enfin, une couche entièrement connectée avec une fonction d'activation Softmax est utilisée pour effectuer la classification finale [69]. La figure suivante illustre le schéma d'un MobileNetV2 :

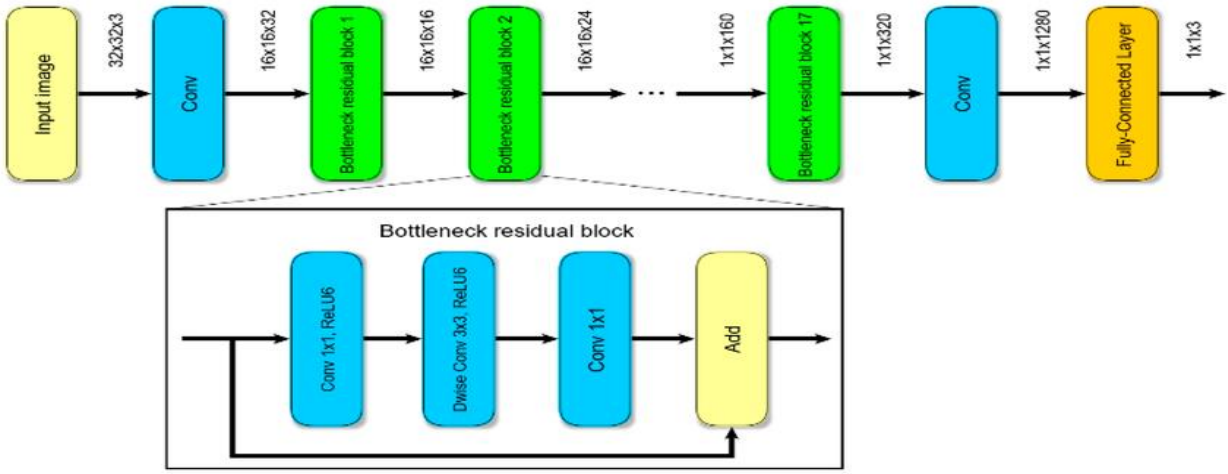


Figure II.12: Schéma de l'architecture d'un MobileNetV2

4) EfficientNetB0 :

EfficientNet-B0 est une architecture de réseau de neurones convolutifs (CNN) développée par Google AI en 2019, dans le but de réaliser une classification d'images avec une efficacité optimale. Ce modèle est reconnu pour sa capacité à offrir une haute précision tout en étant peu gourmand en ressources computationnelles. Il se distingue par l'utilisation d'une méthode de scaling composé qui équilibre la profondeur, la largeur et la résolution du réseau pour maximiser les performances tout en minimisant les coûts computationnels. L'architecture d'EfficientNet-B0 est construite autour de blocs appelés MBConv (Mobile Inverted Bottleneck Convolution), inspirés des blocs inversés de MobileNetV2. Chaque MBConv est conçu pour être léger et efficace, combinant des convolutions séparables en profondeur avec des mécanismes d'attention appelés Squeeze-and-Excitation (SE), qui permettent au réseau de se concentrer sur les caractéristiques les plus importantes [70] [71].

Le réseau commence par une couche de convolution standard de taille 3x3 avec 32 filtres et un pas de 2, suivie d'une normalisation par lot (Batch Normalization) et d'une fonction d'activation Swish. Cette couche initiale réduit la taille de l'image d'entrée tout en extrayant les premières caractéristiques [70]. Ensuite, le réseau est composé de plusieurs blocs MBConv .

Chaque bloc MBConv commence par une convolution 1x1 suivie d'une convolution depthwise pour extraire les caractéristiques spatiales, puis d'une convolution 1x1 pour réduire les canaux (projection). Le mécanisme « SE » est appliqué entre la convolution « depthwise" et la projection pour recalibrer les canaux en fonction de leur importance.

Après ces blocs, le réseau utilise une convolution 1x1 avec 1280 filtres, suivie d'une couche de pooling global pour réduire chaque carte de caractéristiques à une seule valeur, puis d'une couche entièrement connectée avec une fonction d'activation Softmax pour effectuer la classification finale. L'entrée du réseau est une image couleur (RVB) de taille 224x224x3.

EfficientNet-B0 est largement utilisé dans la littérature pour des tâches de classification d'images, offrant un excellent compromis entre précision et efficacité. Son architecture modulaire et optimisée le rend adapté à une variété d'applications [72].

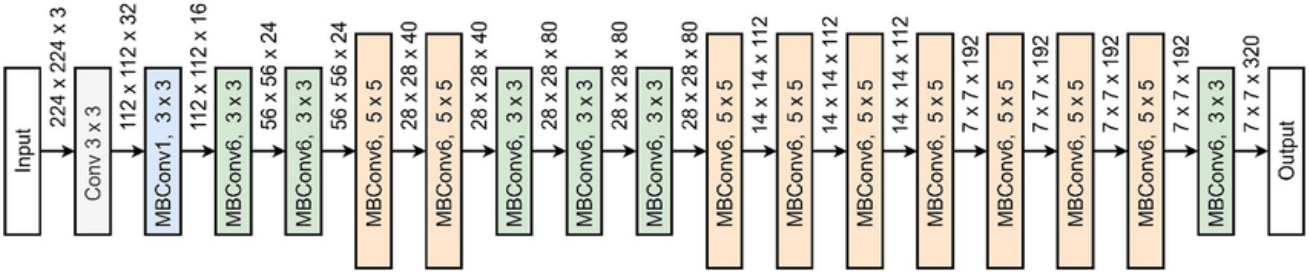


Figure II.13: Schéma d'un EfficientNetB0

5) InceptionV3 :

InceptionV3 est une architecture CNN développée par Google en 2015, dans le but d'améliorer la précision tout en optimisant l'efficacité computationnelle. Ce modèle est reconnu pour sa capacité à extraire des caractéristiques à différentes échelles grâce à l'utilisation de modules Inception, qui combinent plusieurs convolutions de tailles variées en parallèle. L'entrée du réseau est une image couleur (RVB) de taille 299x299x3, son architecture comprend plusieurs blocs [73]. Chacun ayant un rôle spécifique dans le traitement de l'image :

- **Bloc Stem** : Ce bloc initial effectue une série de convolutions et de pooling pour réduire la taille de l'image tout en augmentant la profondeur des caractéristiques.

- **Blocs Inception-A** : Ces blocs utilisent des convolutions factorisées, remplaçant par exemple une convolution 5x5 par deux convolutions 3x3, afin de réduire le nombre de paramètres tout en capturant des caractéristiques complexes.
- **Bloc de Réduction-A** : Ce bloc réduit la taille spatiale des cartes de caractéristiques, permettant une extraction plus profonde des informations.
- **Blocs Inception-B** : Ces blocs introduisent des convolutions asymétriques, telles que des convolutions 1x7 suivies de 7x1, pour capturer des motifs plus complexes tout en maintenant une efficacité computationnelle.
- **Bloc de Réduction-B** : Similaire au Bloc de Réduction-A, il diminue davantage la taille des cartes de caractéristiques, préparant les données pour les couches suivantes.
- **Blocs Inception-C** : Ces blocs combinent des convolutions 1x1 et 3x3 pour extraire des caractéristiques fines et spécifiques.
- **Classificateur auxiliaire** : Situé au milieu du réseau, ce classificateur aide à la rétropropagation du gradient et améliore la convergence pendant l'entraînement.
- **Bloc de sortie** : Après un pooling global, une couche entièrement connectée avec une fonction d'activation Softmax est utilisée pour la classification finale [74].

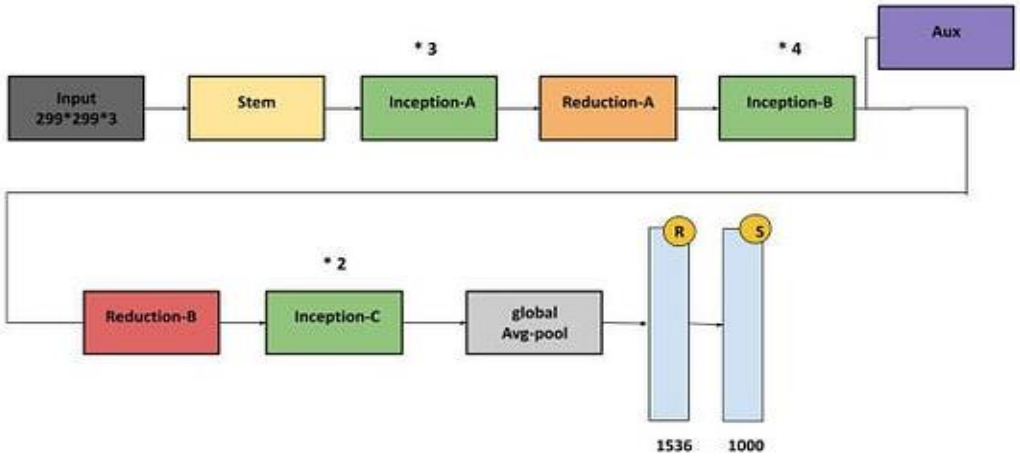


Figure II.14: Architecture d'un InceptionV3

6. L'apprentissage par transfert

L'apprentissage par transfert (ou transfer learning) est une technique qui consiste à réutiliser les connaissances acquises par un modèle préalablement entraîné sur une grande quantité de données

pour en faciliter l'entraînement sur une nouvelle tâche similaire dans le but de résoudre un problème ou effectué une tache spécifique au lieu de créer un nouveau modèle à partir de zéro. Cette approche est particulièrement efficace dans le domaine de l'apprentissage profond (deep learning), où les modèles nécessitent généralement de grandes quantités de données et des ressources computationnelles importantes. En réutilisant un modèle pré-entraîné comme point de départ, on peut obtenir de bonnes performances même avec un nombre limité de données ce qui permet aussi un gain de temps [75].

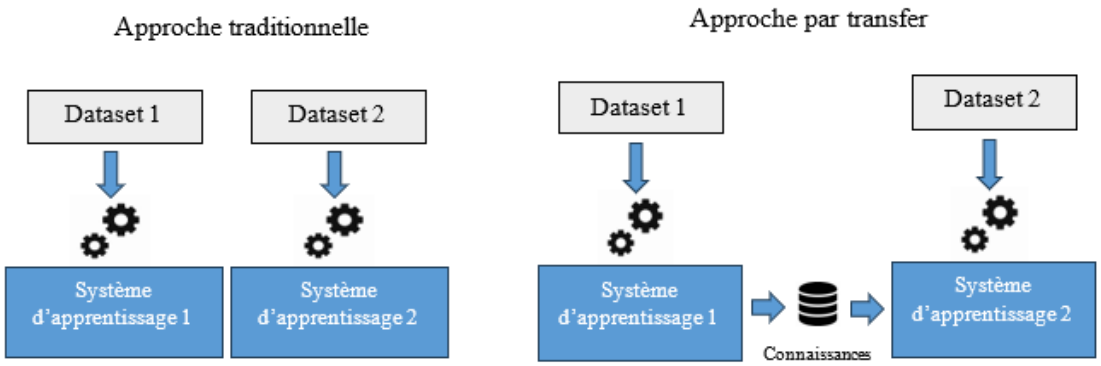


Figure II.15: Illustration du Transfert Learning

Dans le cadre de l'apprentissage par transfert, plusieurs stratégies se distinguent en fonction de ce que l'on souhaite transférer, du moment où le transfert est effectué et de la manière dont il est réalisé. Les deux principales techniques sont :

1. **Extraction de caractéristiques (Feature Extraction)** : Cette approche consiste à utiliser les représentations apprises par un modèle pré-entraîné pour extraire des caractéristiques significatives d'un nouvel échantillon. Un nouveau classificateur est ajouté et entraîné à partir de zéro, en réutilisant les caractéristiques apprises. Les paramètres du modèle pré-entraîné restent figés pendant l'entraînement.

2. **Ajustement fin (Fine-Tuning)** : Lors de l'ajustement fin, les paramètres du modèle pré-entraîné sont affinés sur le jeu de données de la tâche cible, ainsi que les couches supplémentaires. Les couches supérieures (spécifiques à la tâche) sont généralement réentraînées, tandis que les couches inférieures (plus générales) peuvent être fixées ou ajustées légèrement.

Ces techniques permettent d'exploiter les connaissances d'un modèle pré-entraîné pour résoudre efficacement de nouvelles tâches.

Adapter les poids d'un modèle pré-entraîné, à une nouvelle tâche peut améliorer sa spécialisation c'est-à-dire que le modèle devient spécialisé aux données d'entrainement (il apprend beaucoup). Cependant, si le jeu de données est trop restreint, le nombre élevé de paramètres du modèle peut entraîner un sur-apprentissage (*overfitting*) [76].

7. Surapprentissage et sous-apprentissage

Le sur-apprentissage (*overfitting*) survient lorsque le modèle devient trop spécifique aux données d'entraînement, apprenant non seulement les tendances générales mais aussi les détails et le bruit propres à cet ensemble. Cela réduit sa capacité à généraliser et à faire des prédictions précises sur de nouvelles données. Il peut également se manifester lorsqu'un modèle est formé uniquement sur un groupe particulier, entraînant des prédictions biaisées ou contradictoires pour des groupes sous-représentés.

À l'inverse, le sous-apprentissage (*underfitting*) se produit lorsqu'un modèle est trop simple pour capturer les relations complexes entre les variables d'entrée et de sortie. Cela se traduit par une erreur élevée tant sur les données d'entraînement que sur les données de validation, indiquant que le modèle n'a pas appris les structures pertinentes dans les données. Ce phénomène peut résulter d'un modèle insuffisamment puissant, d'une régularisation excessive ou d'un entraînement insuffisant [77] .

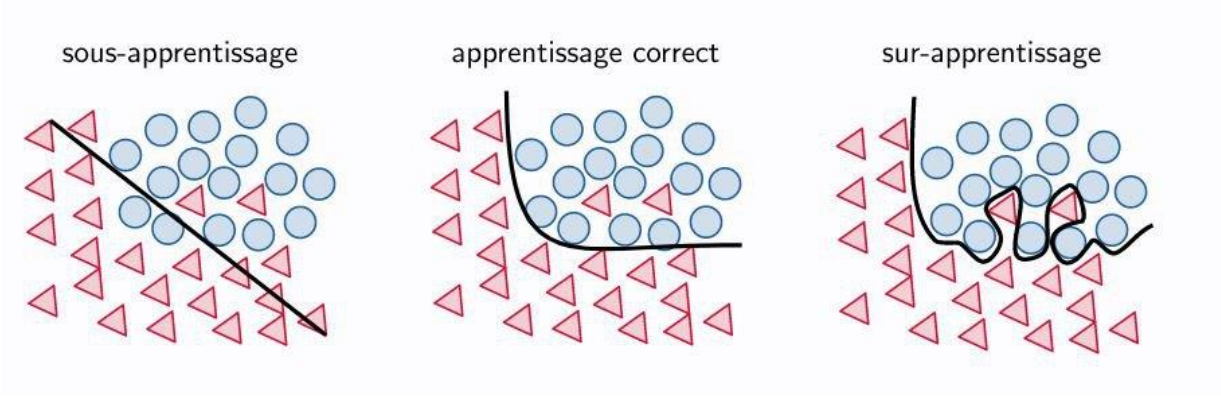


Figure II.16: Sous-apprentissage et sur-apprantissage

Pour prévenir le surajustement (*overfitting*), il est essentiel de disposer d'un ensemble de données d'entraînement suffisamment représentatif et diversifié. En effet, lorsque le nombre d'exemples est limité, le modèle risque d'apprendre non seulement les motifs pertinents, mais aussi les bruits ou détails spécifiques aux données d'entraînement, ce qui peut nuire à ses performances

sur de nouvelles données (tests ou validations) [78]. Plusieurs techniques peuvent être mises en œuvre pour atténuer ce phénomène lors de l'apprentissage :

- **Augmentation des données (Data Augmentation)** : Générer de nouvelles données à partir des existantes en appliquant des transformations (rotations, zooms, etc.) pour enrichir l'ensemble d'entraînement.
- **Utilisation du Dropout** : Désactiver aléatoirement certains neurones du réseau pendant l'entraînement pour éviter une trop forte dépendance à certaines caractéristiques.
- **Ajustement des hyperparamètres** : Modifier des paramètres tels que le nombre de couches, le type de fonction d'activation, l'optimiseur (SGD, Adam, RMSprop, etc.), le nombre d'époques, la taille des lots (*batch size*), afin d'optimiser la performance du modèle.
- **Validation croisée (K-Fold Cross Validation)** : Diviser les données en *k* sous-ensembles et entraîner le modèle *k* fois, chaque fois en utilisant un sous-ensemble différent pour la validation, afin d'évaluer la capacité de généralisation du modèle.
- **Arrêt précoce (Early Stopping)** : Surveiller la performance du modèle sur un ensemble de validation et interrompre l'entraînement lorsque la performance cesse de s'améliorer, pour éviter que le modèle n'apprenne le bruit des données d'entraînement.

8. Les métriques d'évaluation

Dans le domaine de l'intelligence artificielle, évaluer les performances des algorithmes est une étape cruciale pour assurer leur efficacité et leur fiabilité. Les métriques d'évaluation sont des outils quantitatifs essentiels pour mesurer la performance d'un modèle ou d'un algorithme de l'intelligence artificielle. Elles permettent de comparer différentes approches ou configurations, notamment dans le cadre des réseaux de neurones convolutifs (CNN) [32]. Ces mesures sont obtenues en confrontant les prédictions du modèle aux valeurs réelles fournies par des experts, ce qui permet d'évaluer la pertinence des résultats et l'efficacité du modèle pour résoudre un problème spécifique. Parmi lesquelles nous citons :

Un outil central dans cette évaluation est la matrice de confusion, une structure carrée qui résume les performances du modèle, elle présente une vue générale des erreurs du modèle lors de la partie du test.

		Classes prédites	
		Positive	Négative
Classes Réelles	Positive	VP	FN
	Négative	FP	VN

Figure II.17 : Matrice de convolution

- **Vrais Positifs (VP)** : cas correctement identifiés comme positifs ;
- **Faux Positifs (FP)** : cas incorrectement identifiés comme positifs ;
- **Vrais Négatifs (VN)** : cas correctement identifiés comme négatifs ;
- **Faux Négatifs (FN)** : cas incorrectement identifiés comme négatifs.

À partir de ces quatre éléments, plusieurs métriques d'évaluation peuvent être calculées [32] :

- **Exactitude (Accuracy)** : proportion de prédictions correctes sur l'ensemble des cas.
- **Sensibilité (Recall ou Taux de Vrais Positifs)** : capacité du modèle à identifier correctement les cas positifs.
- **Spécificité** : capacité du modèle à identifier correctement les cas négatifs.
- **Précision** : proportion de cas réellement positifs parmi ceux prédits comme positifs.
- **F-mesure (F1 Score)** : moyenne harmonique de la précision et du rappel, utile pour évaluer l'équilibre entre ces deux mesures, notamment en cas de classes déséquilibrées.
- **Indice de Jaccard (IoU)** : mesure de similarité entre les ensembles prédits et réels, souvent utilisée en segmentation d'images.
- **Courbe ROC (Receiver Operating Characteristic)** : représentation graphique du taux de vrais positifs en fonction du taux de faux positifs, permettant d'évaluer la performance du modèle à différents seuils.
- **AUC (Area Under the Curve)** : aire sous la courbe ROC, indiquant la capacité globale du modèle à distinguer entre les classes.

Il est crucial de choisir les métriques d'évaluation appropriées en fonction du problème étudié et du type de modèle utilisé, afin d'obtenir une évaluation fiable et précise de la performance du modèle.

9. Conclusion

Ce chapitre a présenté une revue approfondie des avancées récentes en intelligence artificielle appliquée au diagnostic de l'œdème maculaire diabétique à partir d'images rétiniennes, notamment la tomographie par cohérence optique (OCT) et la rétinographie. Les travaux analysés mettent en lumière l'efficacité des modèles de deep learning. Ces approches surpassent les méthodes traditionnelles en termes de sensibilité, spécificité et rapidité, tout en automatisant l'extraction de caractéristiques complexes (exsudats, fluides intra-rétiniens).

Les prochaines étapes de cette étude consisteront à collecter une base de données multimodales et à proposer une méthodologie innovante, s'appuyant sur les forces identifiées dans cet état de l'art, pour développer un modèle performant et explicable dédié à la détection de cette maladie.

III. Chapitre III : Bases de données

1. Introduction

Les bases de données rassemblent généralement un ensemble d'informations issues de diverses sources, souvent variées et réparties sous différents formats. Elles fournissent une vue d'ensemble des données, permettant aux analystes et aux experts de différents domaines de les exploiter efficacement. Dans le domaine médical, les bases de données jouent un rôle très important, car elles permettent de faciliter l'analyse des informations cliniques, le suivi des patients et l'aide au diagnostic. Elles peuvent aussi servir pour mener des recherches statistiques ou pour améliorer les méthodes et pratiques médicales [79].

La création d'une base de données ne se résume pas simplement à rassembler des informations : c'est un processus qui suit plusieurs étapes essentielles. D'abord, il faut identifier les sources de données les plus pertinentes et bien comprendre les besoins des utilisateurs qui vont les exploiter. Ensuite, ces données doivent être organisées de manière claire et structurée, en créant des sous-catégories adaptées à chaque type d'information. Enfin, il est indispensable de mettre en place des outils permettant d'interroger et d'analyser efficacement ces données. Chaque étape comporte ses propres défis, comme la nécessité de standardiser les informations, de garantir la confidentialité des données médicales et de gérer la complexité des structures utilisées [79].

Dans le cadre de ce projet de fin d'étude, nous nous concentrons sur la collecte, la sélection et l'organisation des données au sein d'une base multimodale dédiée à la détection de l'œdème maculaire diabétique. Cette base regroupe à la fois des images OCT et des rétinographies grand champ, dans le but d'appliquer des techniques de traitement d'images et d'approches basées sur le Deep Learning afin d'aider les médecins dans la détection de l'OMD.

Ce chapitre détaille la conception de cette base de données, son organisation, ainsi que les défis rencontrés lors de sa création.

2. Définition des bases de données

Une base de données (BDD) est un système où on peut stocker des informations de manière organisée et structurée, tout en minimisant la redondance des données. Elle est conçue pour que ces informations puissent être facilement consultées, modifiées ou mises à jour par des utilisateurs, comme les professionnels de santé. Pour faciliter le partage et l'accès à ces données. Elles sont souvent regroupées en catégories ou en tableaux qui sont liés entre eux, ce qui permet de faire des

recherches rapides et d'obtenir des rapports précis. Les bases de données sont souvent reliées à des réseaux, d'où l'appellation "base". En général, une base de données fait partie d'un système d'information plus vaste, qui regroupe l'ensemble des ressources nécessaires pour gérer et partager des données de manière efficace. Cela inclut à la fois le stockage des informations et les outils permettant leur exploitation. Les bases de données offrent aux utilisateurs la possibilité de consulter, de saisir ou de mettre à jour les informations, tout en contrôlant les droits d'accès pour garantir la sécurité des données.

Il existe deux types de bases de données : les bases de données locales et les bases de données distribuées.

- Une base de données locale est accessible sur une seule machine par un utilisateur présent sur place.
- Base de données distribuée stocke des informations sur plusieurs machines distantes, accessibles via un réseau, ce qui permet à plusieurs utilisateurs de travailler simultanément sur les mêmes données.

Dans le domaine médical, les bases de données servent à stocker des informations sur les patients, des images médicales, des résultats d'examens et des données statistiques. Elles facilitent le suivi des patients, aident au diagnostic et soutiennent la recherche médicale en fournissant des informations fiables et centralisées.

Rappelons que dans ce projet, notre objectif est de développer un outil intelligent capable d'aider les médecins dans la détection de l'œdème maculaire diabétique (OMD). Cet outil sera validé et évalué à partir de la base de données (BDD) que nous avons construite.

3. Description de la base de données collectée

Notre base de données a été collecté dans *la Clinique Lazouni d'ophtalmologie* sise à Tlemcen en collaboration avec plusieurs médecins ophtalmologues, où nous avons pu obtenir des images médicales de patients atteints d'OMD dans le cadre du suivi de ces patients pendants leurs consultations. Ainsi que des images de patients sains qui ne représentent aucun signe pathologique.

L'objectif était d'avoir une base de données suffisamment équilibrée entre cas pathologiques et cas normaux, tout en tenant compte des variations inter patients afin d'améliorer la robustesse du modèle.

Les données collectées proviennent de deux modalités d'imagerie :

- Les images OCT (Optical Coherence Tomography), obtenues à l'aide de deux types d'appareils de marque « Solix » et « RTvue ».
- Les images de rétinographies eux-mêmes provenant de deux types de rétinographes : « Eidon » et « Optos ».

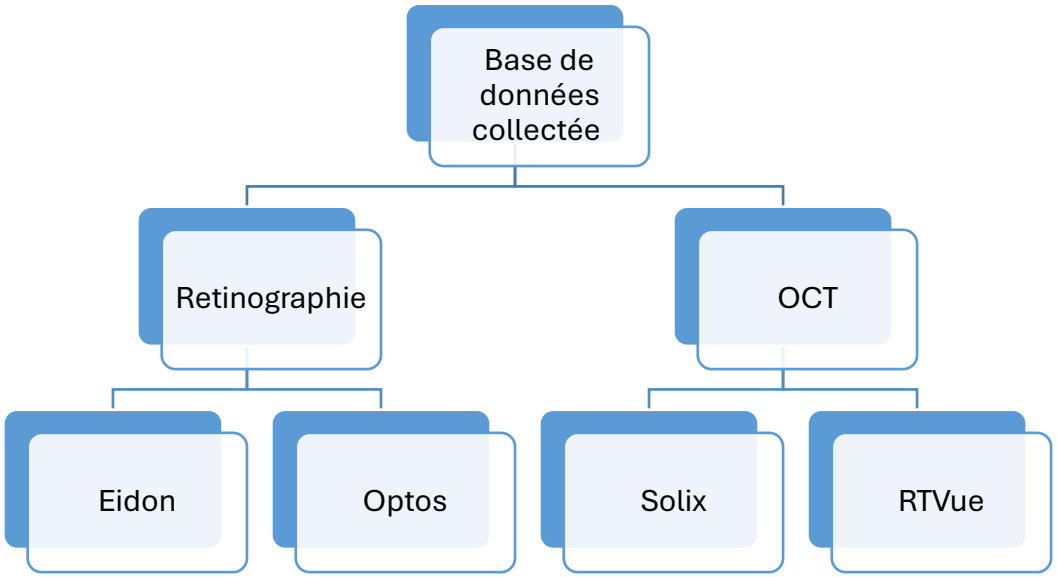


Figure III.1: Organisation de la base de données collectée selon le type d'imagerie et le dispositif utilisé

La collecte des données s'est déroulée sur une période d'environ trois mois, ce qui a permis de garantir la cohérence et la qualité des informations recueillies. Au total, 14 239 images ont été collectées, réparties entre deux groupes (patients atteints d'OMD et cas normaux), provenant de 700 patients en total. Le diagramme ci-dessous (Figure III.2) représente en détails le nombre d'images provenant de chaque appareil.

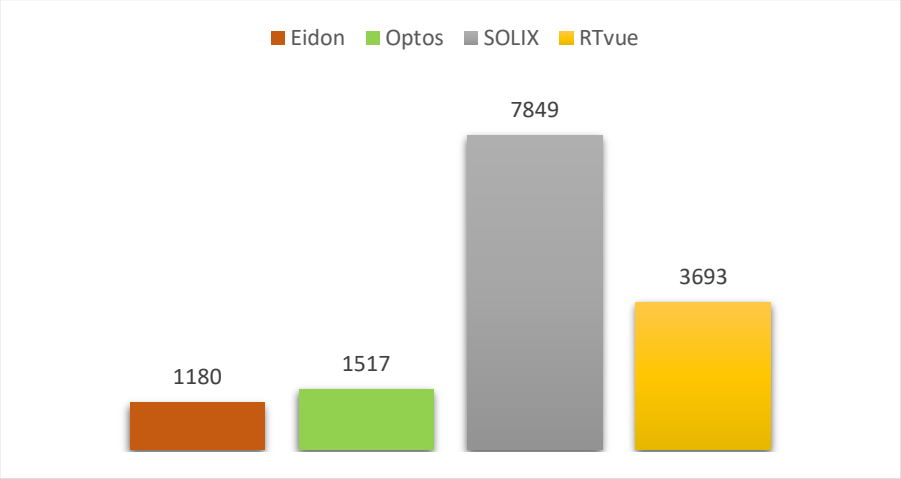


Figure III.2: Répartition du nombre d'images selon le type d'appareil d'acquisition

Le tableau suivant détaille la distribution des images selon les modalités utilisées et les cas (OMD et NORMAL) :

Tableau III-1: Distribution des images collectées selon la modalité et selon les cas

Type d'images	OMD	NORMAL	Total
OCT	8 943	2 599	11 542
Rétinographie	1 697	1 000	2 697

Comme le montre le Tableau III-1 ci-dessus, il y a un grand nombre d'images OCT pour les cas d'OMD (8 943 images), mais beaucoup moins pour les cas normaux (2 599 images). Cela s'explique par le fait que l'examen OCT est principalement utilisé pour diagnostiquer et suivre des pathologies rétiniennes comme l'OMD. Par contre les patients sans problèmes de rétine, n'ont pas nécessairement besoin de passer cet examen, ce qui explique pourquoi il y a moins d'images OCT pour les cas normaux. En d'autres termes, l'OCT est surtout utilisé pour les patients ayant des symptômes ou des risques de maladies rétiniennes. Les patients sans symptômes évidents ne sont généralement pas soumis à cet examen, ce qui fait que le nombre d'images pour les cas normaux est plus faible. Afin d'éviter tout biais de classification, nous avons opté pour une augmentation des données sur les images normales. Cette approche permet de rééquilibrer le jeu de données et d'améliorer la capacité de généralisation du modèle (voir le Chapitre 4).

D'un autre côté, on constate que les images OCT représentent la majorité de la base de données. Cela s'explique par le fait que chaque examen OCT comprend plusieurs coupes par œil, le solix effectue 12 coupes pour chaque œil et le RTvue effectue 10 coupes par oeil, ce qui permet une

analyse plus détaillée et précise des différentes couches de la rétine. Cette base constitue un atout majeur pour l'entraînement et l'évaluation de systèmes d'analyse assistée par ordinateur, notamment dans le contexte du dépistage précoce et du suivi de l'OMD.

En ce qui concerne les modalités d'imagerie. Les figures qui suivent illustrent la proportion de patients atteints d'OMD et NORMAUX disposant d'images OCT, de rétinographies ou des deux.

La figure III.3 qui suit illustre le diagramme de la répartition des images selon les modalités d'imagerie chez les patients atteints d'OMD.

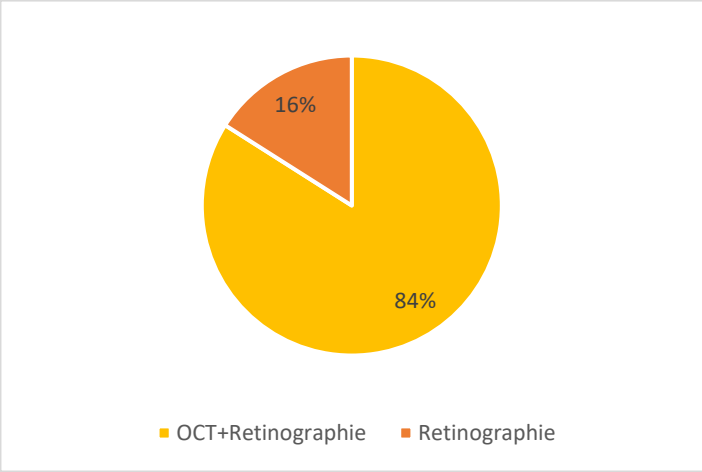


Figure III.3: Répartition des modalités d'imagerie disponibles chez les patients atteints d'OMD

La figure ci-dessus montre que 84% des patients disposent à la fois d'images OCT et de rétinographies, tandis que 16% ne possèdent que des rétinographies. Cependant, cette répartition pourrait être influencée par un manque de données pour certains patients qui pourraient limiter la représentativité des résultats. Cela constitue l'un des inconvénients majeurs d'une collecte de données.

De même la figure qui suit illustre le diagramme de la répartition des images selon les modalités d'imagerie chez les patients normaux.

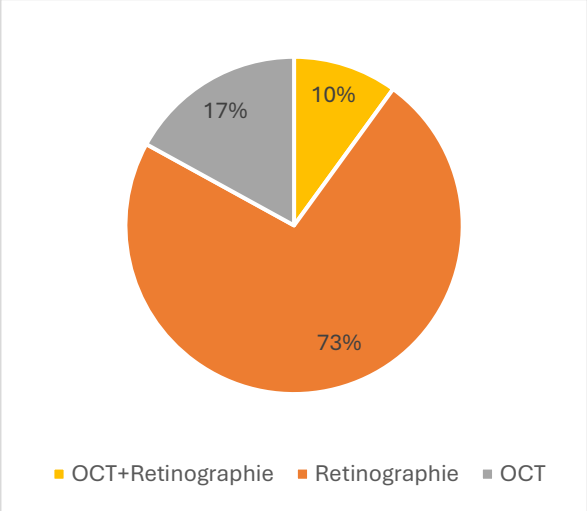


Figure III.4: Répartition des modalités d'imagerie disponibles chez les patients NORMAUX

La figure illustre la répartition des types d'imagerie disponibles chez les patients normaux, en distinguant trois modalités : la rétinographie seule, l'OCT seule, et la combinaison des deux examens. On observe que la majorité des cas, soit 73 %, disposent uniquement d'une rétinographie, tandis que 17 % ont bénéficié d'un OCT seul, et seulement 10 % d'une double modalité (OCT + rétinographie).

Cette répartition met en évidence une réalité clinique importante : chez les patients ne présentant pas de signes pathologiques, la rétinographie constitue généralement un examen suffisant. Elle permet un contrôle visuel de la rétine à moindre coût, sans recourir à des techniques plus complexes. Largement utilisée dans les bilans de routine ou les campagnes de dépistage, elle s'adapte bien au profil des patients normaux qui ne nécessitent pas d'exploration approfondie.

En revanche, l'OCT, qui offre une visualisation fine des couches rétiniennes, est une technique plus ciblée, généralement réservée à la détection ou au suivi de pathologies spécifiques comme l'OMD. Sa faible utilisation chez les patients normaux (seuls ou en complément de la rétinographie) témoigne donc d'un usage rationnel des ressources d'imagerie, adapté au contexte clinique.

Afin de garantir une analyse précise et représentative de nos données collectés, ces derniers ont été réparties en deux groupes distincts : les images OCT, et les images rétinographiques que nous allons détaillés par la suite.

3.1 Première base : Données de l'OCT

Notre base de données comprend des images OCT acquises à l'aide de deux appareils le premier est de la marque « RTvue », repose sur la technologie Spectral-Domain OCT(SD-OCT), et le deuxième de la marque « SOLIX » d'une technologie plus avancée Swept-Source OCT (SS-OCT), deux systèmes d'OCT de dernière génération, qui permettent d'obtenir des images en coupe très détaillées de la rétine. Au total, nous avons collecté les données de 181 patients tous possèdent des images provenant de SOLIX et RTvue. Parmi eux, des patients sont atteints d'OMD, tandis que d'autres présentent une rétine normale sans pathologie apparente. Figure III.5 suivante illustre la répartition des patients selon les cas (NORMAL, OMD).

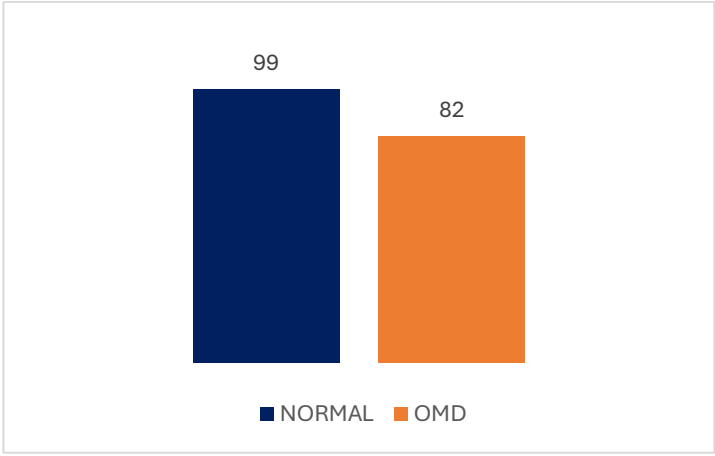


Figure III.5: Effectif des patients collectés de l'appareil OCT : NORMAL et OMD

Les images obtenues avec les appareils SOLIX et RTVue ont été enregistrées au format « .jpg », un format qui conserve les détails fins de l'image sans perte d'informations. Ce choix technique garantit une qualité optimale pour les traitements d'images et les analyses à venir, en préservant la finesse des structures visibles.

Les images issues du SOLIX présentent une résolution de 1536 × 1024 pixels, correspondant à une capacité d'imagerie en profondeur d'environ 6 mm au niveau de la rétine, avec une résolution axiale d'environ 5 µm. Cette définition permet une observation précise des structures rétiniennes comme la fovéa, les couches internes, ainsi que les signes caractéristiques de l'OMD, tels que les zones d'épaississement ou les kystes intrarétiniens. Le SOLIX offre des acquisitions ultra-rapides et des coupes rétiniennes très détaillées, ce qui en fait un outil de choix pour l'analyse fine de la microstructure rétinienne.

Le RTVue quant à lui, permet l'acquisition rapide d'images transversales de la rétine avec une résolution d'image d'environ 512 × 1024 pixels par coupe, avec une résolution axiale d'environ 5 µm, facilitant la détection précoce de pathologies rétiniennes comme l'OMD. Bien qu'il soit antérieur au SOLIX en termes de génération, le RTVue reste un appareil performant, apprécié pour sa fiabilité, sa précision et sa capacité à visualiser les micro-détails tissulaires.

Ainsi, grâce à leur complémentarité, le SOLIX et le RTVue ont permis de constituer une base d'images riche et diversifiée. Ces deux appareils offrent une excellente qualité d'imagerie rétinienne, indispensable pour une étude approfondie de l'OMD. La figure ci-dessous illustre un exemple des images obtenues par ces systèmes d'acquisition. La Figure III.6 ci-dessous illustre un exemple des images prise par ces deux appareils.

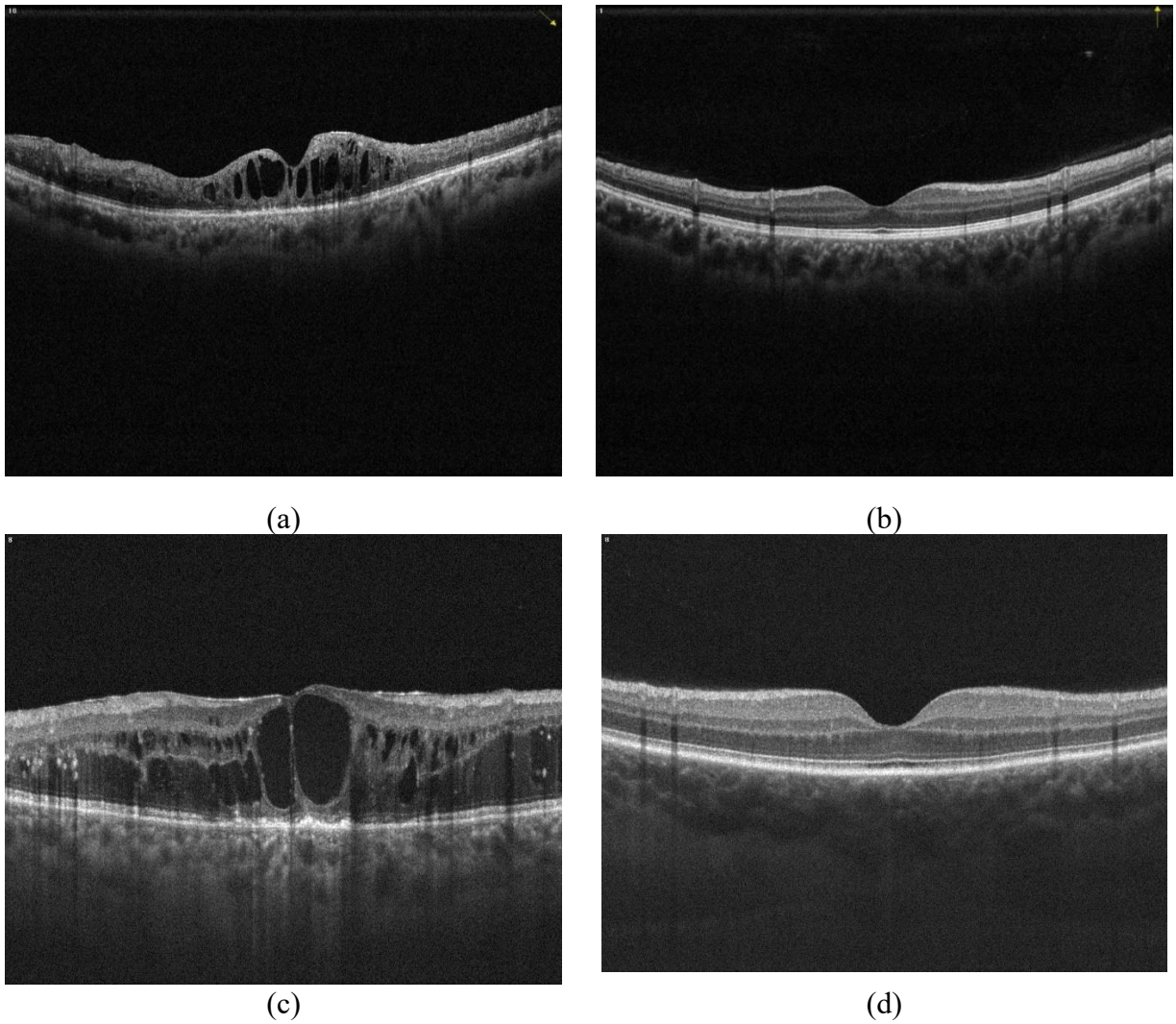


Figure III.6: Exemples d'images OCT illustrant les deux classes considérées. (a) Image OCT d'un patient atteint d'OMD prise avec par SOLIX;(b) Image d'une reine saine prise par SOLIX ;(c) Image OCT d'un patient atteint d'OMD prise avec par RTvue;(d) Image d'une reine saine prise par RTvue

3.2 Deuxième base : Données de la rétinographie

Comme nous avons dit précédemment, nos images rétinographiques proviennent de deux machines différentes, nous allons par la suite détailler chacune d'entre elles.

3.2.1 Données issues de l'EIDON

Notre base de données comprend également des images obtenues avec l'appareil EIDON, qui est un système moderne de capture d'images du fond d'œil. Au total, nous avons collecté les données de 125 patients, répartis en cas d'œdème maculaire diabétique (OMD) et des cas normaux, c'est-à-dire sans signes cliniques évidents de pathologie rétinienne. Les images ont été enregistrées au format « .jpg », un format largement utilisé car il permet de garder une bonne qualité visuelle tout en prenant peu de place en mémoire. Cela facilite leur manipulation et leur traitement par la suite. La Figure III.7 suivante illustre la répartition des patients selon les cas (NORMAL, OMD).

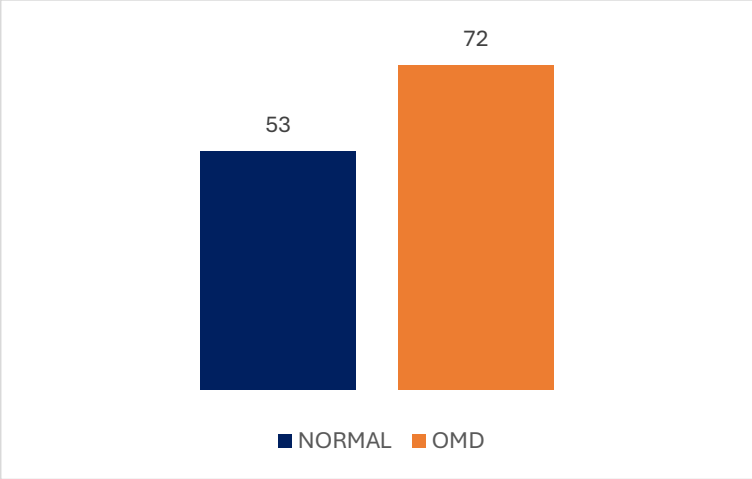


Figure III.7: Effectif des patients collectés de l'appareil EIDON : NORMAL et OMD

Les images prises avec l'EIDON bénéficient d'une haute résolution de 4608 x 3288 pixels (environ 14 mégapixels), ce qui permet de visualiser avec une grande précision les détails fins de la rétine. Cette qualité est particulièrement utile pour repérer les signes discrets d'anomalies, notamment dans le cadre du dépistage de l'OMD.

Chaque image couvre un champ de vision de 90°, mais grâce à la fonction SmartMosaic, plusieurs images peuvent être automatiquement assemblées pour obtenir une vue panoramique étendue jusqu'à 200°. Cela permet d'examiner non seulement la macula et la papille, mais aussi des zones plus périphériques de la rétine, souvent négligées par les systèmes plus classiques. Ce

type d'imagerie offre une résolution élevée et un bon rendu des détails anatomiques de la rétine, ce qui en fait un outil précieux pour le dépistage précoce**,** le diagnostic et le suivi de l'OMD.

La Figure III.8 ci-dessous représente un exemple des images prise à l'aide d'un appareil EIDON.

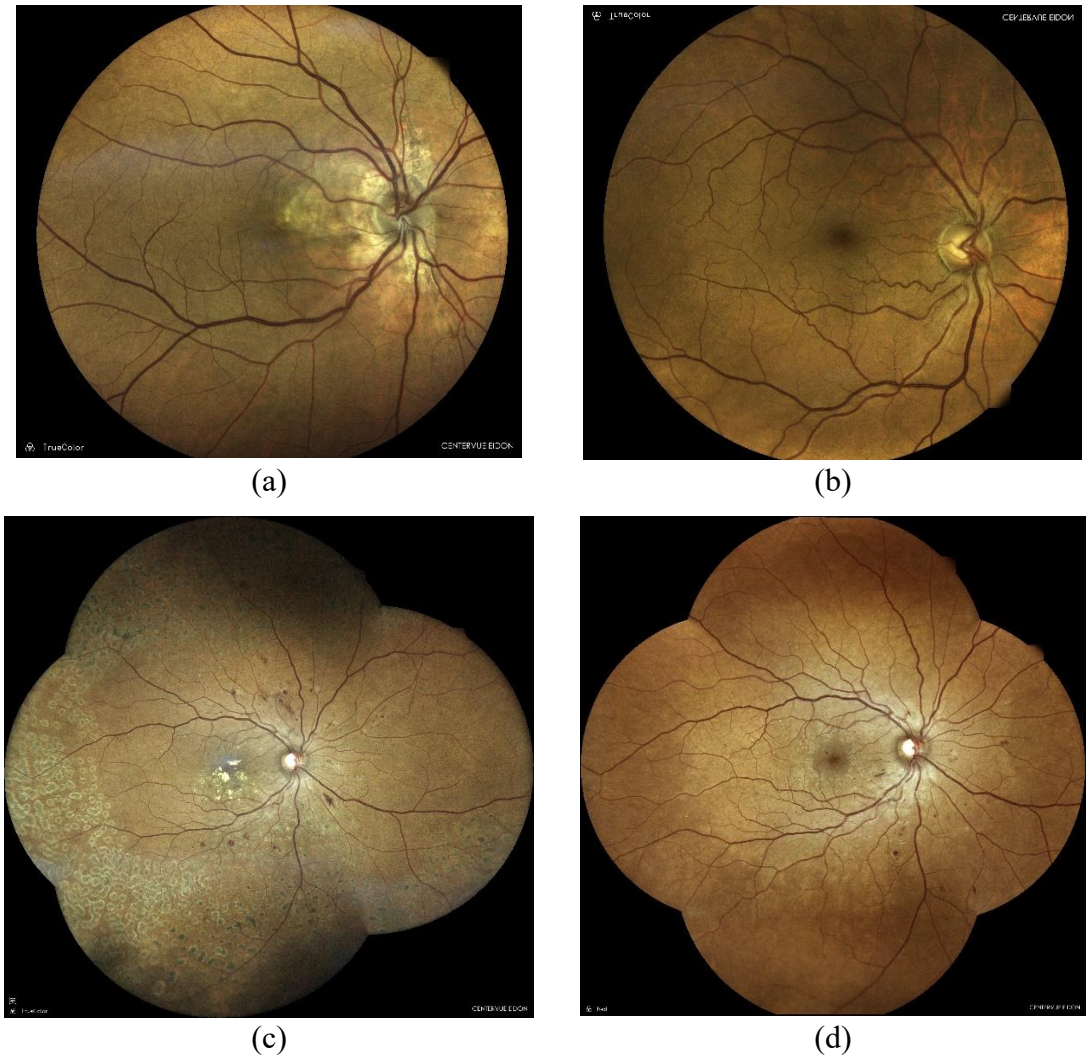


Figure III.8: Exemples d'images collectées avec EIDON. (a) Image 120° d'un cas atteint d'OMD ;(b) Image 120° d'un cas NORMAL ;(c) Image 200° d'un cas OMD ;(d) Image 200° d'un cas NORMAL

3.2.2 Données issues de l'OPTOS

Une autre partie importante de notre base de données provient de l'imagerie réalisée avec le système OPTOS, un appareil reconnu pour sa capacité à capturer des images ultra-grand champ de la rétine. Nous avons collecté les données de 394 patients avec cet appareil, répartis en cas

NORMAL et cas atteint d'OMD. La figure suivante illustre la répartition des patients selon les cas (NORMA, OMD).

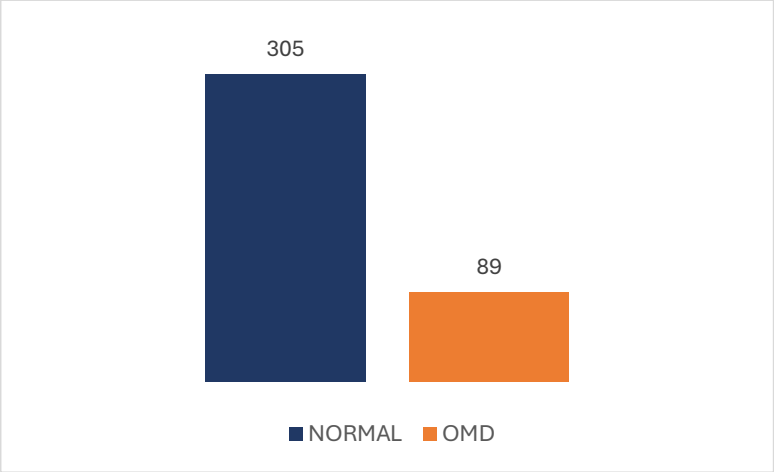


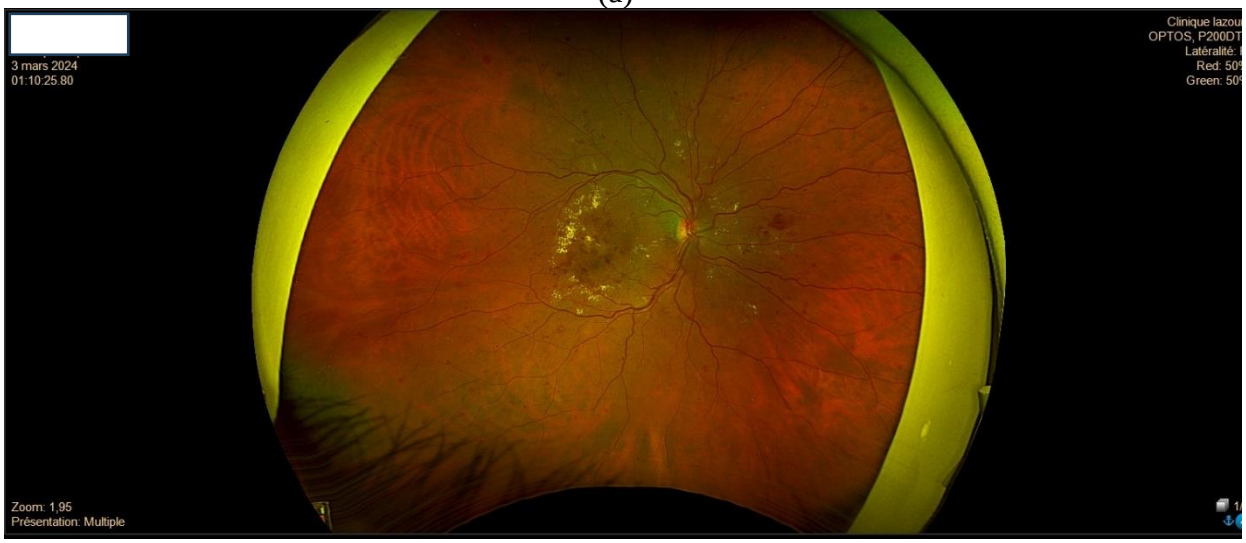
Figure III.9: Effectif des patients collectés de l'appareil Optos : NORMAL et OMD

Toutes les images ont été enregistrées au format « .jpg », un format léger et largement utilisé, qui permet de conserver une bonne qualité visuelle tout en facilitant l'archivage et le traitement informatique. Les clichés obtenus avec OPTOS se distinguent par leur champ de vision pouvant atteindre jusqu'à 200°, ce qui offre une couverture très large de la rétine. Cette particularité est précieuse, car elle permet d'observer des zones souvent négligées, notamment en périphérie, où certaines lésions de l'OMD peuvent apparaître de façon subtile.

Chaque image a une taille moyenne de 3900 x 3072 pixels, ce qui correspond à une résolution d'environ 14 mégapixels. Cette haute définition garantit une visualisation claire et précise des détails anatomiques de la rétine, essentielle pour un diagnostic fiable.

Un exemple d'image rétinienne obtenue avec OPTOS est illustré ci-dessous(figureIII.10).

(a)



(b)

Figure III.10: Exemples d'images collectées avec Optos. (a) Rétine saine ; (b) Rétine affectée par l'OMD

4. Analyse descriptive de la population étudiée

Afin de mieux comprendre la diversité et la représentativité de notre base de données, nous avons réalisé une analyse descriptive de l'ensemble des patients inclus, qu'ils soient atteints d'OMD ou non. Cette analyse s'appuie sur trois critères : la répartition géographique (wilayas), la répartition selon le sexe, et selon les tranches d'âge. Elle permet de s'assurer que notre base couvre différentes catégories de la population, ce qui renforce sa qualité et sa fiabilité pour les analyses à venir.

4.1 Répartition géographique

Nos patients proviennent de différentes wilayas représentées comme suit dans figure ci-dessous.

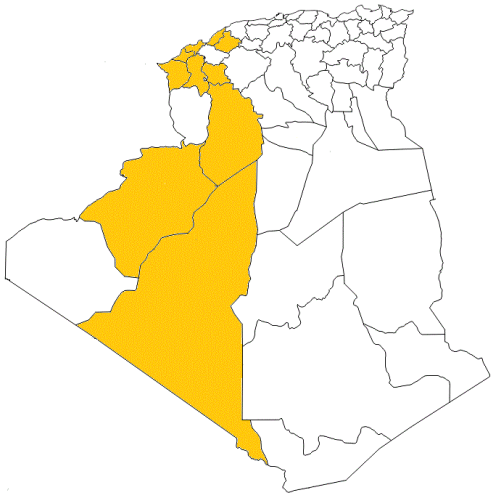


Figure III.11: Répartition géographique des patients inclus dans la base de données (OMD et normaux)

La carte ci-dessus illustre la provenance géographique des patients ayant participé à la constitution de notre base de données, qu'ils soient atteints d'œdème maculaire diabétique (OMD) ou ne présentant aucune anomalie oculaire. Les wilayas en jaune représentent les régions d'origine des patients inclus. Cependant, nous notons que certaines régions sont nettement plus représentées que d'autres, On constate une prédominance claire de la région Ouest de l'Algérie, avec une forte concentration de cas dans la wilaya de Tlemcen ce qui justifie la localisation de notre clinique. D'autres wilayas comme Adrar, El Bayadh, Béchar et Saïda comptent chacune entre 5 et 6 patients, montrant une présence modérée de cas. À l'inverse, certaines wilayas enregistrent un nombre plus faible de patients, comme Sidi Bel Abbès, Relizane, Mostaganem et Aïn Témouchent. Cette disparité peut s'expliquer par plusieurs facteurs, notamment la disponibilité des infrastructures médicales spécialisées, ou même la sensibilisation au dépistage.

Cette étude révèle une diversité géographique intéressante qui assure une représentativité des données et renforce la robustesse ainsi la cohérence de notre BDD. Elle permet de prendre en compte des patients issus de milieux de vie variés, ayant potentiellement des habitudes

alimentaires, des expositions environnementales et des conditions sanitaires différentes. Ces éléments peuvent influencer l'apparition et l'évolution des maladies rétiniennes comme l'OMD.

4.2 Répartition selon le sexe

La Figure III.12 suivante illustre la répartition des patients inclus dans notre base de données selon le sexe (homme/femme) et le type de cas (OMD ou normal) :

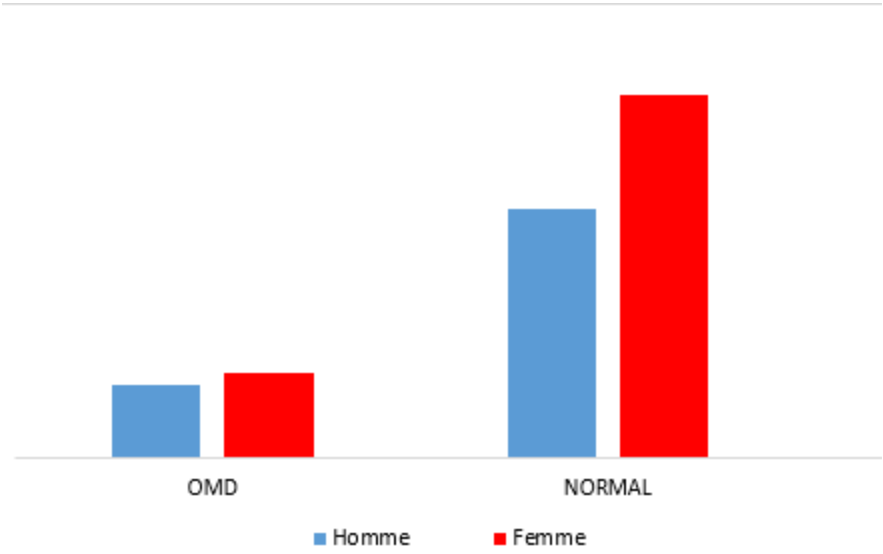


Figure III.12: Répartition des patients selon le sexe et le statut pathologique (OMD ou Normal)

Cette figure compare le nombre d'hommes et de femmes dans l'échantillon étudié des patients. Cependant, pour les patients atteints d'OMD, on constate que les femmes sont légèrement plus nombreuses que les hommes. Cette différence peut s'expliquer par une espérance de vie plus élevée chez les femmes, ce qui augmente leur exposition aux complications du diabète comme l'OMD. De plus, les femmes consultent souvent plus régulièrement pour leur santé visuelle, ce qui peut favoriser un diagnostic plus fréquent.

Du côté des patients sains, l'écart est encore plus marqué, avec une nette majorité de femmes. Cette tendance peut refléter une plus grande participation féminine aux examens de routine ou une sensibilisation accrue aux consultations préventives.

Ces observations mettent en évidence l'importance de la représentation des deux sexes dans l'analyse des données médicales. Elles contribuent à une meilleure compréhension des comportements de consultation et renforcent la qualité statistique de la base de données.

4.3 Répartition selon les tranches d'âge

Afin de mieux cerner la structure démographique de l'échantillon, la Figure III.13 suivante illustre la distribution des patients atteints d'OMD ainsi que des patients normaux selon quatre tranches d'âge.

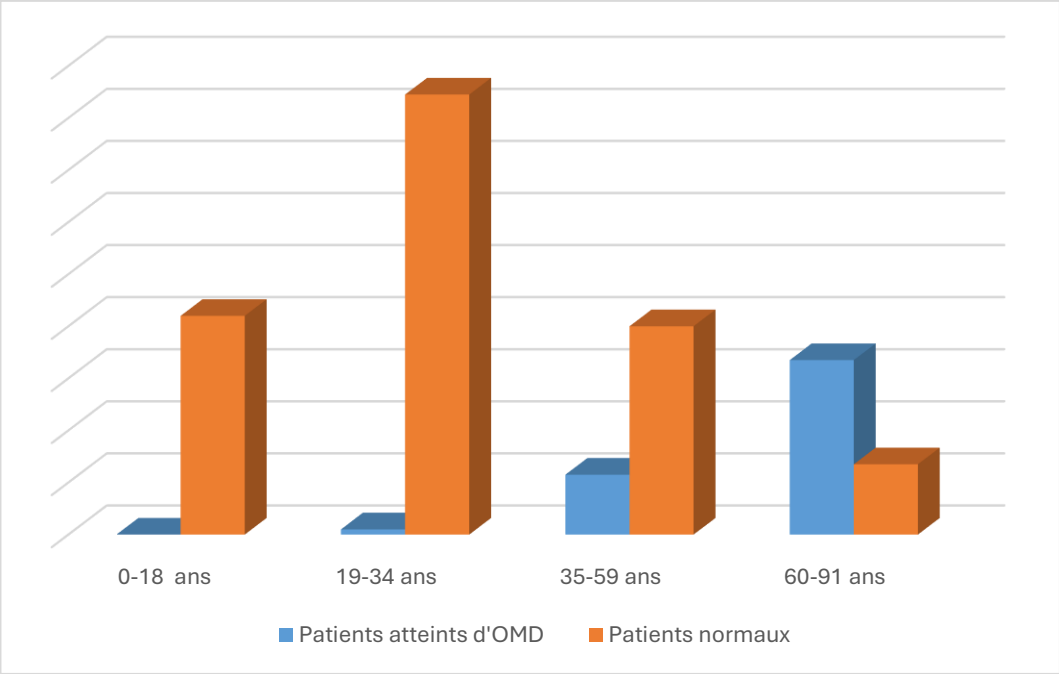


Figure III.13: Répartition des patients selon les tranches d'âge

L'analyse met en évidence une nette différence entre les deux populations. Les cas normaux sont majoritairement concentrés dans les tranches jeunes et adultes, notamment entre 19 et 34 ans, suivie de près par les 0–18 ans et les 35–59 ans. En revanche, les cas atteints d'OMD apparaissent plus fréquents dans la tranche d'âge de 60 à 91 ans. Ce résultat reflète la réalité clinique de cette pathologie, qui survient en général après plusieurs années d'évolution du diabète. Les complications ophtalmologiques comme l'OMD touchent donc plus souvent les personnes âgées, ce qui explique la forte présence de cette tranche dans les cas pathologiques.

Ce contraste entre les deux profils d'âge met en lumière l'importance de l'âge comme facteur de risque dans l'apparition de l'OMD et souligne également l'hétérogénéité de la base de données en termes de représentation démographique.

5. Structure de la Base de Données

Les données ont été organisées de manière hiérarchique pour faciliter leur manipulation et leur traitement dans le cadre de l'entraînement du modèle. La base de données est divisée en deux dossiers principaux correspondant aux deux modalités d'imagerie :

- Dossier OCT : Contient toutes les images OCT collectées.
- Dossier Rétinographies : Contient les images issues de la rétinographie grand champ.

À l'intérieur de chaque dossier, les images sont classées en fonction du patient, de la date d'examen et de l'œil concerné comme illustré dans la Figure III.14 suivante.

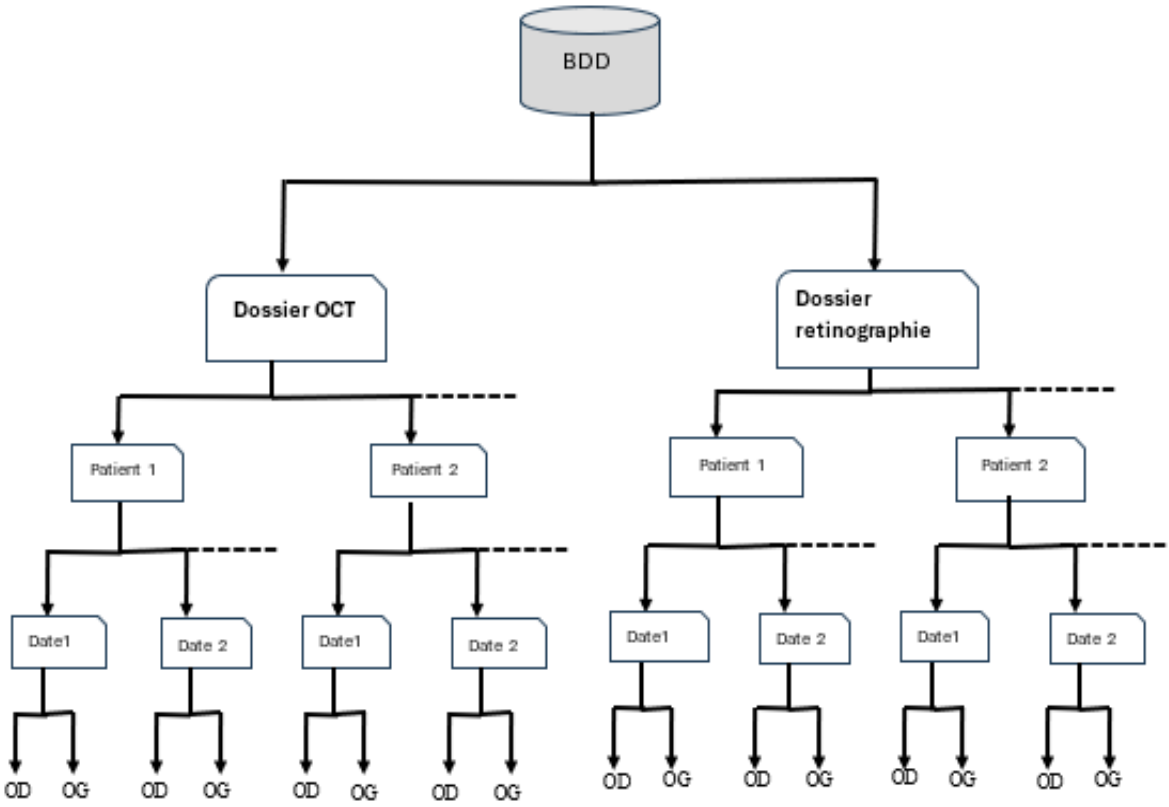


Figure III.14: Organisation de la base de données collectées

6. Description de la base de données publique

En plus des données que nous avons collectées dans notre étude, nous avons utilisé une base de données publique, la base OCT2017, proposée par Kermany et ses collègues en 2018 [80]. Elle est disponible gratuitement sur la plateforme Kaggle. Cette base est largement utilisée dans la recherche en ophtalmologie et en intelligence artificielle, notamment pour la détection automatisée des maladies rétiniennes. Les images de cette base de données ont été prises avec un appareil OCT spectral-domain (SD-OCT) de la marque Heidelberg Engineering, un équipement souvent utilisé en clinique pour examiner la rétine.

Chaque image correspond à une coupe transversale de la rétine, ce qui permet de bien observer ses différentes couches et de repérer des anomalies comme celles causées par l'OMD.

La base contient **84 484 images**, classées en quatre catégories :

- **CNV** : Choroidal Neovascularization ou néovascularisation choroïdienne,
- **DME** : Diabétique Macular Edema ou œdème maculaire diabétique,
- **Drusen** : Dépôts liés à la DMLA,
- **Normal** : rétine saine.

Ces images sont réparties en trois groupes :

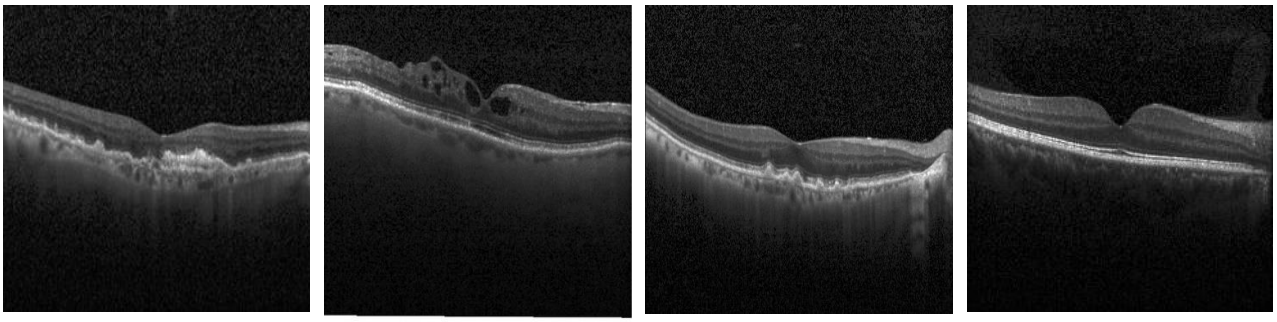
- Un groupe d'apprentissage (train) avec 83 484 images,
- Un groupe de validation (val) avec 32 images,
- Un groupe de test (test) avec 968 images.

Leur répartition est représentée dans le tableau III-2 ci-dessous :

Tableau III-2: Distribution des images de la base de données publique(originale)

Catégorie / Images	Train	Val	Test
CNV	37 205	9	243
DME	11 348	9	243
Drusen	8 616	9	243
NORMAL	26 315	9	243

Les images sont enregistrées au format « .jpeg » un format couramment utilisé qui permet de conserver une bonne qualité visuelle tout en réduisant la taille des fichiers, facilitant ainsi leur manipulation et leur traitement**,** avec une résolution moyenne de 500 x 500 pixels, ce qui garantit un bon compromis entre qualité d'image et vitesse de traitement, tout en facilitant leur intégration dans des algorithmes de classification. La figure suivante illustre un exemple des images disponible sur cette base de données.



(a) (b) (c) (d)

Figure III.15: Exemples d'images de la base publique. (a) CNV ; (b) DME ; (c) DRUSEN ; (d) NORMAL

La répartition des classes est déséquilibrée dans la version originale, avec un grand nombre d'images « CNV » et « Normal », et moins d'images pour les classes « DME » et « Drusen ». Pour remédier au déséquilibre des classes dans la version originale, une version équilibrée a été proposée par Mohamed Berrimi [81], elle aussi est disponible gratuitement sur la plateforme Kaggle. Cette version contient au total images répartis comme suit dans le tableau III-3 suivant :

Tableau III-3: Distribution des images de la base de données publique(équilibrée)

Catégorie / Images	Train	Test
CNV	8 016	242
DME	8 016	242
Drusen	8 016	242
NORMAL	8 016	242

7. Environnements de travail et bibliothèques Python utilisées :

Une fois la base de données constituée avec les images OCT et les rétinographies collectées à la clinique, plusieurs étapes ont été nécessaires avant de pouvoir entraîner un modèle de détection. Il a fallu préparer les images, cette phase de préparation a inclus le tri, le renommage, le redimensionnement, ainsi que le prétraitement des images et une organisation rigoureuse des

dossiers. Ces opérations ont été réalisées dans un environnement de travail local utilisant le logiciel Anaconda, associé à l'éditeur de code Spyder. Anaconda est un logiciel qui installe Python ainsi qu'un ensemble d'outils spécialisés pour le traitement d'image, l'analyse de données et le machine learning. Il fournit tout ce qu'il faut pour développer un projet scientifique en Python. Plus précisément, c'est une distribution complète de Python qui contient :

- Le langage Python lui-même,
- Des bibliothèques déjà installées (comme NumPy, Pandas, OpenCV, etc.),
- Des environnements de développement comme Spyder.
- Un gestionnaire de paquets (Conda) pour installer ou mettre à jour facilement d'autres outils.

Spyder quant à lui, inclus avec Anaconda, est un environnement de développement interactif qui facilite l'écriture du code, l'affichage des résultats, et le suivi des données en mémoire.

Ensuite, l'étape de détection proprement dite c'est-à-dire l'entraînement et l'évaluation des modèles de Deep learning a été effectuée sur la plateforme en ligne Kaggle, qui permet d'utiliser des ressources puissantes comme les GPU.

7.1 Environnement local : Anaconda (Spyder)

L'environnement Anaconda, associé à Spyder, a été utilisé pour le prétraitement des images. Cet environnement a permis :

- De charger et modifier les images (grâce à des bibliothèques comme OpenCV),
- De redimensionner les images,
- D'améliorer la qualité visuelle (amélioration du contraste),
- D'organiser les dossiers selon la catégorie (OMD ou normal),
- Et de renommer les images.

Cette phase était essentielle pour garantir la qualité des données utilisées lors de l'entraînement.

7.2 Plateforme Kaggle

Dans ce projet, nous avons utilisé Kaggle pour lancer les modèles, visualiser les résultats et tester plusieurs architectures.

Kaggle est une plateforme gratuite très populaire pour les projets de machine learning et de deep learning. Autrement dit c'est un environnement de programmation qui permet de coder directement en ligne sans avoir besoin d'installer quoi que ce soit sur son ordinateur. L'un des avantages majeurs de Kaggle est la possibilité d'utiliser des GPU gratuitement, des cartes graphiques qui permettent d'entrainer des modèles de deep learning de manière rapide et efficace en accélérant les calculs. Dans le cadre de nos expériences, nous avons bénéficié de la carte graphique NVIDIA Tesla P100, qui permet d'accélérer significativement les temps d'entraînement des modèles. Grâce à sa grande capacité de calcul parallèle, cette carte est particulièrement adaptée aux tâches complexes de deep learning, comme la classification d'images médicales. Elle permet de gagner un temps précieux tout en assurant de bonnes performances lors de l'apprentissage.

7.3 Bibliothèques Python

Pour ce projet, nous avons choisi d'utiliser le langage Python, apprécié pour sa simplicité d'écriture et la richesse de son écosystème en science des données. Parmi les bibliothèques que nous avons utilisées, TensorFlow occupe une place centrale. Développée par Google, cette bibliothèque open-source nous a permis de concevoir, entraîner et évaluer nos modèles de deep learning de manière souple et efficace. Grâce à ses nombreuses fonctionnalités, TensorFlow facilite la gestion des réseaux de neurones complexes, tout en assurant une bonne compatibilité avec les ressources matérielles disponibles. Elle s'est révélée particulièrement adaptée à notre tâche de classification d'images médicales.

8. Conclusion

La collecte de données est une étape cruciale dans tout projet de recherche, en particulier dans le domaine médical, où la qualité des données a un impact direct sur la précision des résultats. Les données doivent être soigneusement sélectionnées, organisées et préparées afin d'assurer leur pertinence et leur utilisabilité. Cela implique non seulement l'acquisition des images, mais aussi leur annotation par des experts médicaux et leur structuration de manière à faciliter l'analyse. Dans ce chapitre, nous avons présenté la base de données collectée ainsi que la base publique. Nous avons également détaillé l'environnement de travail choisi, avec l'utilisation de Kaggle pour

l’entraînement des modèles, et des outils comme Anaconda et Spyder pour le traitement et l’organisation des données ainsi que les bibliothèques Python, qui ont été essentielles pour le développement des modèles et l’analyse des images.

Cette organisation rigoureuse des données et des outils permet de préparer les prochaines étapes du projet, qui auront pour objectif d'appliquer des techniques de deep learning afin de détecter l'œdème maculaire diabétique.

Dans le chapitre qui suit, nous allons présenter nos algorithmes développés pour la détection automatique de l’OMD à partir des images rétiniennes en utilisant les techniques de traitement d’images et l’apprentissage profond. Ainsi qu’une discussion et comparaison des résultats obtenus.

IV. Chapitre IV : Méthodes et résultats

1. Introduction

L'œdème maculaire diabétique est une complication fréquente et invalidante de la rétinopathie diabétique, affectant la région centrale de la rétine appelée macula. Il se manifeste par une accumulation de liquide au niveau de cette région, entraînant une altération progressive de l'acuité visuelle. Une détection automatique, précoce et fiable de cette pathologie représente un enjeu majeur pour prévenir la perte de vision chez les patients diabétiques.

Ce chapitre est consacré à la présentation des différentes méthodes développées dans le cadre de ce projet pour la détection automatique de l'OMD à partir d'images rétiniennes. Les travaux menés s'appuient principalement sur des techniques de traitement d'images et d'apprentissage profond, appliquées à deux modalités d'imagerie : OCT (Optical Coherence Tomography) et rétinographie. L'approche proposée repose sur plusieurs étapes clés : le prétraitement des images pour l'extraction de la région d'intérêt, l'entraînement de modèles de deep learning pour la classification des images normales et pathologiques, ainsi que l'évaluation rigoureuse des performances sur des bases de données locales et publiques. Des expérimentations ont été menées avec et sans techniques d'augmentation de données, et une analyse comparative a été effectuée en croisant les sources de données (base locale vs base publique). Les résultats expérimentaux sont présentés de manière détaillée, en mettant en évidence l'impact de chaque stratégie sur les métriques de performance, notamment l'exactitude, la sensibilité, la spécificité et l'AUC.

2. Méthodes proposées

Dans ce travail, nous avons développé une approche de classification automatique des images rétiniennes en deux classes : normales et pathologiques (OMD), en utilisant des techniques de Deep Learning appliquées à deux modalités d'imagerie :

- OCT (Tomographie en Cohérence Optique).
- Rétinographie.

Dans les sections suivantes, nous allons décrire séparément chaque méthodologie appliquée selon le type d'images, en détaillant les traitements spécifiques, les modèles utilisés, les résultats obtenus et les analyses comparatives associées.

2.1 Détection de l'OMD par les images OCT

L'imagerie par tomographie en cohérence optique (OCT) est l'un des outils les plus puissants pour l'analyse fine de la structure rétinienne. Elle permet de visualiser les couches internes de la rétine avec une très haute résolution, ce qui en fait un moyen privilégié pour la détection de l'OMD. Dans cette section, nous décrivons les méthodes proposées pour identifier cette pathologie à partir des images OCT, en mettant en œuvre un pipeline basé sur le prétraitement, l'extraction de la région d'intérêt, puis l'application d'un algorithme d'apprentissage profond permettant de distinguer entre les cas normaux et pathologiques.

L'organigramme (Figure IV.1) ci-dessous présente notre système développé pour la détection et la classification de l'OMD.

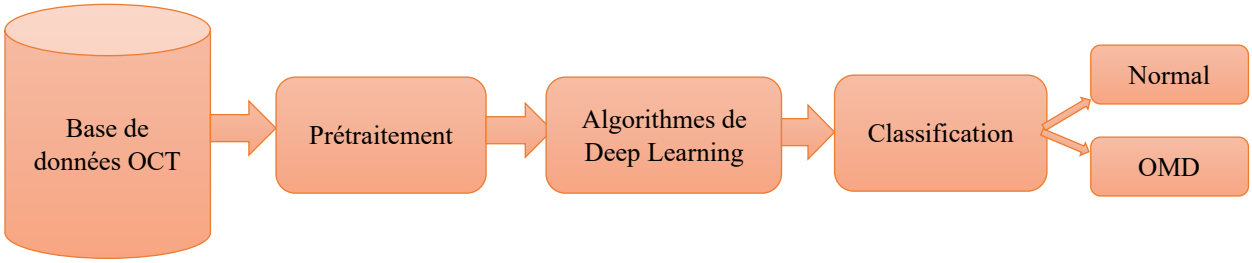


Figure IV.1: Organigramme de nos méthodes pour la détection de l'OMD à partir des images OCT

Dans la partie qui suit nous allons décrire chaque étape de ce pipeline proposé.

2.1.1 Prétraitement des images OCT

Avant d'entraîner un modèle de deep learning, il est essentiel d'effectuer un prétraitement rigoureux des images OCT afin d'améliorer la qualité des données et d'éliminer les éléments susceptibles d'introduire du bruit ou des biais, et ne garder que la région d'intérêt essentielle à la détection des lésions liées à cette pathologie. Cette étape permet de garantir que le modèle se concentre uniquement sur les informations visuelles pertinentes pour la détection de l'OMD. Plusieurs opérations de prétraitement ont été appliquées dont nous allons citer par la suite :

1) Suppression des informations indésirable :

Les images résultantes de l'examen OCT contiennent des flèches visibles, situées en haut à droite. Ces flèches représentent des informations indésirables qui peuvent perturber l'analyse et entraîner un biais dans le modèle d'apprentissage. Par conséquent, il est essentiel de supprimer ces flèches pour garantir que seules les caractéristiques pertinentes des images soient prises en compte

par le modèle. L'idée est de masquer cette flèche de taille 80x80 localisée en haut à droite, en la remplaçant par une zone noire.

De plus, les images OCT sont de grande taille, avec un fond vaste qui inclut des zones non pertinentes. Ce fond excessif augmente la surface d'image à traiter, ce qui peut entraîner un temps d'exécution plus long lors de l'entraînement du modèle. Cette étape vise à extraire la région d'intérêt en supprimant les marges inutiles de la rétine (30 lignes en haut et 30 en bas), ce qui permet de mieux se concentrer sur la rétine et d'améliorer l'efficacité du modèle.

2) Normalisation

Les images issues de l'examen OCT peuvent parfois présenter des niveaux de gris très différents en raison des conditions de prise de vue ou des variations entre dispositifs. Ces différences de luminosité et de contraste peuvent nuire à la qualité de l'apprentissage et compromettre la capacité de généralisation du modèle, en le poussant à interpréter des variations non pertinentes comme des caractéristiques discriminantes. En effet, un même type de structure rétinienne peut apparaître plus sombre ou plus clair d'une image à l'autre, alors qu'il s'agit de la même information anatomique, le modèle risque de se perdre face à ces incohérences et de mal apprendre les véritables motifs liés à l'œdème maculaire diabétique.

Pour y remédier, nous avons appliqué une normalisation min-max, une technique qui consiste à ramener toutes les valeurs de pixels dans une plage commune, ici entre 0 et 255. Cela permet de standardiser la dynamique de gris de chaque image, quelles que soient ses conditions d'acquisition, et d'assurer une meilleure cohérence dans les données utilisées pour l'entraînement. La transformation est définie par la formule suivante :

$$I_{normalise'} = \frac{I - I_{Min}}{I_{Max} - I_{Min}} \times 255 \qquad \textbf{(IV-1)}$$

Où :

- I est la valeur du pixel original,
- I_{Min} est la **valeur minimale** des pixels dans l'image (le plus sombre),
- I_{Max} est la **valeur maximale** des pixels dans l'image (le plus clair),
- $I_{normalise'}$ la nouvelle valeur du pixel, après transformation.

3) Amélioration de contraste

Les détails dans les images OCT ne sont pas toujours bien visibles due au problème de contraste, ce qui rend la visualisation des informations difficile. Pour cela, nous avons utilisé la méthode CLAHE (*Contrast Limited Adaptive Histogram Equalization*) qui est une méthode avancée d'égalisation d'histogramme. Contrairement à l'égalisation classique, qui traite toute l'image globalement, **CLAHE divise l'image en petites zones appelées *tuiles*** (par exemple 8×8 pixels). Pour chaque tuile, l'histogramme est égalisé individuellement afin d'améliorer le contraste local. Une **interpolation bilinéaire** est ensuite utilisée pour recoller ces zones de manière fluide et éviter les discontinuités. De plus, CLAHE limite l'amplification du bruit grâce à un **paramètre de coupure appelé *clipLimit***. Ce paramètre empêche une augmentation trop forte du contraste dans des zones déjà uniformes, ce qui est important pour les images médicales. Dans notre cas nous avons pris :

- clipLimit = 2.0 : empêche le contraste de devenir excessif,
- tileGridSize = (8, 8) : divise l'image en 64 régions locales.

La Figure IV.2 qui suit, illustre un exemple des transformations effectuées sur les images OCT.

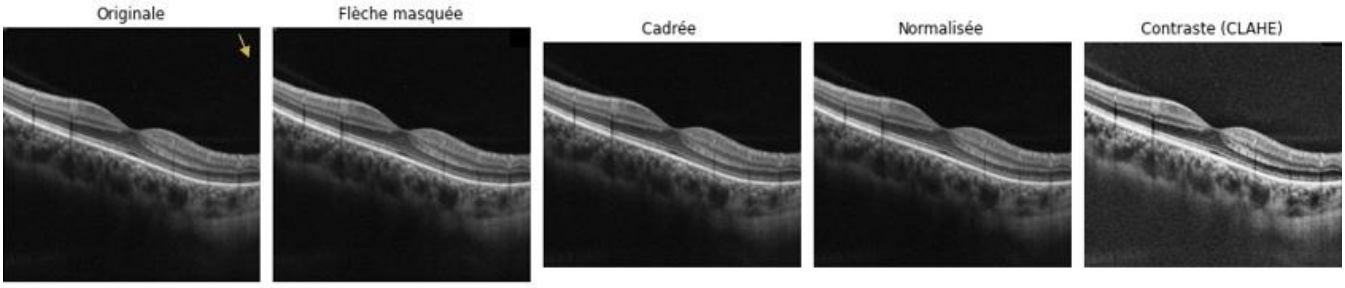
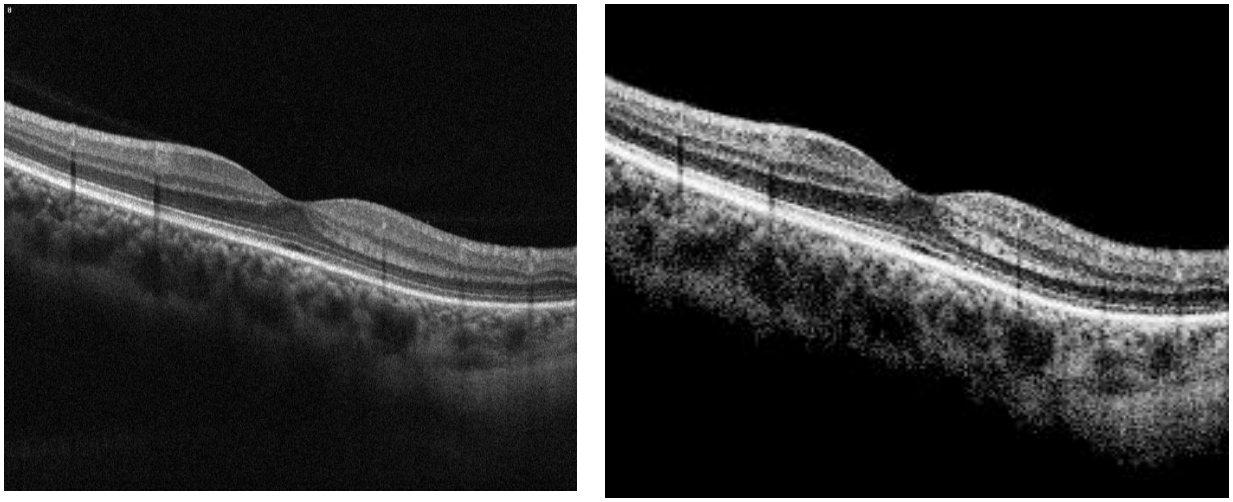


Figure IV.2 : Etapes de prétraitement des images OCT

4) Extraction de la région d'intérêt

Pour extraire la région d'intérêt finale, notamment la rétine et éliminer les parties inutiles comme le fond et les bords, nous avons appliqué un **seuillage global,** une technique qui consiste à distinguer les zones claires des zones sombres en fixant un seuil d'intensité égale à 15, afin d'améliorer les performances des modèles.

La Figure IV.3 ci-dessous illustre le résultat final du prétraitement, après extraction de la région d'intérêt.



(a) (b)

Figure IV.3: Extraction de la région d'intérêt dans une image OCT. (a) Image originale ; (b) Image traitée

2.1.2 Algorithmes d'apprentissage profond pour la détection de l'OMD

Contrairement aux méthodes classiques qui nécessitent une extraction manuelle des caractéristiques, les réseaux de neurones convolutifs apprennent automatiquement à identifier les zones suspectes à partir des données brutes. Ces algorithmes sont aujourd'hui largement utilisés dans le domaine médical pour leur capacité à repérer avec précision des anomalies subtiles.

Dans cette partie, nous présentons les modèles d'apprentissage profond que nous avons utilisés pour la détection de l'OMD à partir des images de notre base de données OCT, notamment CNN construit, VGG16, ResNet50, Inception v3, MobileNet v2.

2.1.2.1 Augmentation de données

Notre base de données est déséquilibrée (nombre d'images NORMALES était inférieur à celui l'OMD), ce qui peut perturber les performances des modèles. Donc, avant d'entraîner nos modèles d'apprentissage profond, une première étape essentielle consiste à équilibrer notre base de données en appliquant une augmentation de données ciblée uniquement sur les images normales. Cela nous a permis d'améliorer la représentativité des deux classes et d'éviter que le modèle n'apprenne à favoriser la classe majoritaire. L'augmentation a été réalisée à partir des images prétraitées, en appliquant une série de transformations géométriques destinées à simuler la variabilité naturelle des acquisitions. Plus précisément, nous avons utilisé :

- Des **rotations** de ±15°,
- Un **retournement horizontal** (effet miroir),
- Des modifications de **contraste** (+30%, -30%),
- Des ajustements de **luminosité** (+20, -20),
- Et l'ajout d'un **bruit gaussien**.

Ces opérations, illustrées dans la Figure IV.4 ci-dessous, ont permis de générer artificiellement de nouveaux échantillons tout en conservant les caractéristiques essentielles de l'imagerie OCT, ce qui a renforcé la capacité de généralisation de nos modèles.

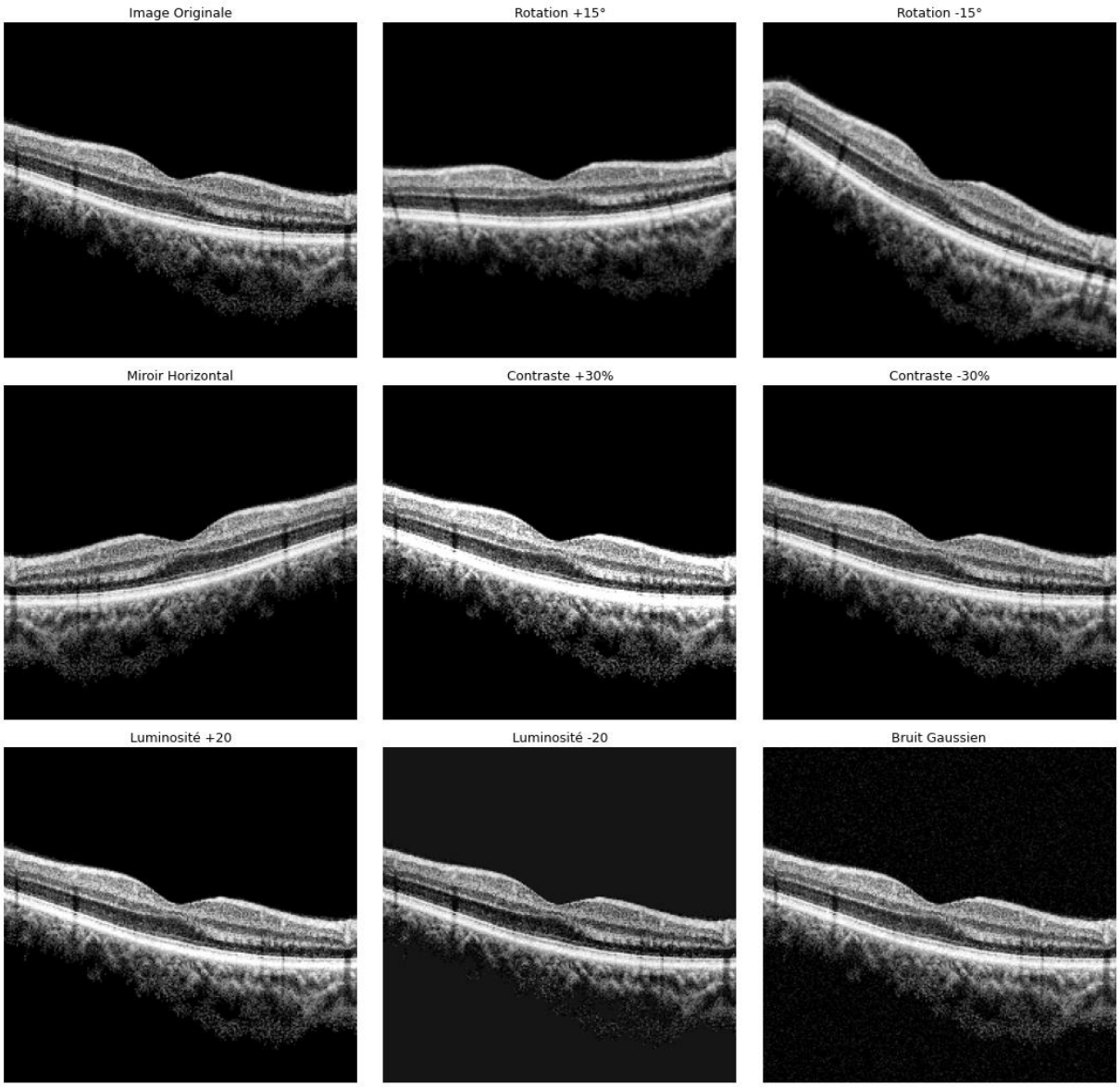


Figure IV.4: Exemples de transformations utilisées pour l'augmentation des données dans la classe NORMAL

2.1.2.2 Modèle CNN proposé

Nous avons construit une architecture CNN avec des poids aléatoires (en anglais from scratch), composée de quatre étages convolutionnels, chacun suivi d'une opération de max pooling de taille 2×2. Chaque couche de convolution est activée par la fonction ReLU, et la profondeur des filtres augmente progressivement : 32 filtres de taille 3×3 dans le premier étage, 64 filtres dans le deuxième, et 128 filtres dans le troisième et quatrième étage. Après l'extraction des caractéristiques, une couche de vectorisation (flattening) est appliquée, suivie d'une couche DropOut avec un taux de 40 % pour atténuer le sur-apprentissage. Le réseau se termine par une couche dense contenant 512 neurones avec activation ReLU, suivie d'une couche de sortie sigmoïde adaptée à la classification binaire (OMD / normal). L'optimiseur utilisé est Adam avec un faible taux d'apprentissage (1e-5), ce qui permet un entraînement progressif et stable.

La Figure IV.5 ci-dessous illustre un diagramme représentatif de l'architecture de ce modèle.

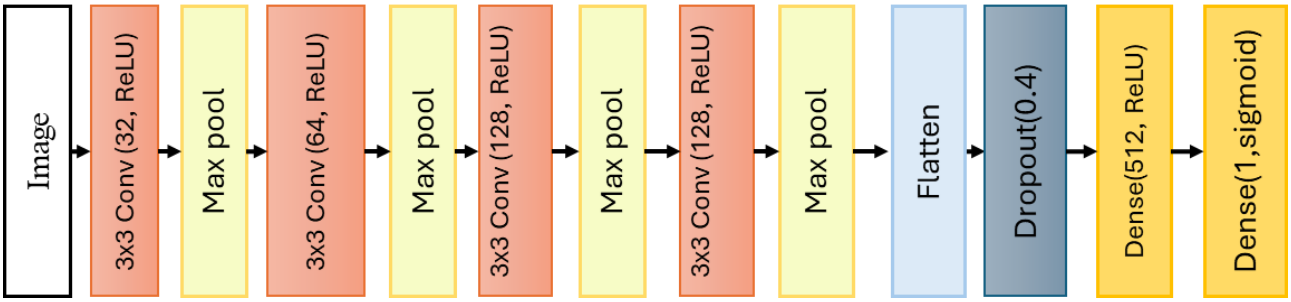


Figure IV.5: Notre architecture du CNN from scratch pour la classification des images OCT

2.1.2.3 Modèle VGG16 proposé

Dans le cadre de notre travail, nous avons exploité l'architecture VGG16 pré-entraînée, et nous l'avons adaptée pour une tâche de classification binaire centrée sur la détection de l'OMD. Cette adaptation repose sur une version modifiée du modèle original, dans laquelle certaines couches ont été conservées, d'autres gelées, et de nouvelles couches spécifiques ont été ajoutées. Plus précisément, nous avons conservé la partie convolutionnelle de VGG16, composée de cinq blocs de convolution, afin de bénéficier des capacités d'extraction de caractéristiques déjà acquises. Toutefois, pour optimiser l'apprentissage sur notre jeu de données, seuls les blocs 4 et 5 ont été rendus entraînables, tandis que les blocs 1 à 3 ont été gelés. Cette stratégie permet d'équilibrer transfert d'apprentissage et spécialisation sur notre tâche cible.

Par ailleurs, nous avons supprimé la tête de classification d'origine, initialement conçue pour la classification multi-classes, et l'avons remplacée par une tête personnalisée adaptée à notre tâche binaire. Celle-ci se compose d'une couche de Global Average Pooling, qui permet de réduire la dimensionnalité tout en conservant les informations spatiales essentielles. Elle est suivie d'une couche de Batch Normalization pour améliorer la stabilité de l'entraînement, puis d'une couche dense de 256 neurones avec la fonction d'activation ReLU et une régularisation L2. Une couche de Dropout avec un taux de 0.3 est également ajoutée pour limiter le sur apprentissage. Enfin, une couche de sortie à un seul neurone, activée par une fonction sigmoïde, produit une probabilité correspondant à la classe prédite. Cette version modifiée du VGG16 combine ainsi la puissance de l'apprentissage profond pré-entraîné avec une architecture optimisée pour notre problématique spécifique. La Figure IV.6 suivante illustre en détail la structure du modèle final utilisé.

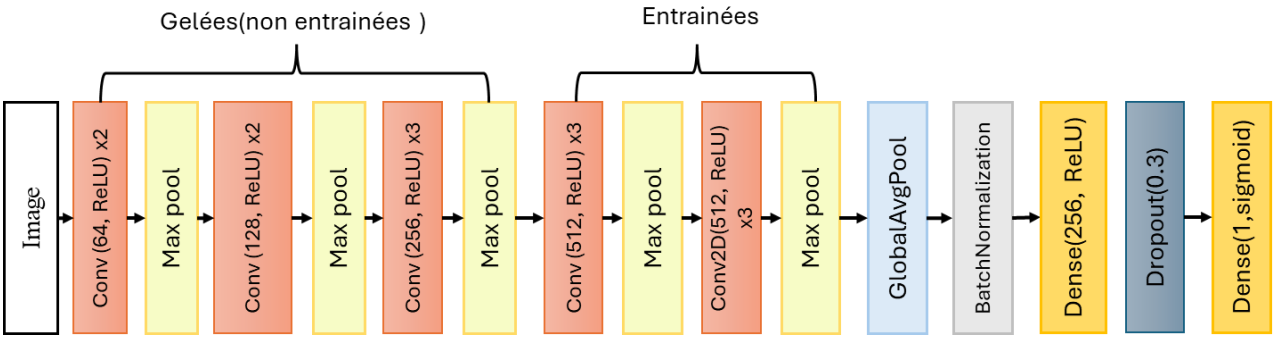


Figure IV.6: Architecture du modèle VGG16 modifié pour la classification des images OCT

2.1.2.4 Modèle ResNet50 proposé

Une autre architecture ResNet50 pré entrainée est utilisé en modifiant la tête de classification initiale. Les couches du backbone sont partiellement gelées, à l'exception des trois dernières, qui restent entraînables. À la sortie du ResNet50, une couche de Global Average Pooling est ajoutée, suivie de Batch Normalization. Ensuite, une couche *Dense* de 256 unités avec une activation ReLU et une régularisation L2 est introduite, accompagnée d'un *Dropout* de 0.3. Enfin, une couche de sortie *Dense* avec une activation sigmoïde permet la classification binaire. Le modèle est entraîné avec l'optimiseur Adam, et un callback *ReduceLROnPlateau* est utilisé pour ajuster dynamiquement le taux d'apprentissage. L'entraînement est effectué en *mixed precision* pour améliorer l'efficacité computationnelle. La Figure IV.7 suivante illustre en détail la structure du modèle final utilisé.

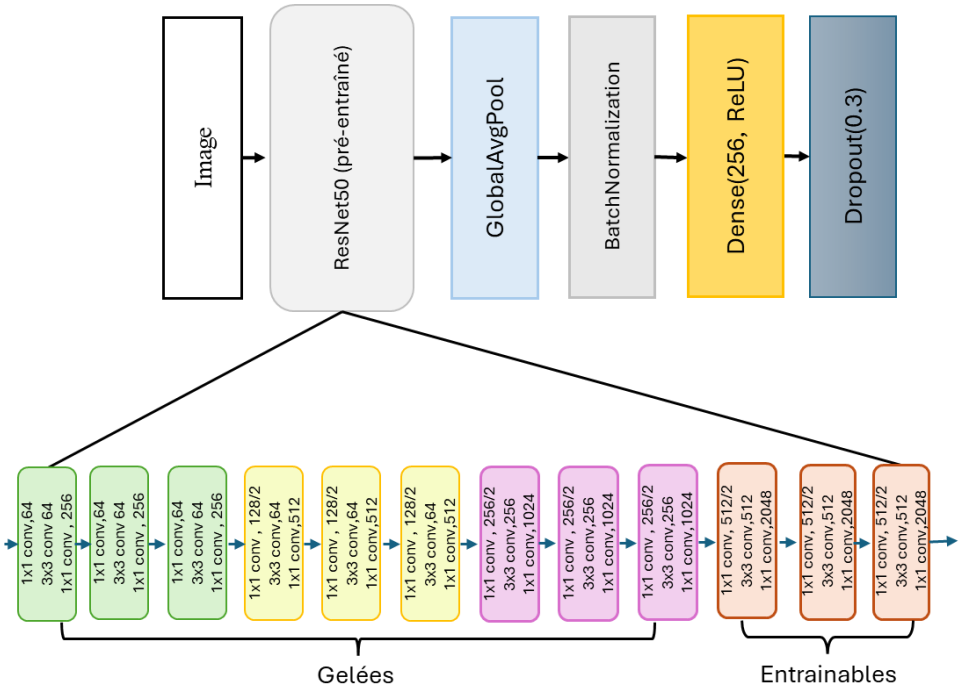


Figure IV.7: Architecture du modèle ResNet50 modifié pour la classification des images OCT

2.1.2.5 Modèle InceptionV3 proposé

Dans le cadre de notre expérimentation, nous avons utilisé une architecture basée sur le modèle InceptionV3, dont nous avons gelé les couches convolutionnelles afin de préserver les poids appris. Le modèle original a été modifié pour s'adapter à notre tâche de classification binaire. La nouvelle tête de classification comprend les couches suivantes :

- Une couche **GlobalAveragePooling** appliquée à la sortie du backbone InceptionV3 ;
- Une couche **Dense** avec 256 neurones et la fonction d'activation ReLU ;
- Une couche **Dropout** avec un taux de 0.4 pour réduire le surapprentissage ;
- Une couche **Dense de sortie** avec une seule unité et une activation sigmoïde, configurée pour produire une sortie de type float32 (compatibilité avec le mode mixed precision).

Le modèle est compilé avec l'optimiseur Adam, une fonction de perte binaire, et la métrique d'exactitude. L'entraînement est effectué en mixed precision pour améliorer la performance. Un callback ReduceLROnPlateau est utilisé pour ajuster dynamiquement le taux d'apprentissage. La Figure IV.8 suivante illustre en détail la structure du modèle final utilisé.

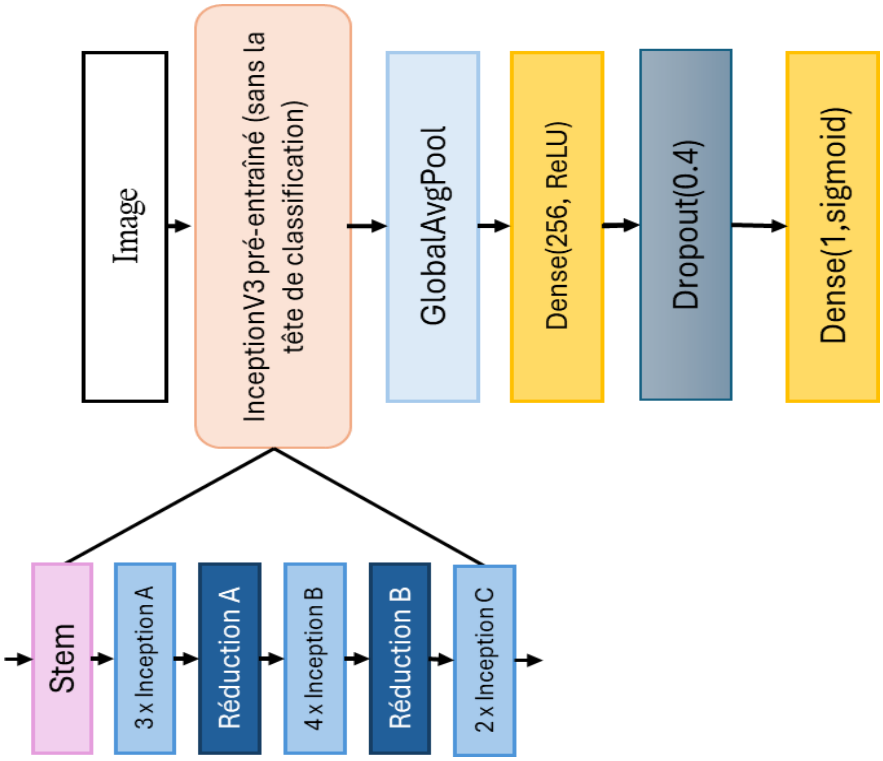


Figure IV.8: Architecture du modèle InceptionV3 modifié pour la classification des images OCT

2.1.2.6 Modèle MobileNetV2 proposé

Dans cette expérience, nous avons utilisé l'architecture MobileNetV2 pré-entraîné comme base, que nous avons adaptée pour une tâche de classification binaire. Le modèle d'origine a été modifié en excluant la tête de classification initiale (include_top=False) afin d'ajouter des couches personnalisées adaptées à notre problématique.

Nous avons appliqué une couche GlobalAveragePooling pour réduire la dimension des cartes de caractéristiques extraites par le réseau. Cette sortie a ensuite été normalisée à l'aide d'une couche BatchNormalization afin de stabiliser l'apprentissage. Une régularisation par Dropout avec un taux de 0.3 a été appliquée pour limiter le surapprentissage. Une couche Dense de 64 neurones avec la fonction d'activation ReLU et une régularisation L2 a été ajoutée pour apprendre des représentations intermédiaires pertinentes. Une seconde couche Dropout (0.3) a été insérée avant la couche finale afin de renforcer la régularisation. Enfin, une couche de sortie Dense avec un seul neurone et une activation sigmoïde permet de générer une probabilité d'appartenance à la classe positive, en cohérence avec la nature binaire du problème.

Le modèle a été compilé avec l'optimiseur Adam utilisant un taux d'apprentissage de 1e-4, la fonction de perte binary_crossentropy, et la métrique d'évaluation accuracy. Pour améliorer les performances, deux callbacks ont été utilisés : ReduceLROnPlateau afin d'ajuster dynamiquement le taux d'apprentissage en cas de stagnation de la performance en validation, et ModelCheckpoint pour sauvegarder automatiquement le meilleur modèle obtenu pendant l'entraînement.

La fonction d'activation sigmoïde choisie en sortie permet d'interpréter la prédiction comme une probabilité entre 0 et 1, ce qui facilite la prise de décision via un seuil fixé, typiquement à 0.5, parfaitement adapté à une classification binaire. La Figure IV.9 suivante illustre en détail la structure du modèle final utilisé.

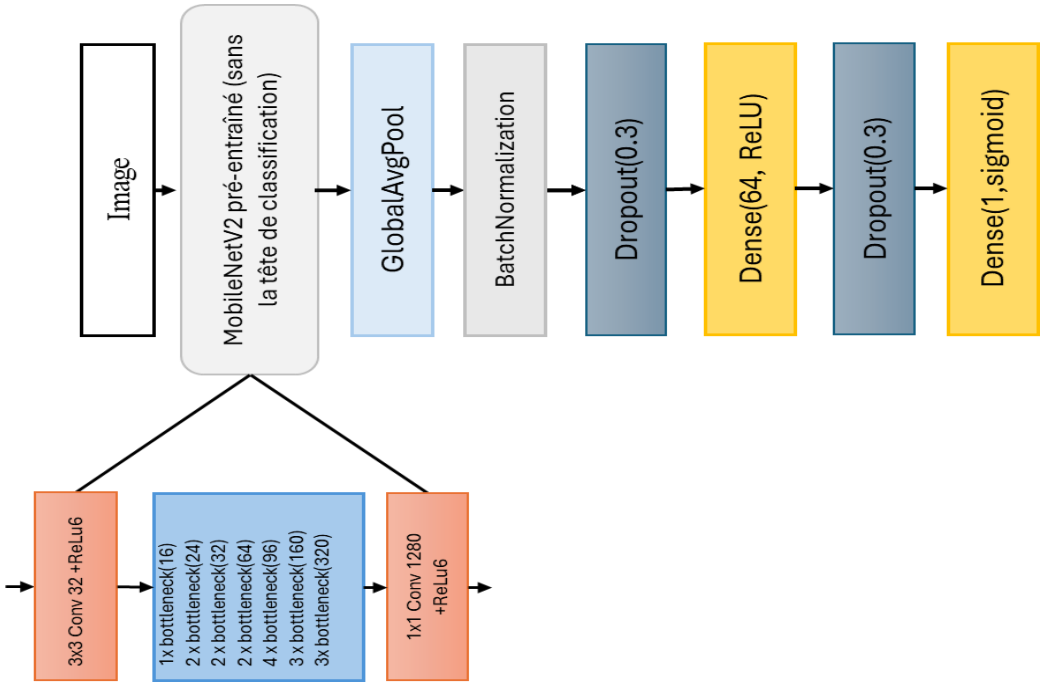


Figure IV.9: Architecture du modèle MobileNetV2 modifié pour la classification des images OCT

2.1.3 Expérimentations sur les images OCT

Afin d'évaluer leur efficacité dans la détection de l'œdème maculaire diabétique (OMD). Les algorithmes de classification que nous avons développés ont été appliqués à plusieurs scénarios expérimentaux sur des images OCT provenant à la fois de notre propre base de données locale et d'une base publique disponible en ligne. L'idée était de tester leur performance dans plusieurs situations : en utilisant les données avec ou sans augmentation, en les entraînant sur une seule base ou en combinant les deux. Ces différentes configurations nous ont permis de mieux comprendre

comment les modèles se comportent selon les cas, et d'évaluer leur capacité à s'adapter à des données variées. Dans ce qui suit nous allons présenter les différentes expérimentations réalisées sur les images OCT, ainsi que les résultats obtenus.

2.1.3.1 Entraînement sur la base locale

La première série d'expériences a été réalisée à partir de notre base de données locale, comprenant un total de 17870 images OCT, dont 8935 images représentant des cas d'OMD et 8935 images des cas normaux (après augmentation de données). Cette expérience a pour objectif de tester nos modèles de classification sur des données provenant d'une même source, dans un environnement contrôlé. Nous avons divisé notre base de données sur 3 ensembles : 70% de données pour l'entrainement, 15% pour validation, et 15% pour le test. Et nous avons entrainé et testé nos modèles sur 2 configurations, la première sans augmentation de données et la deuxième avec augmentation de données qui sont des transformations appliquées sur les deux classes (OMD et normal) pour enrichir le jeu de données et améliorer la généralisation des modèles.

Les performances ont été évaluées à l'aide de cinq métriques classiques de classification binaire : Accuracy, Sensibilité (Se), Spécificité (Sp), AUC et F1-score. Les résultats sont regroupés dans le tableau ci-dessous., notamment la sensibilité (Se), la spécificité (Sp), la précision globale (Acc), l'aire sous la courbe ROC (AUC) et le F1-score. Les résultats obtenus sont illustrés dans le Tableau IV-1 ci-dessous.

Tableau IV-1: Performances comparées des modèles de classification sur la base locale, avec et sans augmentation de données

	Sans augmentation					**Avec augmentation**				
Modèle	Acc(%)	Se(%)	Sp(%)	AUC(%)	F1score(%)	Acc(%)	Se(%)	Sp(%)	AUC(%)	F1score(%)
CNN	95.19	95.90	94.48	99.22	95.22	81.88	100	63.76	97.10	84.66
VGG16	98.94	98.73	99.15	99.93	98.95	99.25	98.28	99.74	99.78	98.87
ResNet50	97.13	97.50	96.80	100	97.10	93.14	98	88	99	93
MobileNetV2	97.84	98	97.70	100	97.80	93.92	97.61	90.23	97	94.14
InceptionV3	96.72	96.94	96.50	99.65	96.73	95.82	97.09	94.56	99.42	95.88

Globalement, les résultats obtenus montrent que les performances des modèles sont élevées dans les deux configurations, mais l'effet de l'augmentation varie considérablement selon l'architecture utilisée. Le CNN simple affiche des résultats très satisfaisants sans augmentation,

avec une précision de 95,19 %, une sensibilité de 95,90 % et une AUC de 99,22 %. Toutefois, après augmentation, ses performances chutent significativement : la précision tombe à 81,88 %, la spécificité s'effondre à 63,76 %, et le F1-score diminue fortement. Bien que la sensibilité atteigne 100 %, cela révèle un fort déséquilibre du modèle, qui tend à prédire systématiquement la classe positive (OMD). Ce comportement indique un cas typique de surapprentissage sur la classe dominante, probablement dû à une sensibilité excessive aux transformations appliquées aux images. L'augmentation a donc eu un effet contre-productif sur ce modèle peu complexe.

À l'inverse, VGG16 démontre une stabilité remarquable. Il obtient les meilleurs résultats. Sans transformation, il atteint une précision de 98,94 %, une spécificité de 99,15 %, et une AUC de 99,93 %. L'augmentation ne fait que renforcer sa robustesse, avec une précision portée à 99,25 %, et un AUC de 99,78 %, tout en maintenant un bon équilibre entre sensibilité et spécificité. Ces résultats traduisent la capacité de VGG16 à généraliser efficacement même en présence de données variées, faisant de lui un excellent candidat pour une application en milieu clinique.

ResNet50, de son côté, affiche également de très bonnes performances en configuration de base, avec une précision de 97,13 % et une AUC parfaite de 100 %. Cependant, contrairement à VGG16, l'augmentation induit une légère baisse de ses performances : la précision chute à 93,14 %, et la spécificité diminue à 88 %. Cela suggère une sensibilité du modèle aux modifications introduites par l'augmentation.

MobileNetV2 montre lui aussi un bon comportement, avec une précision initiale de 97,84 %, une AUC de 100 %, et des performances relativement stables après augmentation. Bien qu'une légère baisse de la spécificité soit observée (de 97,70 % à 90,23 %), les scores restent élevés, et le modèle conserve une bonne capacité de classification.

InceptionV3 a montré de très bons résultats, mais l'augmentation de données ne lui a pas vraiment été bénéfique. Sans augmentation, il atteint une précision de 96,72 %, avec une bonne sensibilité et spécificité. Mais après l'ajout des transformations (comme les flips et rotations), sa précision baisse légèrement à 95,82 %. Certes, la sensibilité augmente un peu (97,09 %), ce qui montre que le modèle devient un peu meilleur pour détecter les cas positifs d'OMD, mais cette amélioration se fait au détriment de sa capacité à bien reconnaître les cas normaux. Autrement dit, il devient plus "sensible", mais un peu moins "équilibré".

En conclusion, ces résultats montrent que l'impact de l'augmentation de données dépend fortement du modèle utilisé. Parmi eux, VGG16 se distingue comme le plus fiable et le plus performant, alliant précision, robustesse et équilibre, ce qui en fait un excellent candidat pour une application pratique à la détection automatisée de l'OMD sur des images OCT issues de bases locales.

2.1.3.2 Entraînement sur une base publique

Afin d'évaluer la capacité de généralisation de nos modèles développés, nous avons mené une deuxième série d'expériences en utilisant une base de données OCT publique, rappelons que cette base est composée de 16516 repartis en 2 classes équilibrées (OMD et NORMAL), dont nous avons divisée en 70% pour l'entrainement, 15% pour la validation et 15% pour le test. Comme précédemment, nous avons testé deux configurations : sans augmentation de données, puis avec augmentation, afin d'observer l'impact de l'enrichissement artificiel du jeu d'apprentissage sur les performances. Les résultats obtenus sont illustrés dans le Tableau IV-2 ci-dessous.

Tableau IV-2: Performances comparées des modèles de classification sur la base publique, avec et sans augmentation de données

	Sans augmentation					**Avec augmentation**				
Modèle	Acc(%)	Se(%)	Sp(%)	AUC(%)	F1score(%)	Acc(%)	Se(%)	Sp(%)	AUC(%)	F1score(%)
CNN	88.74	93.77	83.71	95.83	89.28	83.82	100	67.64	97.37	86.07
VGG16	98.13	99.17	97.10	98.14	99.17	98.84	98.81	98.88	99.95	98.84
ResNet50	87.07	90.36	83.79	94	87	93.14	98.21	88.07	99	93.47
MobileNetV2	93.72	96.67	90.78	98	93.90	93.92	97.61	90.23	99	94.13
InceptionV3	93.81	94.68	92.93	98.26	93.86	95.82	97.09	94.56	99.42	95.88

Quand on utilise des images venant d'une base différente, on voit que les modèles profonds comme VGG16 et InceptionV3 arrivent à garder de très bonnes performances. Cela montre qu'ils savent bien s'adapter à de nouvelles images qu'ils n'ont jamais vues, ce qu'on appelle une bonne capacité de généralisation. Globalement, les performances sont satisfaisantes, mais varient selon l'architecture. Le CNN simple montre un comportement assez instable. Sans augmentation, il obtient une précision de 88,74 %, une sensibilité de 93,77 % et une spécificité de 83,71 %, ce qui traduit une meilleure capacité à détecter les cas positifs que les négatifs. Cependant, après

augmentation, la sensibilité atteint 100 %, mais au prix d'une chute de la spécificité à 67,64 %, et d'une baisse globale de la précision à 83,82 %. Ce déséquilibre est dû à un cas de surapprentissage sur la classe positive. Il devient extrêmement sensible, mais perd en capacité à faire la distinction correcte entre les deux classes.

Le modèle VGG16, en revanche, reste très performant et stable dans les deux configurations. Sans augmentation, il atteint déjà d'excellentes performances : 98,13 % de précision, 99,17 % de sensibilité et une AUC de 98,14 %. Après augmentation, ses scores se maintiennent avec une légère amélioration de l'AUC (99,95 %), et des métriques très équilibrées. Cela montre que VGG16 est capable de généraliser efficacement même sur des bases publiques, et que l'augmentation ne le perturbe pas, au contraire, elle renforce sa fiabilité.

ResNet50, de son côté, réalise une amélioration notable après augmentation. Sa précision passe de 87,07 % à 93,14 %, avec une sensibilité atteignant 98,21 % et une AUC de 99 %. Cette progression montre que ce modèle, bien que sensible à la variabilité dans la configuration de base, bénéficie largement de l'augmentation des données lorsqu'il est confronté à une base externe.

MobileNetV2, quant à lui, reste constant dans les deux scénarios. Sa précision passe légèrement de 93,72 % à 93,92 %, tandis que la sensibilité monte de 96,67 % à 97,61 %. Ces résultats indiquent une bonne robustesse et un équilibre intéressant entre performance et efficacité, ce qui confirme son potentiel pour des systèmes embarqués ou mobiles.

Enfin, InceptionV3 montre des résultats élevés et cohérents. Sans augmentation, il atteint une précision de 93,81 %, avec une sensibilité de 94,68 % et une AUC de 98,26 %. Après augmentation, ces valeurs s'améliorent légèrement, avec une précision de 95,82 %, une sensibilité de 97,09 % et une AUC de 99,42 %. Contrairement à ce que nous avons observé sur la base locale, l'augmentation a ici un effet positif, renforçant à la fois la précision globale et l'équilibre du modèle.

En conclusion, cette série d'expériences confirme que l'augmentation de données est bénéfique pour les modèles profonds comme VGG16, ResNet50, InceptionV3 et MobileNetV2, en améliorant leur généralisation sur une base externe. En revanche, pour les architectures simples comme le CNN de base, l'augmentation tend à déséquilibrer le comportement du modèle, au détriment de sa fiabilité. VGG16, ici encore, se distingue par sa stabilité et ses performances constantes, ce qui en fait un excellent choix pour des tâches de classification OCT sur des bases réelles, variées et hétérogènes.

2.1.3.3 Entraînement combiné et test croisé

Dans le but d'augmenter la diversité des données et de favoriser la généralisation du modèle, nous avons réalisé un entraînement combiné en fusionnant les deux bases de données OCT (locale et publique). Après l'entraînement sur l'ensemble unifié, deux évaluations distinctes ont été menées : test sur la base locale et test sur la base publique.

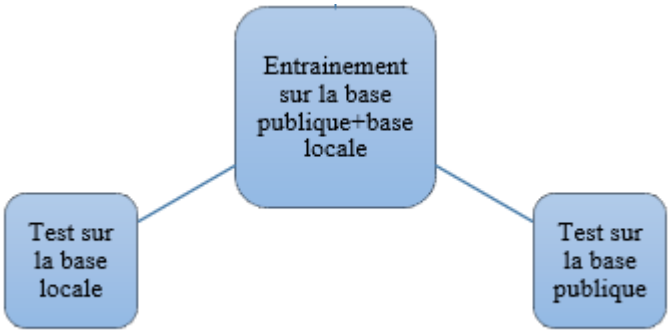


Figure IV.10: Schéma de l'entrainement combiné et le test croisé

En combinant les deux bases d'images (la base locale et la base publique), les modèles de classification sont confrontés à des données plus variées. En effet, chaque base contient des images prises avec des appareils différents, dans des conditions cliniques différentes, et avec des résolutions ou qualités d'image variables. Cette diversité rend l'apprentissage plus riche mais aussi plus difficile, car le modèle doit apprendre à reconnaître les mêmes signes de maladie dans des images qui ne se ressemblent pas toujours.

Cette stratégie a permis d'analyser la robustesse du modèle face à des variations de provenance des images, en évaluant sa capacité à maintenir de bonnes performances lorsqu'il est appliqué séparément à chaque source de données. Les performances ont été mesurées selon les métriques classiques : sensibilité (Se), spécificité (Sp), précision globale (Acc), aire sous la courbe ROC (AUC) et F1-score. Les résultats obtenus sont illustrés dans le Tableau IV-3 ci-dessous.

Tableau IV-3: Résultats des tests croises des modèles de classification sur les bases OCT : locale et publique

	Base locale					**Base publique**				
Modèle	Acc(%)	Se(%)	Sp(%)	AUC(%)	F1score(%)	Acc(%)	Se(%)	Sp(%)	AUC(%)	F1score(%)
CNN	96.02	98.73	93.23	96.19	99.41	87.78	90.94	84.62	95.10	88.15
VGG16	99.24	99.48	99	99.99	99.26	98.59	98.25	98.92	99.89	98.58
ResNet50	96.40	97	96	100	96	86.66	84	90	94	86
MobileNetV2	97.69	98.58	96.77	100	97.74	93.81	95.76	91.85	99	93.93
InceptionV3	97.27	97.17	97.38	99.75	97.31	94.60	93.60	95.59	98.71	94.54

Les résultats montrent que la plupart des modèles conservent de bonnes performances sur les deux ensembles, bien que l'on observe une légère baisse de précision lorsqu'ils sont testés sur la base différente de celle sur laquelle les images ont été collectées à l'origine.

Le CNN simple montre une bonne adaptation sur la base locale avec une précision de 96,02 %, une sensibilité élevée de 98,73 % et un F1-score remarquable de 99,41 %. En revanche, ses performances chutent de façon marquée sur la base publique (précision à 87,78 %, F1-score à 88,15 %), traduisant une moins bonne généralisation vers une source d'images différente. Ce comportement suggère que, bien qu'enrichi, ce modèle reste limité dans sa capacité à gérer des données trop hétérogènes.

À l'inverse, VGG16 se distingue une fois de plus par sa stabilité et son efficacité, avec une précision de 99,24 % sur la base locale et de 98,59 % sur la base publique. Les métriques associées (AUC, F1-score, sensibilité et spécificité) restent très élevées et équilibrées, confirmant que ce modèle est particulièrement adapté à des environnements multi-source.

De manière similaire, MobileNetV2 et InceptionV3 conservent de très bonnes performances sur les deux bases, avec seulement une légère baisse sur la base publique. MobileNetV2 atteint par exemple une précision de 97,69 % sur la base locale, contre 93,81 % sur la base publique, tout en maintenant une bonne balance entre sensibilité et spécificité. Ces résultats confirment la capacité de ces modèles à s'adapter à des variations importantes dans les données tout en conservant leur fiabilité.

Quant à ResNet50, il montre une bonne précision sur la base locale (96,40 %), mais ses performances baissent davantage sur la base publique (86,66 %), ce qui pourrait indiquer une

sensibilité aux différences de distribution entre les deux ensembles. Ce modèle semble légèrement moins stable dans un contexte de généralisation croisée.

En résumé, l'entraînement combiné sur les bases locale et publique a permis d'améliorer la capacité des modèles à apprendre sur des données variées. Toutefois, tous les modèles ne réagissent pas de la même manière face à cette diversité. Les architectures profondes comme VGG16, InceptionV3 et MobileNetV2 tirent le meilleur parti de cette fusion, maintenant des performances élevées même lorsque les images testées sont de différents acquisition. Cela souligne l'importance de l'architecture dans la capacité de généralisation, et valide l'intérêt d'un entraînement multi-source dans des contextes cliniques réels.

Comparaison avec les résultats précèdent :

Les résultats, présentés sous forme de diagrammes en barres, permettent de visualiser l'évolution des performances (Accuracy, Sensibilité, Spécificité, AUC, F1-score) avant et après la fusion des données, et de mesurer l'impact de cette stratégie sur la capacité des modèles à s'adapter à des sources d'images variées.

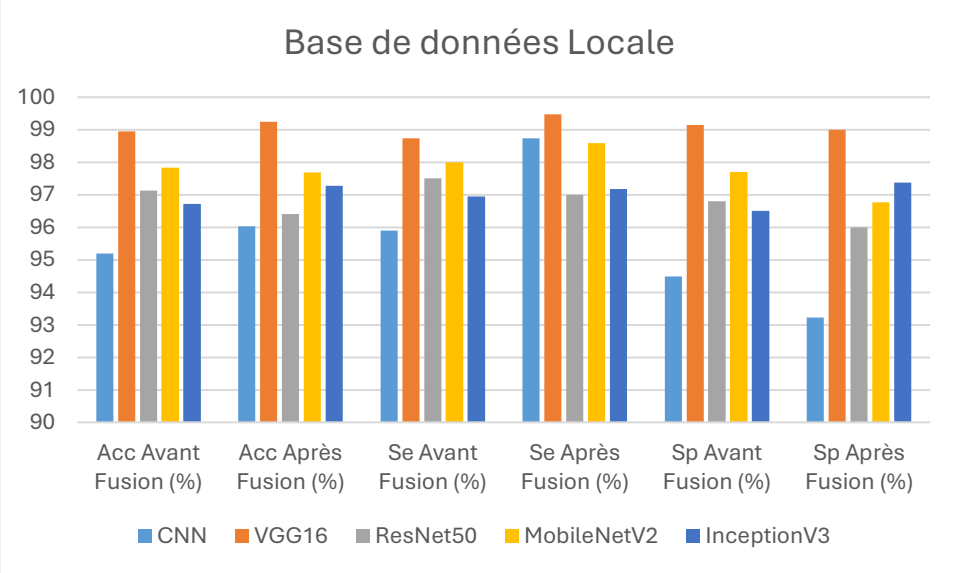


Figure IV.11: Evolution des performances des modèles avant et après fusion des bases OCT (sur la base locale)

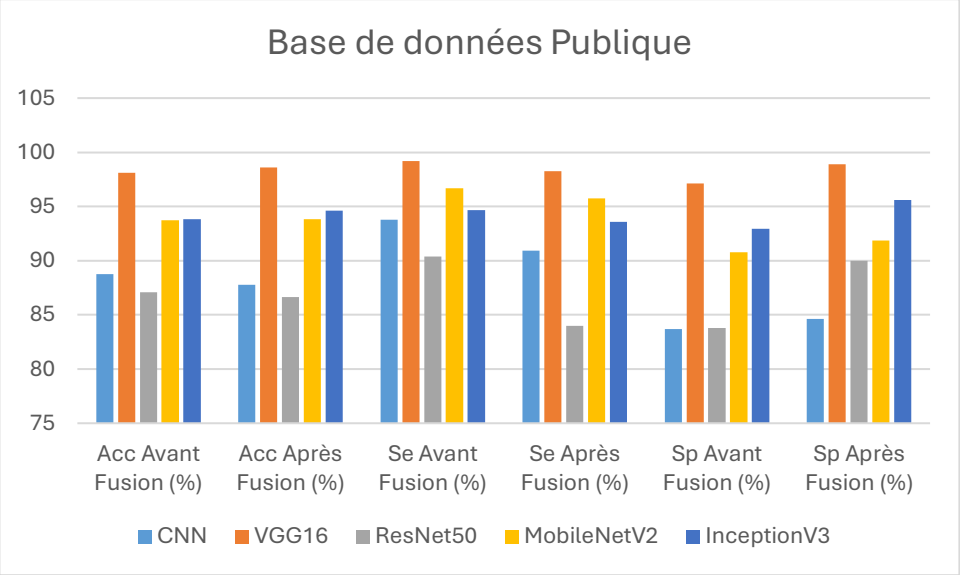


Figure IV.12: Evolution des performances des modèles avant et après fusion des bases OCT (sur la base publique)

Sur la base locale, les résultats montrent que la fusion a eu un effet positif. Les performances des modèles ont généralement augmenté. Par exemple, le modèle CNN est passé d'une précision de 95.19 % (avant fusion) à 96.02 % (après fusion), et sa sensibilité a atteint 98.73 %. Le modèle VGG16, déjà très performant, est passé à 99.24 % de précision. Les autres modèles comme ResNet50, MobileNetV2 et InceptionV3 ont aussi obtenu de bons résultats après fusion. Cela montre que l'ajout de la base publique a enrichi l'apprentissage et permis aux modèles de mieux reconnaître les cas d'OMD, sans nuire à leur performance sur les données locales.

Sur la base publique, l'effet de la fusion est plus variable. Certains modèles, comme VGG16 et InceptionV3, ont gardé de très bonnes performances, proches ou même légèrement supérieures à celles obtenues avant la fusion. Par exemple, VGG16 atteint 98.59 % de précision, presque identique à son résultat d'avant fusion. Cependant, d'autres modèles comme ResNet50 ou CNN ont eu une légère baisse de performance. Le modèle ResNet50 est passé de 93.14 % à 86.66 % de précision, et le CNN a baissé légèrement aussi. Cette baisse peut s'expliquer par les différences entre les deux bases (qualité, résolution, type d'appareil utilisé), qui rendent l'apprentissage plus difficile pour certains modèles.

En résumé, la fusion des deux bases a été bénéfique pour la base locale, en améliorant clairement les résultats des modèles. Pour la base publique, les résultats sont plus partagés : certains modèles ont bien réagi à la fusion, d'autres un peu moins. Cela montre que fusionner plusieurs bases permet de rendre les modèles plus robustes, mais demande aussi une bonne adaptation pour tenir compte

des différences entre les sources de données. Globalement, cette approche reste intéressante et utile dans un contexte réel, où les images médicales peuvent venir de plusieurs hôpitaux ou appareils différents.

2.1.3.4 Comparaison des performances expérimentales avec l'état de l'art

Dans cette section, nous comparons les performances de nos modèles de classification avec celles rapportées dans la littérature. L'objectif est d'évaluer la pertinence de notre approche par rapport aux standards actuels en détection automatique de l'œdème maculaire diabétique (OMD) à partir d'images OCT.

Nos premières expériences ont été réalisées à partir de notre propre base locale, équilibrée entre cas normaux et cas d'OMD. Nous avons évalué les performances de plusieurs architectures, notamment CNN simple, VGG16, ResNet50, MobileNetV2 et InceptionV3. Les résultats montrent que les modèles VGG16 et MobileNetV2 atteignent des performances très élevées, avec une précision globale dépassant 97 % même sans augmentation de données. Après application de techniques d'augmentation, VGG16 atteint 99.25 % de précision, avec une aire sous la courbe ROC (AUC) de 99.78 %. Ces résultats sont largement compétitifs face à ceux rapportés dans la littérature. Par exemple, Feng Li et al Rapportent une précision de 98.80 % avec VGG16 sur une base publique [46]. Notre modèle, appliqué à notre base locale, dépasse ce score. Ce qui témoigne de la qualité de nos images locales ainsi que de la robustesse de notre pipeline de prétraitement et d'apprentissage.

Nous avons ensuite évalué nos modèles sur la base OCT2017 proposée par Kermany et ses collègues, afin de tester leur capacité de généralisation. Là encore, nos modèles, et en particulier VGG16, ont montré des performances très solides. Avec augmentation de données, VGG16 atteint une précision de 98.84 %, une sensibilité de 98.81 %, une spécificité de 98.88 %, et un AUC de 99.95 %.

Plusieurs études antérieures ont été menées sur la base publique OCT2017, proposée par Kermany et al., pour la détection automatique de l'œdème maculaire diabétique (OMD). Toutefois, toutes ces études ont utilisé la version initiale déséquilibrée de cette base de données, dans laquelle les images normales sont largement plus nombreuses que celles présentant un OMD, ce qui peut introduire un biais en faveur de la classe majoritaire.

Par exemple, Kermany et al., auteurs de la base, ont entraîné leur modèle sur cet ensemble déséquilibré et obtenu une précision de 98,20 %, une sensibilité de 96,80 %, une spécificité de 99,60 %, et une AUC de 99,87 % pour une classification binaire entre images normales et OMD [42] .De même, Kaymak et Serener ont utilisé l'architecture AlexNet sur la même base, avec un jeu d'entraînement fortement déséquilibré (26 315 images normales contre 11 348 images OMD) et un test sur 500 images. Leur modèle a atteint une précision de 99,80 %, une sensibilité de 100 % et une spécificité de 99,6 % [82] ,mais sans mention d'un rééquilibrage ou d'une gestion du déséquilibre, ce qui limite la portée réelle de ces performances.

Dans une autre approche, Feng Li et al. (2018) [46] ont utilisé un modèle VGG-16 pré-entraîné, affiné par apprentissage par transfert, sur le même jeu de données déséquilibré. Pour la classification binaire OMD vs NORMAL, ils ont utilisé 500 images de validation (250 par classe) et obtenu une précision, sensibilité et spécificité de 98,8 %, ainsi qu'une AUC de 99,89 % [47]. Bien que leurs résultats soient excellents, leur étude ne mentionne ni équilibrage ni augmentation des données, ce qui soulève les mêmes réserves.

En comparaison, notre étude adopte une démarche plus rigoureuse en équilibrant le jeu de données (8 016 images normales et 8 016 images OMD), ce qui permet une évaluation plus juste et sans biais. De plus, nous avons appliqué une stratégie d'augmentation de données pour renforcer la capacité de généralisation du modèle. En utilisant également un modèle VGG-16 pré-entraîné, notre approche a permis d'atteindre une précision de 98,84 %, une sensibilité de 98,81 %, une spécificité de 98,88 %, et une AUC de 99,95 %. Ces performances, obtenues sur un ensemble équilibré, sont comparables ou supérieures à celles des études précédentes, tout en assurant une meilleure robustesse et fiabilité en condition réelle.

Parmi les travaux récents menés sur la même base, l'étude de Jaimes et al. (2025) a proposé un modèle VGG16 modifié appliqué à une classification multiclasses des pathologies rétiniennes sur des classes équilibrés [45]. Pour la classe OMD, leur modèle a atteint une précision de 98,77 %, une sensibilité de 94,47 %, une spécificité de 97,72 %, Bien que ces résultats soient très satisfaisants, notre approche obtient des performances supérieures sur cette même base de données, confirmant la robustesse et l'efficacité de notre méthode dans la détection de l'œdème maculaire diabétique.

Tableau IV-4: Comparaison entre les performances de notre modèle avec la littérature

Méthode	Précision	Sensibilité	Spécificité
Kermany et al [42]	98,20%	96,80%	99,60%
Kaymak et Serener [82]	99,80%	100%	99,6%
Feng Li et al [46]	98.80%	98.80%	98.80%
Jaimes et al [45]	98.77%	94.47%	97.72%
Notre modèle (VGG16)	**98.84%**	**96,80 %**	**99,60 %**

En conclusion, les résultats obtenus dans notre étude, que ce soit sur base locale, publique ou en condition de fusion, montrent des performances très compétitives, souvent supérieures à celles rapportées dans la littérature. L'utilisation de l'augmentation de données, du fine-tuning, ainsi que l'approche combinée, a permis d'améliorer la robustesse des modèles et leur capacité de généralisation. Nous estimons que cette approche pourrait constituer une base solide pour une future application clinique de détection automatisée de l'OMD.

2.1 Détection de l'OMD par les images rétiniennes

L'imagerie rétiniennes est un outil de référence pour l'évaluation globale de la rétine. Elle permet de capturer des vues en champ large des structures rétiniennes, facilitant la détection de nombreuses anomalies, notamment les signes d'œdème maculaire diabétique (OMD). Dans cette section, nous présentons les méthodes proposées pour détecter l'OMD à partir des images rétiniennes, en suivant une séquence structurée de traitements comprenant le prétraitement, l'extraction de la région d'intérêt, puis l'application d'un algorithme de deep learning permettant de classer les cas en normaux ou pathologiques.

L'organigramme ci-dessous (Figure IV.13) présente notre système développé pour la détection et la classification de l'OMD à partir des images rétinographiques.

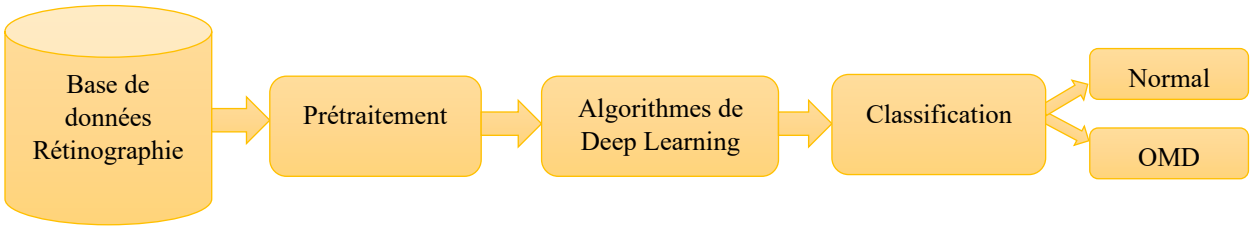


Figure IV.13: Organigramme de nos méthodes pour la détection de l'OMD à partir des images rétiniennes

2.2.1 Prétraitement des images rétiniennes

Avant d'utiliser les images rétiniennes dans nos algorithmes, une étape importante consiste à les préparer afin d'éliminer tout ce qui pourrait gêner l'interprétation ou la détection automatique des lésions. En effet, les images fournies par des dispositifs comme l'Optos couvrent un champ très large, mais contiennent aussi des éléments inutiles en bordure, comme des textes, des logos ou des données du patient. Ces informations ne sont pas utiles pour la détection de l'OMD et risquent d'introduire du bruit dans le modèle. C'est pourquoi nous avons appliqué un prétraitement pour nettoyer les images et nous concentrer uniquement sur la partie centrale de la rétine, celle qui nous intéresse vraiment. Plusieurs étapes de prétraitement ont été appliquées dont nous allons citer par la suite.

1) Suppression des informations indésirable :

Les images rétiniennes issues du dispositif Optos présentent un fond très large et contiennent souvent des éléments visuels non pertinents comme des écritures (informations du patient, date, nom de la clinique, etc.) qui apparaissent en périphérie. Ces éléments peuvent fausser l'analyse ou introduire un biais lors de l'entraînement du modèle. Pour y remédier, comme première étape nous avons isolé la zone centrale utile de l'image (la rétine) à l'aide d'un masque circulaire, excluant ainsi les bordures et les annotations inutiles. Ce recentrage permet de focaliser l'analyse sur la zone pertinente. La figure suivante représente avant et après application du masque.

La Figure IV.14 ci-dessous représente le résultat avant et après du masque circulaire.

(a)

(b)

Figure IV.14: Suppression des informations périphériques dans une image rétiniennes. (a) Image d'origine issue d'un dispositif Optos ;(b) Image recentrée conservant uniquement la zone utile de la rétine.

Après application du masque, certaines marges noires persistent autour de la zone utile et plus précisément dans le fond. Afin de ne conserver que la rétine dans l'image finale, un seuillage global a été appliqué : les pixels d'intensité inférieure à un seuil fixe (15) ont été considérés comme appartenant au fond. Ensuite, les contours présents dans l'image ont été détectés, et seul le plus grand contour correspondant généralement à la rétine a été conservé. Cette segmentation permet de concentrer l'analyse sur la zone utile tout en éliminant les régions inutiles. Le résultat final est représenté dans la Figure IV.15 ci-dessous.

Figure IV.15: Image après prétraitement

2.2.2 Algorithmes d'apprentissage profond pour la détection de l'OMD

2.2.2.1 Augmentation de données

Comme dans le cas des images OCT, cette première étape essentielle a consisté à équilibrer entre les classes NORMAL et OMD de nos deux bases de données Eidon et Optos. Pour corriger ce déséquilibre, nous avons appliqué une augmentation de données ciblée sur les images NORMAL de la base de données Eidon et les images OMD pour la base Optos. Cela nous a permis d'améliorer la représentativité des deux classes et d'éviter que le modèle n'apprenne à favoriser la classe majoritaire. L'augmentation a été réalisée à partir des images prétraitées, en appliquant une série de transformations géométriques et photométriques destinées à simuler la variabilité naturelle des acquisitions.

Ces techniques, illustrées dans les figures ci-dessous (Figure IV.16 et Figure IV.17), ont permis de générer artificiellement de nouveaux échantillons tout en conservant les caractéristiques essentielles de l'imagerie rétinographie, ce qui a renforcé la capacité généralisante de nos modèles.

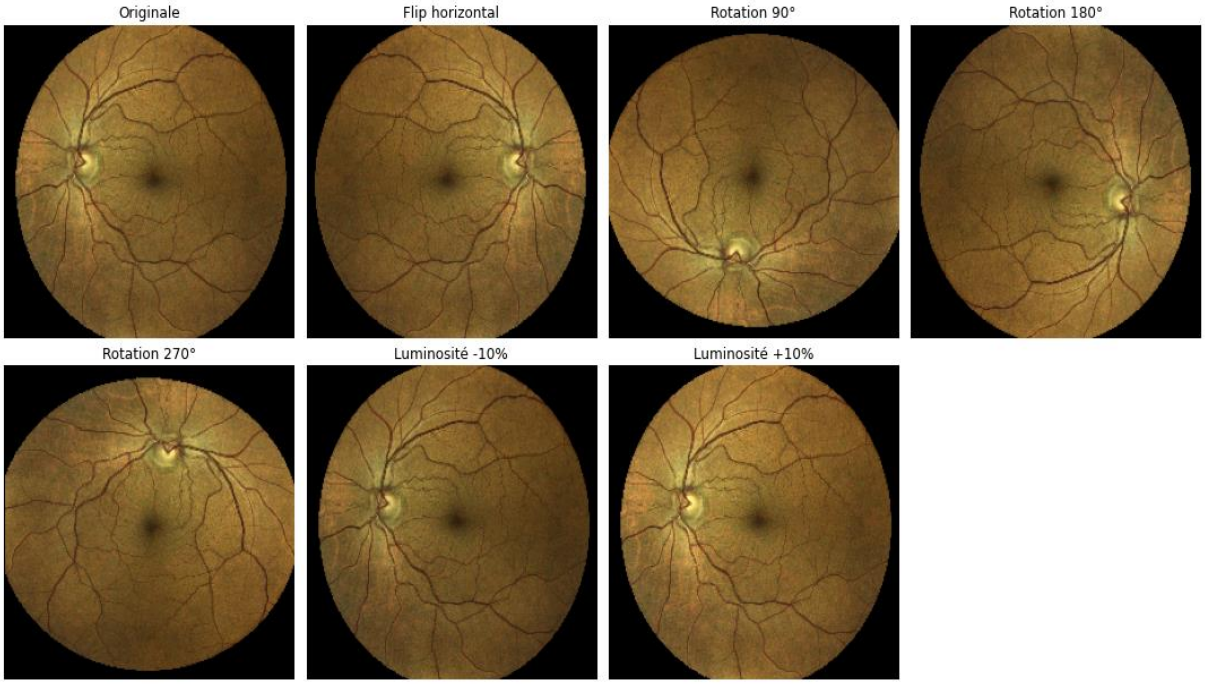


Figure IV.16: Les différents transformations appliquées sur les images issues d'Eidon

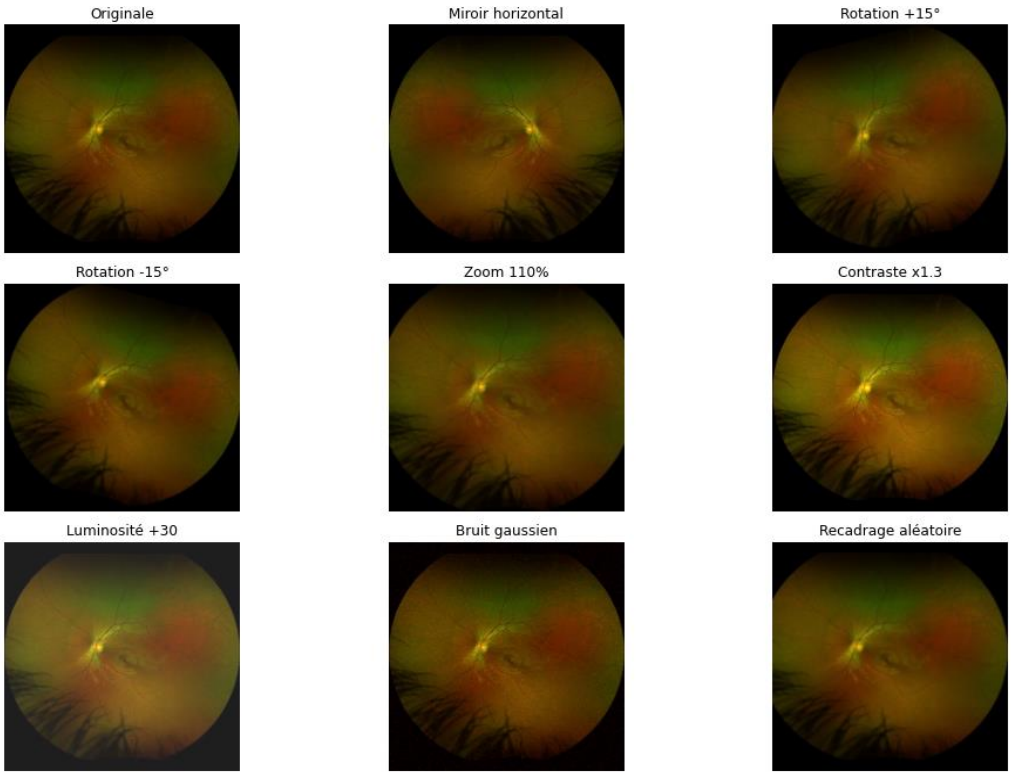


Figure IV.17: Les différents transformations appliquées sur les images issues d'Optos

2.2.2.2 Modèle CNN proposé

Dans le cadre de ce projet, nous avons conçu un réseau de neurones convolutifs (CNN) personnalisé, entraîné depuis zéro, spécifiquement adapté à la classification binaire des images rétinographiques. Le modèle est structuré en quatre blocs convolutionnels successifs, chacun suivi d'une opération de max pooling 2×2, permettant une réduction progressive de la dimension spatiale tout en capturant des caractéristiques de plus en plus abstraites :

- **Bloc 1** : 32 filtres 3×3, activation ReLU
- **Bloc 2** : 64 filtres 3×3, activation ReLU
- **Bloc 3** : 128 filtres 3×3, activation ReLU
- **Bloc 4** : 256 filtres 3×3, activation ReLU

Après l'extraction des caractéristiques, les données sont aplaties via une couche Flatten, puis traitées par une couche dense de 512 neurones avec activation ReLU. Une couche Dropout avec un taux de 50 % est intégrée pour atténuer le sur-apprentissage. Enfin, une couche de sortie avec une seule unité et activation sigmoïde permet de prédire la probabilité d'appartenance à la classe "OMD"ou normal.

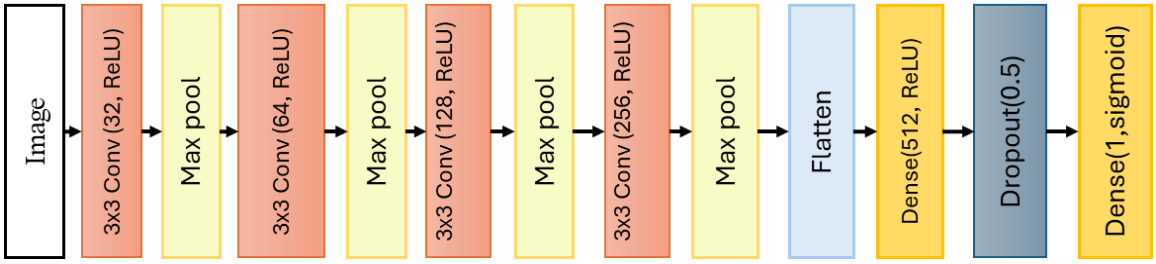


Figure IV.18: Architecture du modèle CNN from scratch pour la classification des images rétiniennes

2.2.2.3 Modèle VGG16 proposé

Nous avons conçu une architecture de classification binaire basée sur le modèle préexistant VGG16, adaptée à notre tâche spécifique. Cette approche repose sur le transfert d'apprentissage, où les couches convolutionnelles de VGG16, préalablement entraînées sur le vaste ensemble de données ImageNet, sont utilisées pour extraire des caractéristiques pertinentes des images d'entrée.

Pour adapter ce modèle à la classification binaire, nous avons modifié la structure en ajoutant des couches personnalisées. Après la couche de pooling global, nous avons intégré une première couche dense de 512 neurones avec activation ReLU, suivie d'une normalisation par lots pour stabiliser l'apprentissage. Une couche de dropout avec un taux de 50% a été ajoutée pour réduire le sur-apprentissage. Ensuite, une deuxième couche dense de 256 neurones avec activation ReLU a été ajoutée, suivie à nouveau d'une normalisation par lots et d'une autre couche de dropout de 50%.

Enfin, une couche de sortie avec une activation sigmoïde a été utilisée pour produire une probabilité entre 0 et 1, indiquant l'appartenance à l'une des deux classes. Par ailleurs, seules les huit dernières couches convolutionnelles de VGG16 ont été rendues entraînables afin de permettre un ajustement fin tout en conservant les connaissances préalables du modèle.

L'optimiseur Adam, avec un taux d'apprentissage de 1e-5, a été choisi pour sa capacité à converger efficacement. Cette architecture modifiée a montré une amélioration notable des performances par rapport à l'utilisation directe de VGG16, notamment en termes de précision, de sensibilité et de robustesse, ce qui la rend particulièrement adaptée à la classification binaire dans des applications spécifiques.

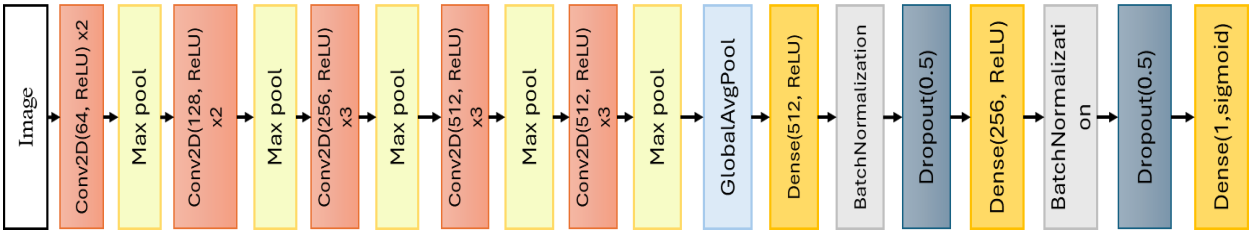


Figure IV.19: Architecture du modèle VGG16 modifié pour la classification des images rétiniennes

2.2.2.4 Modèle EfficientNetB0 proposé

Notre modèle EfficientNetB0 a été modifié pour s'adapter à la tâche de classification binaire (OMD vs normal) sur des images rétiniennes. Bien que le modèle de base EfficientNetB0 soit pré-entraîné sur ImageNet, nous avons ajusté sa structure pour répondre à nos besoins spécifiques.

Tout d'abord, la tête du modèle a été modifiée : la couche de sortie a été remplacée par une couche sigmoïde pour la classification binaire, contrairement à la softmax utilisée pour la classification multi-classes dans le modèle de base.

Ensuite, nous avons ajouté des couches denses après la couche de Global Average Pooling, avec respectivement 512 et 256 neurones, suivies de Batch Normalization et de Dropout pour améliorer la régularisation et éviter le sur-apprentissage. Ces ajouts permettent au modèle de mieux capter les caractéristiques pertinentes des images rétiniennes dans le contexte du diagnostic d'œdème maculaire diabétique. Enfin, l'optimiseur Adam a été utilisé avec un taux d'apprentissage de 1e-4, ajusté pour garantir un apprentissage stable et progressif pendant l'entraînement.

Ainsi, bien que basé sur l'architecture EfficientNetB0, notre modèle a été modifié pour mieux correspondre aux spécificités de la classification binaire dans le cadre de la détection d'OMD (figure IV.20).

2.2.2.5 Modèle MobileNetV2 proposé

Dans notre architecture de MobileNetV2 modifié, nous avons retiré la couche de classification initiale et l'avons remplacée par une couche de sortie sigmoïde pour la classification binaire. Ainsi, après la couche de Global Average Pooling, nous avons ajouté plusieurs couches denses, dont une avec 512 neurones et une autre avec 256 neurones, chaque couche étant suivie d'une Batch Normalization et d'une couche Dropout à 50% pour favoriser la régularisation et éviter le sur-apprentissage. Nous avons utilisé un optimiseur Adam avec un taux d'apprentissage de 1e-4 pour

un entraînement stable et progressif. De plus, nous avons intégré le callback ReduceLROnPlateau, pour ajuster dynamiquement le taux d'apprentissage, afin de garantir que le modèle ne surapprenne pas les données (figure IV.21).

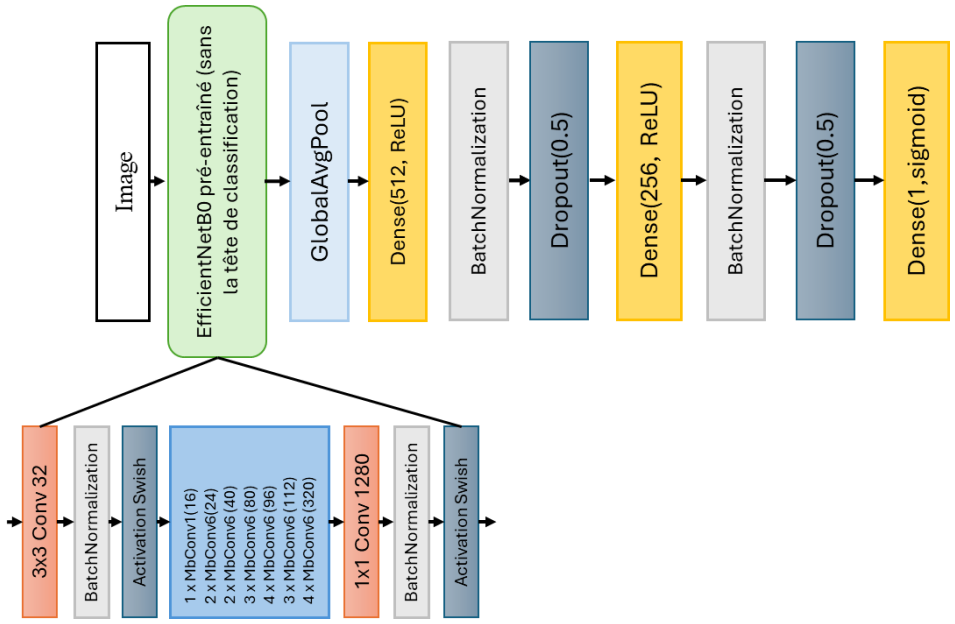


Figure IV.20: Architecture du modèle EfficientNetB0 modifié pour la classification des images rétiniennes

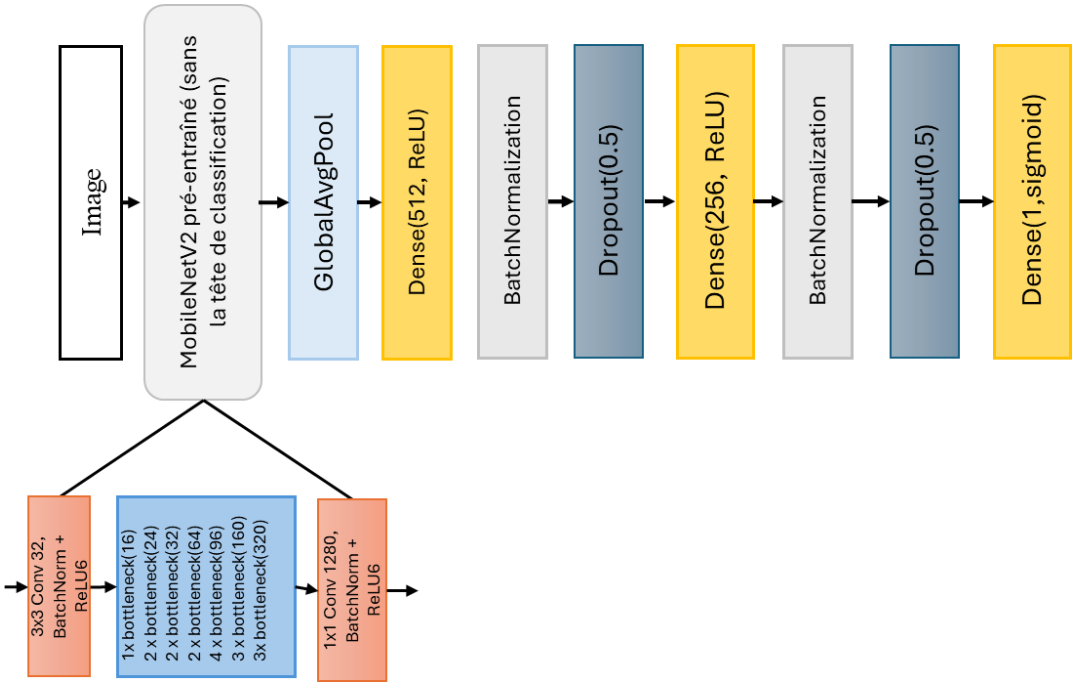


Figure IV.21: Architecture du modèle MobileNetV2 modifié pour la classification des images rétiniennes

2.2.2.6 Modèle InceptionV3 proposé

Dans cette architecture, nous avons utilisé le modèle InceptionV3 pré-entraîné sur ImageNet, avec la tête de classification personnalisée. Après la couche de GlobalAveragePooling2D, deux couches denses ont été ajoutées : l'une avec 512 neurones et l'autre avec 256 neurones, chacune suivie d'une BatchNormalization et d'un Dropout à 50 %. Ces couches permettent de mieux ajuster les caractéristiques extraites par le modèle de base et de réduire le surapprentissage. La sortie du modèle est une couche dense avec 1 neurone et une activation sigmoïde, adaptée à la classification binaire (OMD vs normal). Le modèle est compilé avec l'optimiseur Adam (learning rate = 1e-4) et la fonction de perte binary_crossentropy. Nous avons utilisé un callback ReduceLROnPlateau pour ajuster dynamiquement le taux d'apprentissage en cas de stagnation de la performance sur la validation.

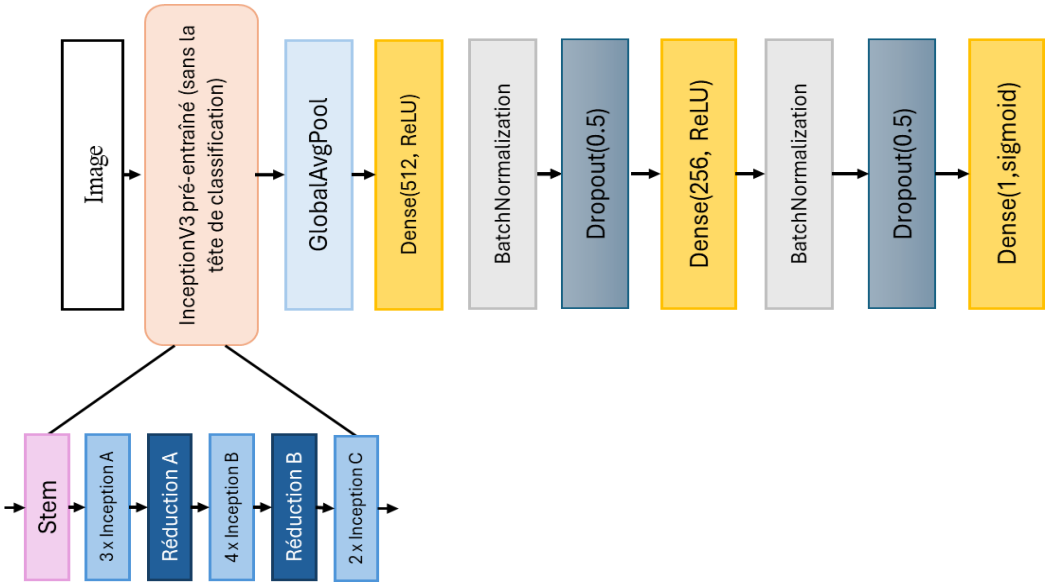


Figure IV.22: Architecture du modèle InceptionV3 modifié pour la classification des images rétiniennes

2.2.3 Expérimentations sur les rétinographies

En complément des expériences menées sur les images OCT, nous avons étudié la détection de l'OMD issues de la rétinographie couleur. Pour cela, nous avons conçu des algorithmes spécifiques, mieux adaptés aux caractéristiques visuelles de ce type d'images. Les données utilisées proviennent de nos deux bases de données locaux : Optos, et Eidon. Ces deux sources nous ont permis de couvrir une diversité de cas cliniques. Ensuite nous avons évalué nos modèles dans plusieurs configurations : d'abord sur chaque base séparément (avec et sans augmentation de

données), puis en combinant les deux pour entraîner un modèle commun. Cette approche visait à tester la robustesse des algorithmes face à des variations d'origine instrumentale et de qualité d'image.

2.2.3.1 Expérimentations sur la base locale Optos

La première série d'expériences a été réalisée à partir de notre base de données locale acquise avec le système Optos qui permet une visualisation périphérique du fond d'œil, qui est essentielle dans la détection de l'œdème maculaire diabétique (OMD). L'ensemble des données a été divisé en 70 % pour l'entraînement, 15 % pour la validation et 15 % pour le test.

Deux configurations ont été explorées : un premier entraînement sans augmentation des données, suivi d'un second où une augmentation artificielle a été appliquée à toutes les images d'entrainement (OMD et normales), dans le but d'enrichir la diversité des exemples d'apprentissage et de renforcer la capacité de généralisation des modèles. Les résultats obtenus sont représentés dans le Tableau IV-5 suivant.

Tableau IV-5: Performances comparées des modèles de classification sur la base locale Optos, avec et sans augmentation de données

	Sans augmentation					**Avec augmentation**				
Modèle	Acc(%)	Se(%)	Sp(%)	AUC(%)	F1score(%)	Acc(%)	Se(%)	Sp(%)	AUC(%)	F1score(%)
CNN	98.83	99.21	98.46	99.67	98.81	99.61	100	99.23	100	99.60
VGG16	99.22	99.23	99.23	99.23	99.22	98.83	97.60	100	98.82	98.80
EfficientNetB0	99.61	100	99.23	100	99.60	99.22	99.21	99.23	99.21	99.99
MobileNetV2	75.39	100	51.54	99.89	80	99.22	100	98.46	100	99.21
InceptionV3	100	100	100	100	100	100	100	100	100	100

Les résultats montrent que sans augmentation, les modèles profonds tels qu'InceptionV3, EfficientNetB0 et VGG16 obtiennent d'excellentes performances, avec notamment InceptionV3 atteignant un score parfait de 100 % sur toutes les métriques. Notre CNN personnalisé, bien que plus simple, affiche également de très bons résultats avec une précision de 98,83 % et une sensibilité supérieure à 99 %, ce qui témoigne de sa bonne capacité à apprendre sur une base homogène. En revanche, MobileNetV2 présente une forte sensibilité de 100 % mais une spécificité limitée à 51,54 %, suggérant qu'il détecte efficacement les cas OMD mais a du mal à éviter les faux positifs.

L'introduction de l'augmentation des données a globalement renforcé les performances des modèles, particulièrement ceux dont l'architecture est plus légère. MobileNetV2 améliore notablement son équilibre, avec une spécificité qui passe à 98,46 % et une précision globale de 99,22 %. Le CNN personnalisé bénéficie aussi de cette diversité accrue, atteignant 99,61 % de précision et un F1-score de 99,60 %. Quant aux modèles plus profonds, EfficientNetB0 conserve ses performances élevées et InceptionV3 maintient ses résultats parfaits, confirmant leur robustesse face à la variation des données. Seul VGG16 enregistre une légère baisse, restant toutefois très performant avec plus de 98 % sur toutes les mesures.

Ainsi, ces expérimentations soulignent que les modèles profonds tels qu'InceptionV3 et EfficientNetB0 sont capables de bien généraliser sur notre base locale, même sans recours à l'augmentation. Toutefois, l'ajout de données artificielles apporte un avantage certain aux modèles plus simples, améliorant leur capacité à gérer la diversité des images et à réduire les erreurs de classification. Cela démontre l'importance de la richesse et de la variabilité des données dans la formation de modèles de détection d'OMD performants et robustes.

2.2.3.2 Expérimentations sur la base locale Eidon

La deuxième série d'expérimentations a été menée sur des images issues du dispositif Eidon, également local, qui fournit des rétinographies de haute résolution centrées sur la région maculaire. Ce type d'image, plus focalisé, offre une autre perspective pour la détection de l'OMD, souvent centrée sur les détails de la fovéa. Un entraînement et un test ont été réalisés exclusivement sur cette base, sans augmentation de données selon le même protocole que celui appliqué à la base Optos. Cela a permis de comparer les performances des modèles sur des images de type et de qualité différentes. Les résultats obtenus sont représentés dans le Tableau IV-6 suivant.

Tableau IV-6: Performances comparées des modèles de classification sur la base locale Eidon, avec et sans augmentation de données

	Sans augmentation					**Avec augmentation**				
Modèle	Acc(%)	Se(%)	Sp(%)	AUC(%)	F1score(%)	Acc(%)	Se(%)	Sp(%)	AUC(%)	F1score(%)
CNN	89.15	89.92	88.37	96.00	89.23	86.82	89.92	83.72	94.67	87.22
VGG16	93.41	90.70	96.12	97.72	93.23	92.25	91.40	93	92.20	92
EfficientNetB0	94.19	92.25	96.12	97.93	94.07	95.74	96.12	95.35	98.61	95.75
MobileNetV2	83.33	68.22	98.45	95.50	80.37	79.07	59.69	98.47	93.74	74.4
InceptionV3	93.02	93.02	93.02	98.38	93.02	94.96	91.47	98.45	98.52	94.78

Les résultats montrent que les modèles EfficientNetB0 et InceptionV3 obtiennent les meilleures performances, avec des précisions supérieures à 93 % sans augmentation, et encore améliorées avec augmentation, notamment EfficientNetB0 qui atteint 95.74 % de précision et un F1-score de 95.75 %. Ces modèles profonds réussissent à extraire efficacement des caractéristiques pertinentes malgré la complexité des images centrées sur la macula. VGG16 se place également en bonne position, avec des performances solides dépassant les 90 % dans la plupart des métriques.

À l'inverse, le modèle MobileNetV2 affiche des résultats nettement inférieurs, surtout en sensibilité (59.69 % avec augmentation), indiquant des difficultés à détecter correctement les cas d'OMD dans ce contexte. L'augmentation des données n'a pas permis d'améliorer ce modèle, qui semble moins adapté à la variabilité et la complexité des images Eidon. Quant au CNN personnalisé, ses performances sont correctes mais restent modestes comparées aux architectures plus profondes, notamment après augmentation, où on observe même une légère baisse de précision.

Comparativement à la base Optos, les performances obtenues sur la base Eidon sont légèrement inférieures. Cela peut s'expliquer par une taille d'échantillon plus restreinte, une variabilité accrue des images ou encore la focalisation maculaire plus prononcée des images Eidon, rendant la détection plus subtile. Malgré cela, les architectures EfficientNetB0 et InceptionV3 ont maintenu de très bons résultats, avec des précisions respectives de 95.74 % et 94.96 %, et des AUC proches de 99 %. À l'inverse, MobileNetV2 s'est révélé peu fiable, n'atteignant que 59.69 % de sensibilité, ce qui indique un grand nombre de cas pathologiques non détectés. Cela confirme les limites de cette architecture dans des contextes cliniques exigeants sans stratégie d'optimisation adaptée.

2.2.3.3 Fusion des bases Optos et Eidon

Dans cette dernière expérimentation, les modèles ont été entraînés sur un ensemble combiné issu de la fusion des bases locales Optos et Eidon, puis testés séparément sur chacune des deux bases pour évaluer leur capacité de généralisation à des sources d'images distinctes. Cette approche simule un scénario réaliste dans lequel un modèle est exposé à des images variées, tant en termes de qualité que de centrage anatomique. Les résultats obtenus sont représentés dans le Tableau IV-7 suivant.

Tableau IV-7: Résultats des tests croisés des modèles de classification sur les bases retinographie : Optos et Eidon

	Base Optos					Base Eidon				
Modèle	Acc (%)	Se (%)	Sp (%)	AUC (%)	F1score (%)	Acc (%)	Se (%)	Sp (%)	AUC (%)	F1score (%)
CNN	98.43	96.03	94.57	99.03	95.28	92.64	93.80	92.25	98.20	93.03
VGG16	97.25	98.41	95.35	99.53	96.88	95..35	94.57	92.25	98.73	93.49
EfficientNetB0	98.82	99.21	99.22	99.23	99.21	96.12	96.90	94.57	99.30	95.79
MobileNetV2	99.22	98.41	98.45	99.88	98.41	91.86	90.70	96.90	97.95	93.60
InceptionV3	98.22	98.41	99.22	99.93	98.80	96.90	94.57	96.12	99.10	95.31

Afin d'évaluer l'impact de la fusion des bases de données Optos et Eidon sur les performances des différents modèles, nous avons comparé les résultats obtenus avant et après fusion, pour chaque base et pour chaque architecture testée. Les diagrammes suivants résument ces comparaisons en indiquant si les performances se sont améliorées, maintenues stables ou ont légèrement diminué après l'entraînement sur la base fusionnée.

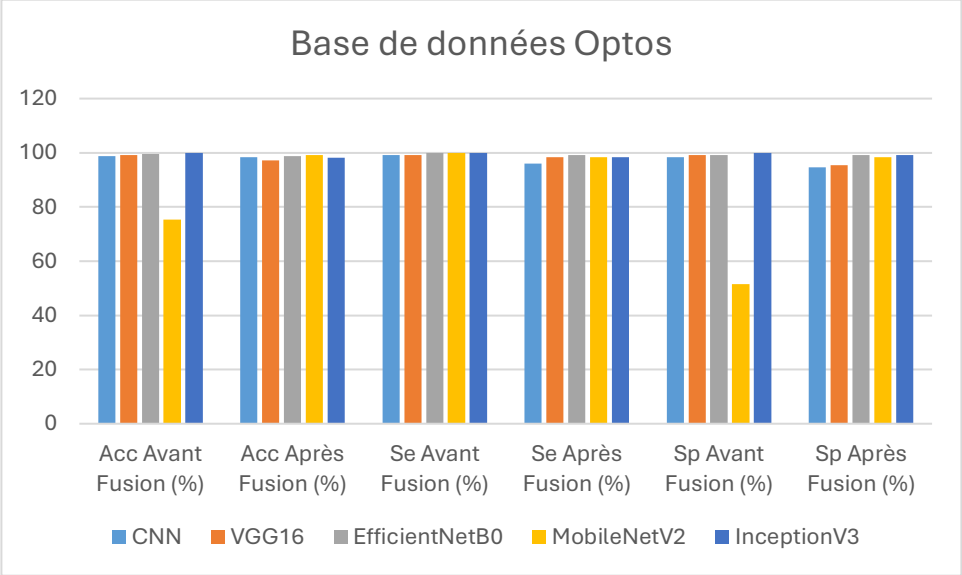


Figure IV.23: Evolution des performances des modèles avant et après fusion des bases rétinographie (sur la base Optos)

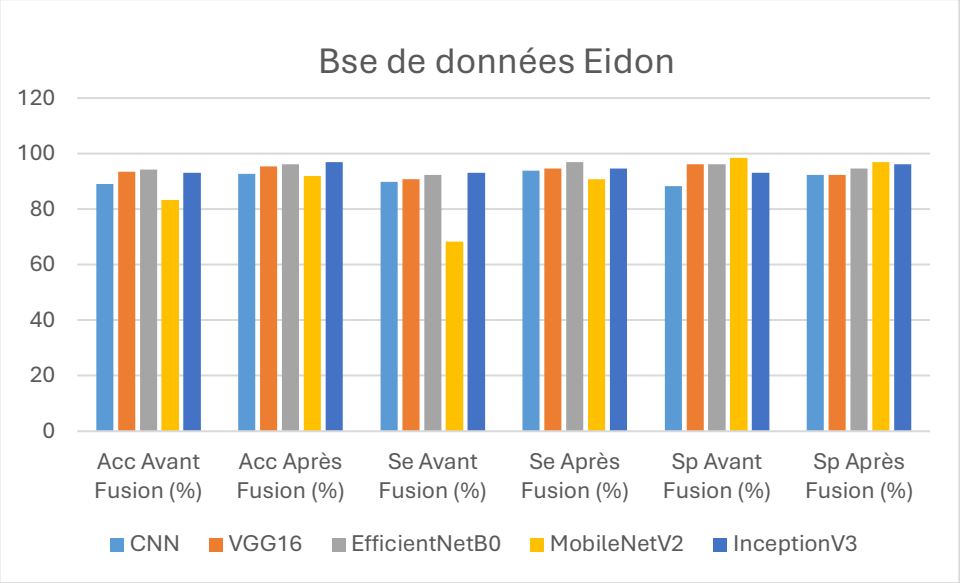


Figure IV.24: Evolution des performances des modèles avant et après fusion des bases rétinographie (sur la base Eidon)

Les résultats démontrent l'intérêt d'un entraînement multi-source pour améliorer la généralisation. Lorsqu'on teste le modèle entraîné sur les deux bases et qu'on l'évalue séparément sur chacune d'elles, on constate des performances élevées, notamment sur la base Optos où les modèles CNN, EfficientNetB0 et InceptionV3 obtiennent des précisions supérieures à 98 %, ce qui indique que l'entraînement avec des images variées n'a pas dégradé la capacité à reconnaître les caractéristiques propres aux images Optos. Sur la base Eidon, les performances restent très bonnes, en particulier pour EfficientNetB0 et InceptionV3, avec des précisions proches de 97 % et des AUC très élevés. MobileNetV2, qui avait montré des faiblesses précédemment, parvient cette fois à de meilleurs résultats sur la base Eidon, ce qui montre que l'exposition à des images de différents types durant l'entraînement permet d'enrichir la représentation interne du modèle et d'améliorer ses performances sur des cas plus complexes. Ces observations confirment que la fusion des bases constitue une stratégie efficace pour développer des modèles robustes et transférables à différents dispositifs d'imagerie rétinienne.

Globalement, ces résultats démontrent que l'entraînement sur une base fusionnée améliore significativement la robustesse des modèles, en particulier pour ceux qui possèdent une architecture suffisamment profonde ou optimisée. Cela met en évidence l'importance d'un apprentissage multi-source pour développer des systèmes de détection d'OMD performants et généralisables en milieu clinique hétérogène.

2.2.3.4 Comparaison des performances expérimentales avec l'état de l'art

Les modèles développés dans ce travail ont été évalués sur des bases d'images provenant des dispositifs Optos et Canon Eidon, puis comparés aux résultats de recherches récentes afin de mesurer leur efficacité par rapport aux méthodes existantes pour la détection de l'œdème maculaire diabétique (OMD).

Parmi tous les modèles testés, InceptionV3, appliqué aux images Optos avec augmentation de données, a obtenu les meilleures performances. Il a atteint 100 % de précision, sensibilité, spécificité, F1-score et AUC, ce qui signifie que le modèle a correctement détecté tous les cas, sans aucune erreur. Ce résultat est supérieur à ceux rapportés dans plusieurs études récentes. Par exemple, Li et al. (2022) ont obtenu une précision de 97,4 % et une AUC de 99,4 % sur une grande base de données [51], tandis que Guo et al. (2022) ont rapporté une précision de 96 % sur la base Messidor [53]. Ces résultats montrent que notre modèle peut surpasser ceux de la littérature.

Le modèle EfficientNetB0, testé sur les images Eidon, a également montré de très bons résultats, avec une précision de 95,74 %, une sensibilité de 96,12 %, une spécificité de 95,35 %, un F1-score de 95,75 % et une AUC de 98,61 %. Ces performances sont proches de celles rapportées par Fu et al. (2022), qui ont obtenu une précision de 97,06 %, mais une sensibilité plus faible (88,64 %) [54], ce qui signifie que leur modèle a manqué davantage de cas d'OMD. À l'inverse, notre modèle détecte mieux les cas positifs, ce qui est très important pour un système de dépistage.

De plus, une étude menée par Manikandan et al. (2023) a montré que la sensibilité moyenne des modèles de détection d'OMD est d'environ 94 % [50]. Les résultats obtenus ici sont donc supérieurs à cette moyenne, ce qui confirme la fiabilité de nos modèles, même lorsqu'ils sont appliqués à des images issues de différents appareils.

En résumé, les modèles que nous avons développés, en particulier InceptionV3 pour les images Optos et EfficientNetB0 pour les images Eidon offrent des performances très solides, parfois meilleures que celles rapportées dans les publications récentes. Ces bons résultats s'expliquent par la qualité du prétraitement des images, l'utilisation d'augmentations adaptées, et le choix d'architectures efficaces. Cela montre que ces modèles pourraient être utilisés dans des systèmes réels de dépistage automatisé de l'OMD, en complément de l'examen clinique.

3. Conclusion

Dans ce chapitre, nous avons présenté en détail la mise en œuvre de notre système de détection automatique de l'œdème maculaire diabétique (OMD) à partir d'images rétiniennes. Plusieurs architectures d'apprentissage profond ont été testées, telles que CNN, VGG16, ResNet50, MobileNetV2, InceptionV3 et EfficientNetB0, en appliquant des techniques de fine-tuning, d'augmentation de données, et des stratégies d'évaluation rigoureuses.

Les résultats obtenus montrent que les performances de classification varient selon les modalités d'images (OCT, rétinographie) et les conditions d'acquisition (base locale vs base publique). Les modèles convolutifs profonds ont généralement fourni de bons résultats, notamment en termes de sensibilité et d'AUC, confirmant leur capacité à détecter les signes d'OMD de manière fiable. L'utilisation d'images issues de notre base locale a permis de tester le système dans des conditions cliniques réelles, tout en soulignant les défis liés à l'hétérogénéité des données.

Nous avons également observé que l'augmentation des données et le prétraitement des images améliorent la robustesse des modèles. Par ailleurs, la combinaison de données multimodales (OCT + rétinographie) ouvre des perspectives intéressantes pour le renforcement de la précision du diagnostic assisté.

Ce chapitre met ainsi en évidence la pertinence des approches par apprentissage profond dans le contexte du dépistage de l'OMD, tout en soulignant la nécessité d'optimisations continues, notamment en matière de généralisation, de traitement d'images bruitées et d'explicabilité des décisions.

Conclusion générale

L'œdème maculaire diabétique (OMD) est une complication fréquente de la rétinopathie diabétique et représente l'une des principales causes de perte de vision chez les patients atteints de diabète. Son diagnostic précoce est essentiel pour éviter une dégradation irréversible de l'acuité visuelle. Dans ce contexte, le recours à des systèmes d'aide au diagnostic automatisés basés sur l'intelligence artificielle (IA) constitue une voie prometteuse, notamment pour les environnements où l'accès à un ophtalmologiste est limité.

Le travail présenté dans ce mémoire consistait à développer un système de détection automatique de l'OMD à partir d'images médicales rétiniennes (OCT et rétinographie), en s'appuyant sur des techniques avancées de traitement d'images et d'apprentissage profond. Une des contributions majeures de ce projet est la collecte et l'annotation d'une base de données locale, constituée d'images rétiniennes acquises en milieu clinique réel. Cette base a permis de travailler sur des données représentatives des conditions d'acquisition courantes dans la pratique ophtalmologique.

Le système proposé repose sur plusieurs étapes successives : prétraitement des images, extraction automatique des régions d'intérêt, entraînement de modèles de deep learning, et classification binaire (OMD / normal). Plusieurs architectures ont été expérimentées, dont CNN, VGG16, ResNet50, MobileNetV2 et EfficientNetB0. Les performances des modèles ont été évaluées sur la base locale ainsi que sur des bases publiques, selon des métriques standards (accuracy, sensibilité, spécificité, AUC, F1-score), et ont montré des résultats satisfaisants, témoignant de la robustesse et du potentiel de généralisation du système.

Malgré ces résultats encourageants, certaines limites doivent être soulignées. En premier lieu, la base locale, bien que précieuse, présente un volume relativement restreint de données, et certaines informations cliniques complémentaires (par exemple, l'évolution temporelle des cas, ou des données fonctionnelles comme l'acuité visuelle) n'ont pas pu être intégrées en raison d'un manque de quelques données lors de la collecte. De plus, la diversité des appareils d'imagerie utilisés engendre une variabilité qui peut influencer la stabilité des modèles. Ces éléments limitent partiellement la portée des résultats obtenus.

Les résultats obtenus sont encourageants et ouvrent de nombreuses voies d'amélioration et de recherche pour les années à venir.

Perspectives

Malgré les progrès réalisés au cours de ce mémoire, il reste encore beaucoup à faire. La recherche continue pour trouver de nouvelles solutions et améliorer les systèmes de dépistage afin qu'ils soient plus efficaces et plus fiables. Les travaux futurs qui peuvent être ajoutés pour enrichir ce travail sont :

- Enrichir la base de données actuelle en augmentant le nombre d'images OCT et rétinographiques.
- Enrichir la base collectée en incluant des données numériques (telle que l'acuité visuelle ou la tension intra-oculaire)
- Développement d'un modèle capable d'exploiter simultanément deux modalités d'imagerie l'OCT et la rétinographie pour établir un diagnostic automatique de l'œdème maculaire diabétique (OMD).
- L'extension du système pour permettre la classification des différents stades de l'OMD, et non plus uniquement une classification binaire (normal / pathologique).
- Classification multi-pathologies en mono- et multimodalité (multilabel)
- L'intégration de techniques de segmentation automatiques permettant d'isoler précisément la macula, les kystes ou les exsudats dans les images OCT ou UWF.
- La collecte de bases de données plus étendues, incluant des images issues de plusieurs dispositifs et couvrant un large éventail de cas normaux, OMD, et d'autres pathologies rétiniennes.
- Conception d'une interface utilisateur intuitive.

En résumé, ce travail constitue une première étape vers l'intégration de l'intelligence artificielle dans la détection automatisée de l'OMD.

Bibliographie

[1] Federation, International Diabetes, «IDF Diabetes Atlas,» International Diabetes Federation, Brussels, Belgium, 2025.

[2] P.-Y. R. R. Paul Robert, «Anatomie et physiologie de l'œil,» *Actualités Pharmaceutiques,* vol. 61, n° %1620, pp. 16-20, 2022.

[3] M. Fadi, «La prise en charge de l'œdème maculaire diabétique:physiopathologie, thérapies et conseils,» CCSD, Marseille, 2022.

[4] R. T. ADDO, «Ocular drug delivery: advances, challenges and applications,» *Springer,* 2016.

[5] C. E. P. D. F. S. e. a. WILLOUGHBY, «Anatomy and physiology of the human eye: effects of mucopolysaccharidoses disease on structure and function–a review,» *Clinical& Experimental Ophthalmology,* vol. 38, pp. 2-11, 2010.

[6] S. Yasmine, «ANALYSE AUTOMATIQUE DES IMAGES RETINIENNES NUMERIQUES POUR L'AIDE AU DIAGNOSTIC DE LA RETINOPATHIE DIABETIQUE,» UNIVERSITE DE SAAD DAHLED DE BLIDA Faculté de Technologie departement d'electronique, BLIDA, 2015.

[7] E. G. L. J. e. P. L. Francine Behar-Cohen, «Anatomie de la rétine,» *Med Sci(Paris),* vol. 36, n° %16-7, pp. 594-599, 2020.

[8] É. BENOIT-BÉLANGER, «Homéostasie des fluides et pathophysiologies dans la rétine: l'épithélium rétinien pigmenté,» Université de Montréal, Montréal, 2023.

[9] S. D. F. K. e. a. Borrelli E, «OCT angiography and evaluation of the choroid and choroidal vascular disorders.,» *Prog Retin Eye Res ,* vol. 67, pp. 30-55, 2018.

[10] societe francaise d'ophtalmologie, «Médicaments et biothérapies en ophtalmologie,» Elsevier Masson, 2023. [En ligne]. Available: https://www.sfo-online.fr/files/rapports-sfo/2023/sforender/B9782294770203000011.html?utm_source=chatgpt.com. [Accès le 10 04 2025].

[11] D. P. L. A. R. e. a. Fields MA, «Interactions of the choroid, Bruch's membrane, retinal pigment epithelium, and neurosensory retina collaborate to form the outer blood-retinal-barrier,» *Prog Retin Eye Res,* vol. 76, p. 100803, 2020.

[12] R. C. A. D. Díaz-Coránguez M, «The inner blood-retinal barrier : Cellular basis and development,» *Vision Res 2017,* vol. 139, pp. 123-37, 2017.

[13] K. G. K. K. V. V. S. R. B. U. K. R. V. &. J. G. M. Selvaraj, «Current treatment strategies and nanocarrier based approaches for the treatment and management of diabetic retinopathy,» *Journal of Drug Targeting,* vol. 25, n° %15, p. 386–405, 2017.

[14] M. E. M. C. N. Khaldi, «Proliferative diabetic retinopathy inauguring gestational diabetes,» *Annales d'Endocrinologie,* vol. 69, n° %15, pp. 449-452, 2008.

[15] Federation, International Diabetes, «Diabetes global report 2000-2050,» IDF Diabetes Atlas , 2024. [En ligne]. Available: https://diabetesatlas.org/data-by-location/global/?. [Accès le 16 Avril 2025].

[16] M. Joanne W.Y. Yau, M. Sophie L. Rogers, P. Ryo Kawasaki, P. Ecosse L. Lamoureux, P. Jonathan W. Kowalski, P. Toke Bek, P. Shih-Jen Chen, P. Jacqueline M. Dekker, P. Astrid Fletcher et M. Jakob Grauslund, «Global Prevalence and Major Risk Factors of Diabetic Retinopathy,» *Diabetes Care,* vol. 35, n° %13, p. 556–564, 2012.

[17] F. F. 3. K. R. e. a. Wilkinson CP, «Proposed international clinical diabetic retinopathy,» *Ophthalmology,* vol. 110, n° %19, pp. 1677-1682, 2003.

[18] « Early Treatment Diabetic Retinopathy Study Research Group. Classification of diabetic retinopathy,» *Ophthalmology,* vol. 98, n° %15, pp. 807-822, 1991.

[19] P. Massin, «Œdème maculaire diabétique :une manifestation locale d’une maladie systémique,» *Les Cahiers d'Ophtalmologie,* n° %1233, pp. 50-52, 2020.

[20] K. B. M. S. C. K. Klein R, «L'étude épidémiologique du Wisconsin sur la rétinopathie diabétique. XV. The long-term incidence of macular edema,» *Ophthalmology,* vol. 102, pp. 7-16, 1995.

[21] W. T. S. M. L. J. W. C. P.-C. I. H. F. F. R. Li JQ, «Prévalence, incidence et projections futures des maladies oculaires diabétiques en Europe : revue systématique et méta-analyse,» *Eur J Epidemiol,* vol. 35, n° %11, pp. 11-23, 2020.

[22] realités ophtalmologiques , «Dossier:ŒDÈME MACULAIRE DIABÉTIQUE,» realités ophtalmologiques , [En ligne]. Available: https://www.realites-ophtalmologiques.com/dossier/dossier-oedeme-maculaire-diabetique. [Accès le 16 Avril 2025].

[23] F. B.-C. Yan Guex-Crosier, «Rétinopathie diabétique : nouvelles possibilités thérapeutiques,» *Revue Medicale Suisse ,* n° %1456-457, 2015.

[24] S. B. C. C.-G. e Pascale MASSIN, «OEDEME MACULAIRE DIABETIQUE : DIAGNOSTIC ET BILAN PRE THERAPEUTIQUE,» societe francaise d'ophtalmologie, Paris, 2021.

[25] Roche, «Comment dépister et traiter un œdème maculaire diabétique ?,» Roche, [En ligne]. Available: https://www.roche.fr/articles/ophtalmologie-oedeme-maculaire-diabetique-depistage-traitement?. [Accès le 11 04 2025].

[26] A. M. T. R. S. Gabsi, «Apport de l'OCT dans les indications thérapeutiques des œdèmes maculaires diabétiques,» *Journal Français d'Ophtalmologie,* vol. 32, p. 1S153, 2009.

[27] W. T. B. H. ,. A. S.Gabsi, «Les moyens d'exploration de l'œdème maculaire diabétique : place de l'OCT,» *Journal Français d'Ophtalmologie,* vol. 31, n° %1Supplement 1, p. 117, Avril 2008.

[28] «iCare Eidon,» iCare, [En ligne]. Available: https://www.icare-world.com/product/icare-eidon/. [Accès le 17 Avril 2025].

[29] T. L. M. B. A. ,. V. C. Thibaud MATHIS MD, «Performance de la rétinophotographie en ultra-grand champ dans ledépistage de la rétinopathie diabétique,» *Journal Français d'Ophtalmologie,* vol. 42, n° %110, pp. 1-6, 2019.

[30] L. K. G. M. K. ,. K. Olivier BRUNET, «Comparaison d'un rétinophotographe ultra grand champ et d'un rétinophotographe conventionnel dans le dépistage de la rétinopathie diabétique,» SFO-ONLINE - SOCIÉTÉ FRANÇAISE D'OPHTALMOLOGIE, 2020. [En ligne]. Available: https://www.sfo-online.fr/media/comparaison-dun-retinophotographe-ultra-grand-champ-et-dun-retinophotographe-conventionnel. [Accès le 11 04 2025].

[31] V. Sarda, «Retour d'expérience :imagerie par rétinographe et tomographie à cohérence optique ultra grand champ,» *Les Cahiers d'Ophtalmologie,* n° %1265, pp. 22-27, 2023.

[32] E. Z. Aziza, «Système d'aide au diagnostic pour la détection automatique du Glaucome et de la Rétinopathie diabétique,» Tlemcen, 2023.

[33] A. A. o. Ophthalmology, « Fundamentals and Principles of Ophthalmology.Chapter 4: Evaluation of the Retina.,» 2021.

[34] F. FAJNKUCHEN, «L'interface vitréomaculaire dans l'œdème maculaire diabétique a-t-elle un rôle?,» *réalités Ophtalmologiques,* vol. 260, pp. 1-5, 2019.

[35] «Photocoagulation for diabetic macular edema. Early Treatment Diabetic Retinopathy Study report number 1. Early Treatment Diabetic Retinopathy Study research group,» *National Librery of medicine ,* vol. 103, n° %112, pp. 1796-806, 1985.

[36] D. M. B. D. M. M. D. S. B. S. P. L. F. A. G. J. S. A. C. R. J. J. H. R. G. R. J. S. E. Quan Dong Nguyen et R. a. R. R. Group, «Ranibizumab for diabetic macular edema: results from 2 phase III randomized trials: RISE and RIDE,» *PubMed,* vol. 119, n° %14, pp. 789-801, 2012.

[37] D. M. B. V. C. J.-F. K. P. K. K. Q. D. N. B. K. A. H. Y. O. G. D. Y. N. S. R. V. A. J. B. Y. S. M. A. G. G. Jeffrey S Heier 1, «Intravitreal aflibercept (VEGF trap-eye) in wet age-related macular degeneration,» *National Librery of Medicine,* vol. 119, n° %112, pp. 2537-48, 2012.

[38] A. E. 2. F. B. 3. J. J. E. B. 4. J. F. 5. E. S. 6. S. W. 7. V. G. 8. N. F. N. 9. G. L. ,. R. T. 1. R. T. 1. D. D. 1. L. W. 1. David M Brown 1, «KESTREL and KITE: 52-Week Results From Two Phase III Pivotal Trials of Brolucizumab for Diabetic Macular Edema,» *National Library of Medicine ,* pp. 157-172, 2022.

[39] Y. H. Y. 2. R. B. J. 3. F. B. 4. R. K. M. 5. A. J. A. 6. X.-Y. L. 7. H. C. 7. Y. H. 7. S. M. W. 7. David S Boyer 1 et O. M. S. Group, «Three-year, randomized, sham-controlled trial of dexamethasone intravitreal implant in patients with diabetic macular edema,» *National Library of Medicine ,* vol. 121, n° %110, pp. 1904-14, 2014.

[40] S. T. 1. E. W. 1. Barbara Sabal 1 2, «Subthreshold Micropulse Laser for Diabetic Macular Edema: A Review,» *National Library of Medicine ,* vol. 12, n° %11, p. 274, 2022.

[41] Y. G. P. 2. S. H. J. 1. S. Y. C. 1. Y.-J. R. 3. Minhee Kim 1, «The efficacy of selective retina therapy for diabetic macular edema based on pretreatment central foveal thickness,» *Nationl Library of Medicine,* vol. 35, n° %18, pp. 1781-1790, 2020.

[42] D. S. e. a. Kermany, «Identifying medical diagnoses and treatable diseases by image-based deep learning,» *Cell,* vol. 172, n° %15, pp. 1122-1131.e9, 2018.

[43] Y. H. 1. K. F. 1. T. H. 1. A. O. 1. Y. S. 1. T. M. 2. K. S. ,. T. K. ,. A. M. ,. J. K. Takumasa Tsuji 1, «Classification of optical coherence tomography images using a capsule network,» *National Library Of Medicine ,* vol. 1, n° %120, p. 1–8, 2020.

[44] R. G. V. P. K. G. K. a. K. M. Kamble, «Automated diabetic macular edema (DME) analysis using fine-tuning with Inception-ResNet-v2 on OCT images,» *2018 IEEE EMBS Conf. Biomed. Eng. Sci. (IECBES),* pp. 442-446, 2018.

[45] A. I. A. J. A. T. N. U. R. a. A. A. D. M. Muneer, «Detection of retinal diseases from OCT images using a VGG16 and transfer learning,» *Discover Applied Sciences,* vol. 5, n° %11, pp. 1-12, 2025.

[46] H. C. Z. L. X. Z. L. J. a. Z. W. F. Li, «Fully automated detection of retinal disorders by image-based deep learning,» *Graefe's Archive for Clinical and Experimental Ophthalmology,* vol. 257, n° %13, p. 495–505, 2019.

[47] H. K. a. D. K. A. Ajaz, «A review of methods for automatic detection of macular edema," Biomedical Signal Processing and Control,» *Biomedical Signal Processing and Control,* vol. 69, p. 102858, 2021.

[48] M. S. K. C. A. C. a. A. K. M. Powroznik, «Automatic Method of Macular Diseases Detection Using Deep CNN-GRU Network in OCT Images,» *Advances in Materials Science and Engineering,* vol. 2024, n° %10074, 2024.

[49] E. M. e. al, «Validation of an Automated Artificial Intelligence Algorithm for the Quantification of Major OCT Parameters in Diabetic Macular Edema,» *Journal of Clinical Medicine,* vol. 12, n° %120, p. 1–14, 2023.

[50] R. R. R. R. S. T. e. R. J. S. S. Manikandan, «Deep learning–based detection of diabetic macular edema using optical coherence tomography and fundus images: A meta-analysis,» *Indian Journal of Ophthalmology,* vol. 71, n° %15, p. 1783–1796, 2023.

[51] Y. W. T. X. L. D. L. Y. M. J. X. Z. H. J. Z. W. a. H. Z. F. Li, «A deep ensemble algorithm for detecting diabetic retinopathy and diabetic macular edema from retinal fundus photographs,» *Eye,* vol. 36, n° %17, pp. 1433-1441, 2022.

[52] V. H. G. K. P. K. G. a. M. K. R. M. Kamble, «A review of methods for automatic detection of macular edema,,» *Biomedical Signal Processing and Control,* vol. 68, p. 102726, 2021.

[53] X. L. B. Z. X. H. a. S. C. X. Guo, «Automatic Detection and Classification of Diabetic Macular Edema Based on a Deep Neural Network,» *Retina,* vol. 42, n° %16, p. 1095–1102, 2022.

[54] X. L. G. Z. Q. L. C. W. a. D. Z. Y. Fu, «Automatic grading of diabetic macular edema based on end-to-end network,» *Expert Systems with Applications,* vol. 213, p. 118835, 2023.

[55] M. S. R. A. R. e. a. D. Bhavadharini, «Automated detection of diabetic retinopathy using custom convolutional neural network,» *J. X-ray Sci. Technol,* vol. 30, n° %16, pp. 903-914, 2022.

[56] S. F. H. W. Q. H. C. L. S. H. J. S. K. L. R. T. Y. W. a. Q. Z. J. Yang, «Artificial intelligence in ophthalmopathy and ultra-wide field image: A survey,» *Expert Systems with Applications,* vol. 182, n° %1115068, 2021.

[57] P.-H. C. M. L. G. Q. a. M. E. H. D. P. Zhang, «Deep learning-based detection of referable diabetic retinopathy and macular edema using ultra-widefield fundus imaging,» *arXiv preprint arXiv:2409.12854,* p. septembre, 2024.

[58] R. A. D. M. G. Y. C. T. I. S. V. S. D. J. R. a. Z. D.-A. I. Bressler, «Autonomous Screening for Diabetic Macular Edema Using Deep Learning Processing of Retinal Images,» *Ophthalmology Science,* vol. 5, n° %14, p. 100722, 2025.

[59] P. B. P. R. P. C. S. V. A. N. J. C. K. K. G. B. M. T. S. S.-a. J. L. V. N. J. L. P. A. K. G. S. C. L. P. D. R. W. A. Varadarajan, «Predicting optical coherence tomography-derived diabetic

macular edema grades from fundus photographs using deep learning,» *Nature Communications,* vol. 11, n° %11, p. 130, 2020.

[60] K. P. Murphy, Machine Learning: A Probabilistic Perspective, Cambridge, Massachusetts, USA: MIT Press, 2012.

[61] I. E. N. &. M. J. Murphy, *What Is Machine Learning? -Machine Learning in Radiation Oncology,* Springer, 2015, pp. 3-11.

[62] Y. B. G. H. Yann LeCun, « Deep learning,» *Nature,* vol. 521, n° %17553, p. 436–444, 2015.

[63] D. Ravì, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo et G. Z. Yang, «Deep learning for health informatics,» *IEEE Journal of Biomedical and Health Informatics,* vol. 21, n° %11, pp. 4-21, 2016.

[64] I. Goodfellow, Y. Bengio et A. Courville, Deep Learning, Cambridge, MA, USA: MIT Press, 2016.

[65] J. Patterson et A. Gibson, Deep learning en action: la référence du praticien, Paris, France: O'Reilly, 2018.

[66] T. Lelong, «VGG et Transfer Learning en Deep Learning,» DataCorner, 31 aout 2020. [En ligne]. Available: https://datacorner.fr/vgg-transfer-learning/. [Accès le 29 Avril 2025].

[67] A. Zaineb, «ResNet50,» DevGenius – Medium, 06 Décembre 2022. [En ligne]. Available: https://blog.devgenius.io/resnet50-6b42934db431. [Accès le 29 Avril 2025].

[68] V. H. Nguyen et L. H. Ngo, «Automatic plant image identification of Vietnamese species using deep learning models,» *International Journal of Engineering Trends and Technology (IJETT),* vol. 68, n° %14, pp. 14-19, 2020.

[69] R. S. Jadon, «What is MobileNetV2?,» Analytics Vidhya, 11 Decembre 2023. [En ligne]. Available: https://www.analyticsvidhya.com/blog/2023/12/what-is-mobilenetv2/. [Accès le 29 Avril 2025].

[70] A. Gupta, «EfficientNetB0 architecture,» Medium – Image Processing with Python, 01 Avril 2023. [En ligne]. Available: https://medium.com/image-processing-with-python/efficientnetb0-architecture-block-1-762c5eba7e12. [Accès le 29 Avril 2025].

[71] J. Rosebrock, «What is EfficientNet?,» Roboflow Blog, 10 Octobre 2023. [En ligne]. Available: https://blog.roboflow.com/what-is-efficientnet/. [Accès le 29 Avril 2025].

[72] A. Saxena, «EfficientNet Architecture,» GeeksforGeeks, 22 decembre 2023. [En ligne]. Available: https://www.geeksforgeeks.org/efficientnet-architecture/?#efficientnetb0-architecture-overview. [Accès le 29 Avril 2025].

[73] A. H. M. Linkon, M. M. Labib, T. Hasan, M. Hossain et M.-E. Jannat, «Deep learning in prostate cancer diagnosis and Gleason grading in histopathology images: An extensive study,» *Informatics in Medicine Unlocked,* vol. 24, 2021.

[74] A. Brital, «Inception V3 CNN architecture explained,» Medium, 23 Octobre 2021. [En ligne]. Available: https://medium.com/@AnasBrital98/inception-v3-cnn-architecture-explained-691c67bbaa08. [Accès le 29 Avril 2025].

[75] S. Guindo et A. Traoré, «Deep learning pour l'analyse et la classification d'images,» Aïn Témouchent, 2024.

[76] A. Robinet, «Comprendre l'overfitting en machine learning,» 13 Fevrier 2023. [En ligne]. Available: https://www.picsellia.fr/post/comprendre-overfitting-machine-learning?. [Accès le 30 Avril 2025].

[77] IBM Editorial Team, «Underfitting-IBM – Think,» 21 fevrier 2023. [En ligne]. Available: https://www.ibm.com/think/topics/underfitting?. [Accès le 30 Avril 2025].

[78] S. Guindo et T. Alimata, «Projet de Fin d'Études:Deep learning pour l'analyse et la classification d'images,» Université Abdelhamid Ibn Badis – Aïn Témouchent, Aïn Témouchent, 2024.

[79] M. L. M. E. Amine, «Aide Au Diagnostic Pour Un Médecin Anesthésiste Réanimateur,» Université Abou Bekr Belkaid Faculté de Technologie Département Génie Électrique et Électronique, Tlemcen, 2014.

[80] P. T. Mooney, «Kermany2018,» Kaggle, 2018. [En ligne]. Available: https://www.kaggle.com/datasets/paultimothymooney/kermany2018. [Accès le 15 Mai 2025].

[81] M. Berrimi, «OCT Images (balanced version),» Kaggle, 2023. [En ligne]. Available: https://www.kaggle.com/datasets/mohamedberrimi/oct-images-balanced-version. [Accès le 15 Mai 2025].

[82] S. Kaymak et A. Serener, «Automated age-related macular degeneration and diabetic macular edema detection on OCT images using deep learning,» *2018 IEEE 14th Int. Conf. on Intelligent Computer Communication and Processing (ICCP),* pp. 265-269, 2018.