

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université de Djilali BOUNAAMA Khemis Miliana



Faculté des Sciences et de la Technologie
Département de Mathématiques et d'Informatique

Mémoire Présenté

Pour l'obtention de diplôme de

Master en Informatique

Option : Génie Logiciel et Systèmes distribués

Titre :

**Detection of Diabetic Retinopathy in fundus images using
deep learning**

Réalisé par :

- HORRI Taqyedine
- MOKRANI Mohamed

Encadré par :

- Mr. BOUKADOUM Omar

Devant les jury :

- Mme. KESSAISSIA Karima
- Mr. KHALFI Ali
- Mr. HARBOUCHE Oussama

Présidente

Examineur 1

Examineur 2

Année Universitaire : 2019/2020

Dédicaces

À mes très chers parents, source de vie, d'amour et d'affection

À mes chers frères et leurs enfants, source de joie et de bonheur

À toute ma famille, source d'espoir et de motivation

À tous mes amis

À vous cher lecteur

Taqyeddine

Dédicaces

Je dédie ce modeste travail à :

*A mes parents .Aucun hommage ne pourrait être à la
hauteur de l'amour Dont ils ne cessent de me combler. Que
dieu leur procure bonne santé et longue vie.*

A mes frères

A toute ma famille, et mes amis,

je vous dis merci.

Mohamed

Remerciements

Nous remercions Allah le tout puissant de nous AVOIR donné LA volonté et le COURAGE pour mener A bien ce TRAVAIL.

Nous tenons A remercier vivement notre promoteur Mr BOUKADOUM OMAR pour AVOIR ACCEPTÉ de diriger ce TRAVAIL et pour ses précieux conseils et ses ENCOURAGEMENTS. qui s'est toujours montré disponible tout AU long de LA REALISATION de ce mémoire, ainsi pour l'INSPIRATION, l'AIDE et le temps qu'il est bien voulu nos consacrer, et SANS qui ce mémoire N'AURAIT JAMAIS vu le jour.

Nos vifs remerciements s'ADRESSENT A tous les membres de jury qui nous ont FAIT l'honneur d'EXAMINER ce TRAVAIL.

Enfin,

A toutes les personnes qui ont contribué de près ou de loin, d'une MANIÈRE directe ou indirecte A l'ÉLABORATION de ce TRAVAIL de fin d'études.

Liste des abréviations

IA : Intelligence Artificielle

CNN : (convolutional neural networks) –réseaux des neurones convolutifs .

ML : Machine Learning

DL : Deep learning

AUC : Area Under the Curve

RAM : Random Access Memory

GPU: Graphics Processing Unit

CPU : Central Process Unit

ROC : Receiver Operating Characteristic

DR : Diabetic retinopathy

RD : Rétinopathid Diabétique

VP : Vrai Positif

VN : Vrai Négatif

FP : Faux Positif

FN : Faux Négatif

LISTE DES TABLEAUX

3.1	Illustration de l'exemple (classes positives et classes négatives).	63
4.1	Nombre d'images APTOS 2019 Blindness Detection	70
4.2	Nombre d'images dans chaque base.	71
4.3	Comparaison des résultats des deux modèles.	78

TABLE DES FIGURES

1.1	Anatomie de l'oeil.	6
1.2	Différence entre un oeil sain (gauche) et oeil avec RD (droite).	10
1.3	Fond d'oeil – cas de rétinopathie diabétique non proliférante.	11
1.4	Fond d'oeil – cas de rétinopathie diabétique proliférante.	12
1.5	La technique de l'angiographie à la fluorescence.	13
1.6	Diagnostic de la RD avec l'angiographie à la fluorescence (fond d'oeil).	14
1.7	L'appareil de rétinographie non mydriatique.	14
2.1	Image noir et blanc (gauche) et sa représentation matricielle (droite).	16
2.2	Image de type Matricielles "bitmap"	17
2.3	Image de type Matricielles "bitmap"	18
2.4	Image de type vectoriel.	18
2.5	Affichez une image bitmap (noir et blanc).	19
2.6	exemple de photo possible en 8 bpp	19
2.7	exemple de photo en Mode couleurs indexées	20
2.8	comparaison couleur indexées 8bit et RVB 24bits	21
2.9	exemple de la pixellisation	21
2.10	exemple de La résolution	22
2.11	a- Image sans bruit (gauche), b- Image avec bruit (droite).	23
2.12	pixel avec ses histogrammes.	24
2.13	Affichage par pixel de l'image en niveaux de gris.	24
2.14	Transformation d'une image en niveaux de gris en histogramme.	25
2.15	exemple avec la Taille des images	26
2.16	exemple avec la compression d'image.	27
2.17	exemple a rgb colors.	30
2.18	tableau de représentation d'Une image numérique en niveau de gris.	30
2.19	table de codage des couleurs	31
2.20	table représente Les différents systèmes de colorimétrie	31
2.21	Composition d'un système de traitement numérique.	32

2.22	Masque de filtre passe-bas.	33
2.23	Masque de filtre passe-haut.	33
2.24	Masque de filtre passe-haut.	34
2.25	Principe du filtre médiane.	35
2.26	Principe du filtre Maximum.	35
2.27	Principe du filtre Minimum.	36
3.1	Schéma de fonctionnement de Deep Learning	42
3.2	Intelligence Artificielle, Machine Learning, Deep Learning	45
3.3	le réseau de neurones	46
3.4	Les réseaux de neurones convolutifs.	46
3.5	Calcul d'une étape de la convolution.	49
3.6	Calcul de deux convolutions sur deux images, avec des kernels.	49
3.7	Pooling avec un filtre 2x2 et un pas de 2.	50
3.8	Calcul du pooling sur une image 4x4.	51
3.9	La fonctionnalité ReLU.	52
3.10	h	53
3.11	Graphe de l'absolu de la fonction tangente hyperbolique.	53
3.12	Graphe de la fonction sigmoïde	53
3.13	Représentation du traitement de tanh	53
3.14	Traitement d'une couche entièrement connectée.	54
3.15	exemple sur l'augmentation de données.	56
3.16	Image avec l'effet de rotation.	57
3.17	Image avec l'effet de translation.	57
3.18	Image avec l'effet de Retournement.	58
3.19	Image avec l'effet de mise en echelle.	59
3.20	En partant de la gauche, nous avons l'image d'origine, image avec Injection de bruit ajouté.	60
3.21	Image originale(à gauche), une section carrée recadrée en haut (au centre), puis une section carrée recadrée en bas à droite.	61
3.22	Images générées par un condition d'éclairage sur une image.	61
3.23	Résultats de matrice de confusion.	64
4.1	Architecture EfficientNet.	73
4.2	Configuration du model 1.	73
4.3	Configuration du model 2.	75
4.4	Graphes de précision (accuracy) à gauche – et de la perte (loss) à droite résultants du modèle 1.	76
4.5	Graphes de précision (accuracy) à gauche – et de la perte (loss) à droite résultants du modèle 1(après augmentation de données).	77
4.6	Graphes de précision (accuracy) à gauche – et de la perte (loss) à droite résultants du modèle 2(100 epochs).	77

TABLE DES MATIÈRES

Introduction Générale	2
1 Le Contexte Médical Rétinopathie Diabétique	4
1 Introduction	5
2 L'anatomie de l'œil :	5
2.1 La cornée :	6
2.2 Le cristallin :	7
2.3 L'iris :	7
2.4 L'humeur aqueuse :	7
2.5 La pupille :	7
2.6 La rétine :	8
2.6.1 Anatomie de la rétine :	8
2.7 Macula et fovéa :	8
2.8 Nerf optique :	9
3 Classification des stades de la RD :	9
3.1 Rétinopathie Diabétique :	9
3.1.1 Rétinopathie non proliférante :	10
3.1.2 Rétinopathie proliférante :	11
4 Techniques d'Acquisition des Images Rétiniennes :	12
4.1 L'angiographie à la fluorescence :	12
4.2 La rétinographie :	12
5 Conclusion :	13
2 Traitement d'image Médicale	15
1 Introduction	16
2 Qu'est-ce qu'une image ?	16
3 Acquisition d'une image :	16
4 Les types d'images :	17
4.1 Les images Matricielles :	17

4.2	Les images Vectorielles :	17
5	Les différents modes colorimétriques :	18
5.1	Mode bitmap (noir et blanc) :	18
5.2	Mode niveau de gris :	18
5.3	Mode couleurs indexées :	19
5.4	Mode couleur RVB :	20
6	Caractéristiques d'une image Médicale :	21
6.1	La pixellisation :	21
6.2	La résolution :	22
6.3	La dimension :	22
6.4	Le bruit :	22
6.4.1	Trois types de bruit d'image :	23
6.5	Histogramme :	23
6.6	Le contraste :	25
6.7	La luminance :	25
6.8	Taille d'une image -compression d'images :	26
6.9	Représentation d'une image :	27
6.10	Format des fichiers image :	28
6.10.1	JPEG (Joint Photographic Experts Group) :	28
6.10.2	PNG (Portable Network Graphics) :	28
6.10.3	GIF (Graphics Interchange Format) :	28
6.10.4	TIFF / TIF (Tagged Image File) :	28
6.10.5	PSD (Photoshop Document) :	28
6.10.6	PDF (Portable Document Format) :	29
6.10.7	EPS (Encapsulated Postscript) :	29
6.10.8	AI (Adobe Illustrator Document) :	29
6.11	Codage de la couleur :	29
7	Système de traitement d'image :	31
8	Techniques de traitement d'images :	32
8.1	Filtrage Numérique :	32
8.2	Filtres linéaire :	32
8.3	Filtre passe-bas (lissage) :	33
8.4	Filtre passe-haut :	33
8.5	Filtre gaussien :	34
8.6	Filtres Non linéaire :	34
8.7	Filtre médian :	34
8.8	Filtre Maximum :	35
8.9	Filtre Minimum :	36
9	Classification des images numériques :	36
9.0.1	Classification d'image numérique :	36
9.1	Les méthodes de classification supervisée :	37
9.2	Les méthodes de classification non supervisée :	37
10	Domaines d'application :	38
11	Applications du traitement d'image numérique :	38
12	Conclusion :	39

3	l'apprentissage en profondeur(deep Learning)	40
1	Introduction	41
2	Qu'est ce que le deep learning ?	41
3	Fonctionnement du deep Learning	41
4	Applications du deep Learning	42
5	Exemples d'applications en histoire de l'art	43
6	la relation entre Intelligence Artificielle, Machine Learning, Deep Learning :	44
6.1	Le Machine Learning et le Deep Learning sont de l'Intelligence Ar- tificielle	44
7	C'est quoi un réseau de neurones ?	45
8	C'est quoi un CNN	46
9	L'architecture d'un CNN	46
10	Les couche de CNN	48
10.1	Les couches convolutives (CONV)	48
10.1.1	Profondeur de la couche :	48
10.1.2	Le pas :	48
10.1.3	La marge (à 0) ou zero padding :	48
10.2	Les couches de regroupement /sous-échenillage (pooling/sub-sampeling)	50
10.3	Les couches non-linéaires :	52
10.3.1	ReLU (Rectified Linear Units) :	52
10.3.2	Les fonctions de déclenchement continues (non linéaires) :	53
10.3.3	Les couches entièrement connectées (FC) ou «dense» : . .	54
10.3.4	La Couche de perte (LOSS) :	55
11	Augmentation de données :	56
11.1	Rotation	56
11.2	Translation	56
11.3	Retournement	57
11.4	Echelle	58
11.5	Injection de bruit	59
11.6	Recadrage (cropping) :	59
11.7	Condition d'éclairage :	60
12	Métriques d'évaluation d'un RN :	61
12.1	Précision de classification (ACCURACY) :	62
12.2	Logarithmic loss : (perte logarithmique)	62
12.3	Matrice de confusion :	64
12.4	F1 score :	64
12.4.1	Précision	65
12.4.2	Rappel	65
12.5	Mean Absolute Error (MAE) :	65
12.6	Mean Absolute Error (MAE) :	66
13	Conclusion :	66
4	Implémentation ,Résultats et évaluation du travail	67
1	Introduction	68
2	Ressources matérielles et logicielles	68
2.1	Ressources logicielles	68
2.1.1	Python :	68

2.1.2	TensorFlow :	68
2.1.3	Keras :	69
2.2	Ressources matérielles :	69
2.2.1	Google Colab :	69
2.2.2	Kaggle :	69
3	Les bases d'images :	70
3.1	APTOS 2019 Blindness Detection :	70
3.2	Diabetic Retinopathy Kaggle competition dataset :	70
3.2.1	Base d'image :	71
4	Traitements préliminaires :	71
4.1	Filtrage :	71
4.2	Redimensionnement :	71
5	L'apprentissage par transfert(Transfer Learning) :	71
5.1	modèle pré-entraîné(pre-trained models) :	72
6	Architecture de notre réseau :	72
6.1	EfficientNet :	72
6.2	Architecture EfficientNet :	72
6.3	Architecture du modèle 1 :	72
6.4	Architecture du modèle 2 :	74
7	Résultats obtenus et discussion :	75
7.1	L'évaluation de notre réseau :	75
7.2	Résultats obtenu pour modèle 1 :	76
7.3	Résultats obtenu pour modèle 2 :	77
7.4	Comparaison des deux modèles :	78
7.5	Résultat final :	78
7.6	L'explication de Code source :	78
8	Conclusion :	79

Conclusion Générale **81**

Bibliographie **84**

INTRODUCTION GÉNÉRALE

Introduction

Imaginez être capable de détecter la cécité avant qu'elle ne se produise.

Des millions de personnes souffrent de rétinopathie diabétique, principale cause de cécité chez les adultes en âge de travailler. Environ quatre cent vingt millions de personnes dans le monde ont reçu un diagnostic de diabète sucré. La rétinopathie diabétique est une cause majeure de malvoyance et de cécité en France. C'est la première cause de cécité avant l'âge de 50 ans, et la troisième après 50 ans. La prévalence de la rétinopathie diabétique augmente avec la durée du diabète et l'importance de l'hyperglycémie chronique.

Il importe donc de chercher à dépister précocément les lésions de rétinopathie diabétique, car la cécité peut d'autant mieux être prévenue que le traitement est réalisé à temps. Pour cela, une surveillance ophtalmologique régulière doit être poursuivie tout au long de la vie du diabétique et le traitement par laser doit être réalisé avant que la rétinopathie diabétique n'ait causé une perte irréversible de la vision .

Une coopération entre l'ophtalmologiste, le diabétologue et le médecin généraliste est indispensable pour motiver le patient à accepter une surveillance ophtalmologique régulière. Les principes de cette surveillance et les modalités du traitement sont bien connus et largement acceptés par la communauté médicale, mais malheureusement de très nombreux patients ne bénéficient pas encore du suivi ophtalmologique nécessaire. Le but de ces recommandations est de contribuer à rationaliser et à généraliser la surveillance ophtalmologique des diabétiques.

Enfin, compte tenu de l'augmentation de la prévalence du diabète et des complications rétiniennes associées dans le monde, les méthodes de diagnostic manuelles risquent de ne pas répondre à la demande de services de dépistage. Les techniques automatisées de diagnostic de la rétinopathie diabétique sont essentielles pour résoudre ces problèmes.

Objective :

Dans ce travail, nous présenterons un classificateur automatisé de rétinopathie diabétique capable de classer les images rétiniennes en fonction des conditions pathologiques et également de détecter le stade de développement de cette maladie, dont le but est de réduire la valeur de la perte et d'augmenter la précision au maximum. par utilisation l'apprentissage en profondeur (deep Learning).

Organisation du travail :

Ce travail est organisé en trois chapitres suivis d'une conclusion générale, chacun représente une partie du travail de ce projet qu'elle soit théorique ou pratique :

Chapitre I : Rétinopathie diabétique

Dans ce chapitre on va parler de l'anatomie de l'œil humain, les maladies des yeux et de la rétinopathie diabétique.

Chapitre II : Traitement d'image Médicale

Ce chapitre va couvrir le domaine de traitement d'images, des définitions de base, les techniques de filtrage et la classification des images numériques

Chapitre III : l'apprentissage en profondeur(deep Learning)

Ce chapitre est consacré à la description l'apprentissage en profondeur et des réseaux de neurones convolutionnels ainsi que leurs l'intérêt dans le domaine de la classification des images.

Chapitre IV : Implémentation : On va montrer la partie expérimentale de notre travail et on discute les différents résultats obtenus et à la fin on termine par une conclusion générale.

CHAPITRE

1

LE CONTEXTE MÉDICAL
RÉTINOPATHIE DIABÉTIQUE

1 Introduction

Le diabète représente l'une des principales causes de cécité (perte totale de la vue) en Europe. La rétinopathie diabétique concernerait, en effet, au moins 6 diabétiques de type 2 sur 10 après une dizaine d'années de diabète. C'est l'une des complications majeures de cette maladie.

La rétinopathie diabétique est une atteinte des yeux du diabétique, plus particulièrement de la rétine. C'est une grave complication du diabète qui concerne une grande majorité des patients atteints de diabète de type 2. En France, cette pathologie représente même la première cause de cécité, c'est-à-dire de perte totale de la vue, avant 65 ans. Cette maladie touche 2 à 2,5% de la population française, dont près de 90% des diabétiques[1].

Donc Dans ce chapitre nous devons expliquer La rétinopathie diabétique et l'anatomie de l'oeil et la relation entre les Deux ?

2 L'anatomie de l'œil :

L'œil humain correspond grossièrement à une petite boule de 2,5 cm de diamètre, pesant sept à huit grammes. le globe oculaire se compose de différentes couches. La plus externe correspond à la sclère, membrane résistante qui donne sa couleur blanche à l'œil en le recouvrant à 80 %, sur laquelle s'insèrent les muscles oculomoteurs. Au centre se trouve la cornée, qui permet de capter les premiers rayons de lumière et de commencer la réfraction.

Sous la sclère, on trouve la choroïde, couche fine irriguée de vaisseaux sanguins et assurant la nutrition de la rétine et de l'iris.

La rétine, située dans le fond du globe oculaire, au contact de la choroïde, correspond à la région directement sensible à la lumière, polarisée principalement par la cornée et le cristallin.

Elle est composée de cellules en cônes et en bâtonnets, sensibles aux détails de l'environnement et à la luminosité.

Au centre de l'œil, derrière la cornée, on trouve l'iris, région qui confère sa couleur particulière au globe oculaire. En son centre, on trouve la pupille, région sombre qui, sous l'action de muscles, se contracte ou se dilate pour laisser passer plus ou moins de lumière, à l'instar d'un diaphragme sur un appareil photo.

L'essentiel du volume de l'œil (90 %) est dû à l'humeur vitrée, liquide gélatineux situé en arrière du cristallin, qui donne au globe oculaire sa consistance. Il contribue également à la réfraction de la lumière[2].

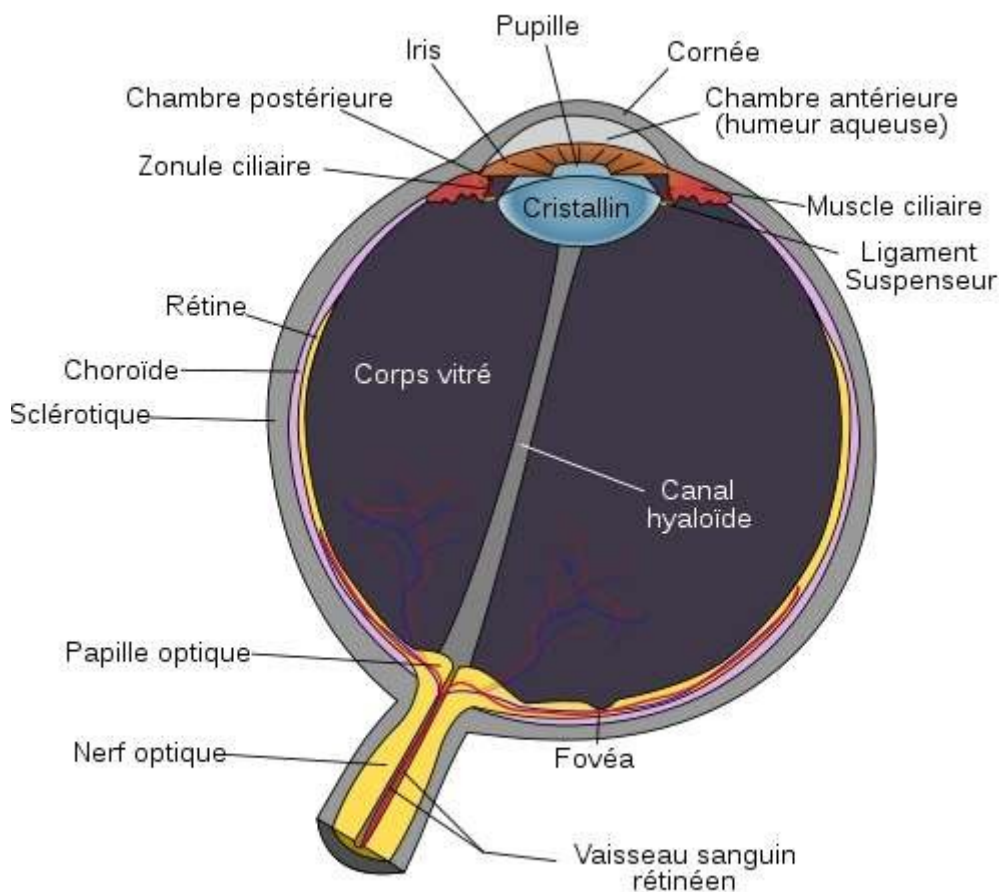


FIGURE 1.1 – Anatomie de l'œil.

2.1 La cornée :

La cornée est un composant oculaire essentiel au fonctionnement de la vision : elle est la première structure que rencontre la lumière qui pénètre l'œil. Son rôle principal est de faire converger les rayons lumineux incidents qui se dirigent ensuite au travers de la chambre antérieure de l'œil, vers le cristallin, avant de rencontrer la rétine et initier la cascade visuelle.

La cornée est un tissu non vascularisé et transparent, de géométrie courbe. Elle est semblable à une coupole hémisphérique.

La cornée est recouverte du film lacrymal et est majoritairement constituée de fibres collagènes entrelacées, qui constituent les 90% de l'épaisseur cornéenne.

Il est recouvert d'un feuillet appelé épithélium, et tapissé à sa face interne d'une monocouche de cellules appelées cellules endothéliales[3].

2.2 Le cristallin :

Le cristallin est une lentille optique naturelle de l'oeil qui participe à la focalisation des rayons lumineux sur la rétine et à l'accommodation (mise au point en vision de près).

Le cristallin jeune est souple, transparent, et possède une forme de pourtour elliptique en coupe. IL est situé dans la chambre postérieure de l'œil, en avant de la cavité vitréenne, et en arrière de l'iris.

Le cristallin est suspendu à l'anneau formé par le corps ciliaire par l'intermédiaire de la zonule cristallinienne, ligament constitué de fibres qui s'insèrent au niveau de l'équateur du cristallin[4].

2.3 L'iris :

L'iris correspond à la zone colorée de l'œil qui est percé en son centre par la pupille, dans laquelle la lumière pénètre. Ces deux éléments jouent un rôle primordial dans la vision. Grâce à l'iris, la pupille peut se dilater ou rétrécir en fonction de la luminosité, ce qui permet de garder une vision nette à tout moment.

L'iris peut prendre un large éventail de couleurs différentes : vert, bleu, marron et même gris ou violet dans certains cas. Sa couleur est due à deux pigments : la mélanine et la lipofuscine (dans le cas de certains yeux verts). [5]

2.4 L'humeur aqueuse :

L'humeur est un liquide transparent et incolore qui transporte les aliments destinés au cristallin et, pour partie, à la cornée. L'humeur aqueuse joue un rôle dans la régulation de la pression intraoculaire.

Elle est sécrétée par le corps ciliaire dans la chambre postérieure et se dirige vers la chambre antérieure par la pupille pour s'évacuer par le trabéculum, sorte de filtre. Ses dysfonctionnements circulatoires sont responsables d'une hausse de la pression intraoculaire et de certains glaucomes[6].

2.5 La pupille :

La pupille est formée de l'iris, la cornée et l'humeur aqueuse qui jouent un rôle complémentaire dans la dilatation ou la rétractation de la pupille.

Si l'on compare les yeux à un appareil photographique, la pupille est un diaphragme par lequel la lumière pénètre. La pupille grandit ou rétrécit en fonction de la luminosité. Avec une lumière intense la pupille ne mesure que 1 à 2 millimètres de diamètre, alors que dans l'obscurité elle peut atteindre 8 mm de diamètre. Les jeux de la pupille permettent de garder une vision nette quelque soit la lumière ambiante.

La lumière n'est pas seule à agir sur la pupille. Une vive émotion (surprise, bonheur, peur ...) peuvent provoquer une dilatation de la pupille. [7]

2.6 La rétine :

La rétine est une fine membrane tapissant le globe oculaire. Par analogie avec un appareil photo, elle est en quelque sorte la « pellicule » de l'œil, chargée de capter les rayons lumineux pour les transmettre au système nerveux central. D'une épaisseur comprise entre 0,1 et 0,4 mm, elle est constituée de 10 couches : quatre de cellules photoréceptrices à l'extérieur et six de cellules nerveuses à l'intérieur.

2.6.1 Anatomie de la rétine :

La rétine est divisée en deux grandes zones :

- la rétine centrale, située au centre, comprend la macula et la fovéa. De 5 à 6 millimètres de diamètre, elle est constituée de cellules photoréceptrices appelées cônes, spécialisés dans l'interprétation des formes et des couleurs en vision diurne. On en dénombre environ 7 millions.
- la rétine périphérique, qui s'étend sur environ 17 millimètres, est plus particulièrement constituée de bâtonnets, les cellules chargées d'analyser la luminosité. Au nombre de 110 à 130 millions, ils ne distinguent pas les couleurs mais permettent la vision de nuit[8].

2.7 Macula et fovéa :

La macula est le centre de la rétine située au fond de l'œil. Exempte de capillaires sanguins, cette zone est aussi appelée "tache jaune" de par sa couleur liée à un pigment jaune nommé xanthophylle, identique au feuillage végétal. La macula est uniquement composée de photorécepteurs visuels appelés « cônes » qui directement exposés à la lumière permettent la précision de vision, la vision des couleurs, la vision de jour, et contribuent à la qualité de vision au centre du champ visuel.

Au centre de la macula se trouve une petite zone de 1,5 mm de diamètre, la fovéa, de forme conique et creuse. Au centre de la fovéa se trouve la fovéola : toute petite zone de la rétine où l'acuité visuelle est maximum. Par rapport à la papille, la macula est située côté temporal. Sa forme est une ellipse d'environ 3 X 2 mm². En son centre, la fovéa est de forme creuse telle un entonnoir. Le point le plus profond est la fovéola[9].

2.8 Nerf optique :

Le nerf optique est responsable de la connexion de l'œil au cerveau. Le nerf optique transporte les impulsions formées par la rétine, la couche nerveuse qui tapisse l'arrière de l'œil et détecte la lumière et crée des impulsions. Ces impulsions sont envoyées par le nerf optique au cerveau, qui les interprète comme des images. À l'aide d'un ophtalmoscope, la tête du nerf optique peut être facilement vue. Il peut être considéré comme la seule partie visible du cerveau[10].

3 Classification des stades de la RD :

La rétinopathie diabétique débute par un stade de rétinopathie diabétique non proliférante minime puis modérée, et évolue vers la rétinopathie diabétique non proliférante sévère (ou préproliférante) caractérisée par une ischémie rétinienne étendue, puis vers le stade de rétinopathie diabétique proliférante caractérisée par la prolifération de néovaisseaux à la surface de la rétine et/ou sur la papille. L'œdème maculaire est associé aux rétinopathies diabétiques non proliférantes ou proliférantes.

les stades de la rétinopathie diabétique :

- pas de rétinopathie diabétique.
- rétinopathie diabétique non proliférante (RDNP).
- rétinopathie diabétique proliférante (RDP).[11]

3.1 Rétinopathie Diabétique :

La rétinopathie diabétique, caractérisée par des lésions de la rétine de l'œil, est une grave complication du diabète qui touche 50 % des patients diabétiques de type 2 (à différents stades), selon la Fédération française des diabétiques.

Le diabète est ainsi responsable de 12 % de l'ensemble des cas de cécité dans les pays occidentaux ! Plusieurs facteurs favorisent la survenue d'une rétinopathie diabétique et accélèrent sa progression : l'ancienneté du diabète, le niveau de glycémie, un diabète

instable, l'hypertension artérielle, le tabagisme, etc.

Cette pathologie peut accélérer la survenue d'autres maladies des yeux comme les glaucomes ou la cataracte, et même conduire à la cécité en l'absence d'un traitement adapté.[12]

RÉTINOPATHIE DIABÉTIQUE

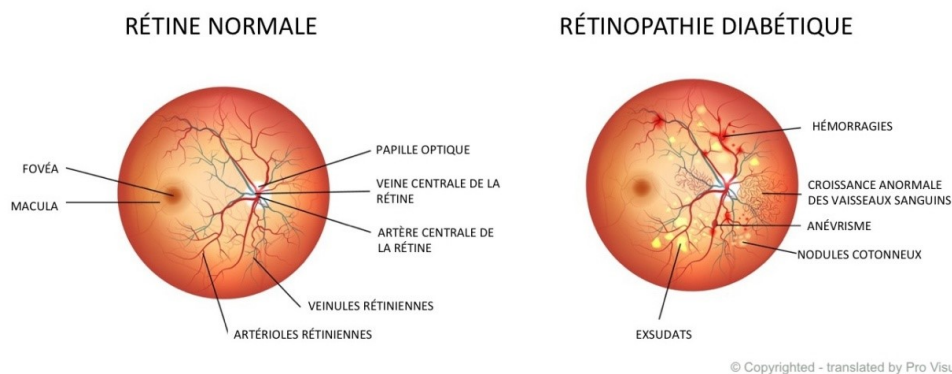


FIGURE 1.2 – Différence entre un oeil sain (gauche) et oeil avec RD (droite).

3.1.1 Rétinopathie non proliférante :

Après plusieurs couches de convolution et de max-pooling, le raisonnement de haut niveau dans le réseau neuronal se fait via des couches entièrement connectées. Des symptômes visuels sont provoqués par l'œdème maculaire ou l'ischémie maculaire. Cependant, les patients peuvent ne pas subir de perte de vision, même en cas de rétinopathie avancée.

Les premiers signes de la rétinopathie non proliférante sont les suivants :

- Des microanévrismes.
- Des microanévrismes
- Des hémorragies rétinienne ponctuelles et en taches,
- Des exsudats durs
- Des nodules cotonneux (exsudats mous)

Les exsudats durs sont des particules jaunes discrètes situées dans la rétine. Lorsqu'ils sont présents, ils évoquent un œdème chronique. Les nodules cotonneux sont des zones de micro-infarctus de la couche des fibres nerveuses de la rétine, qui entraînent l'opacification de la rétine ; ils sont blancs, à bords flous, et masquent des vaisseaux sous-jacents.

Les signes aux stades plus tardifs sont les suivants :

- Un œdème maculaire (observé par biomicroscopie à la lampe à fente comme une élévation et un aspect flou des couches de la rétine)
- Une dilatation veineuse et des anomalies microvasculaires intrarétiniennes

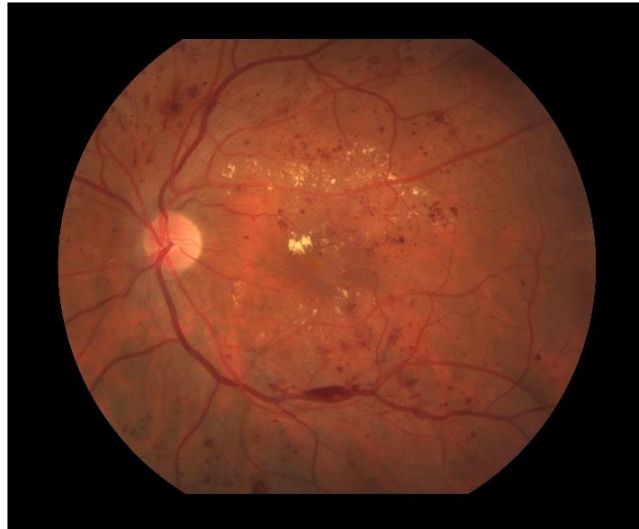


FIGURE 1.3 – Fond d’oeil – cas de rétinopathie diabétique non proliférante.

3.1.2 Rétinopathie proliférante :

Les symptômes peuvent comprendre une vision trouble avec des corps flottants (taches noires) ou des points lumineux dans le champ visuel, et parfois une baisse d’acuité visuelle sévère, brutale et indolore. Ces symptômes sont généralement provoqués par une hémorragie intravitréenne ou un décollement de la rétine par traction.

Contrairement à la rétinopathie non proliférante, la rétinopathie proliférante entraîne la formation d’une néovascularisation des vaisseaux visible sur le nerf optique ou à la surface de la rétine. L’œdème maculaire ou l’hémorragie intravitréenne sont visibles à l’examen du fond d’œil[13].



FIGURE 1.4 – Fond d’oeil – cas de rétinopathie diabétique proliférante.

4 Techniques d’Acquisition des Images Rétiniennes :

4.1 L’angiographie à la fluorescence :

Cet examen étudie l’état des minuscules vaisseaux sanguins qui irriguent la rétine. Pour les visualiser on utilise différents types de colorants fluorescents qui permettront de mieux les étudier. Le colorant est injecté dans une veine. Il est suivi et éclairé avec une lumière pour le rendre fluorescent.

Cet examen est indispensable au diagnostic précoce de nombreuses maladies vasculaires de l’œil.

Les trois principales sont :

- La rétinopathie diabétique.
- La dégénérescence maculaire liée à l’âge.
- L’occlusion veineuse rétinienne.

Cet examen, permettra de dépister les complications de ces maladies avant qu’elle ne soit irréversible. La mise en place de traitement préventif ou curatif pourra alors être décidée[14].

4.2 La rétinographie :

La rétinographie est un examen de diagnostic qui, sous la forme de photographies, permet d’obtenir une image du fond de l’œil ou de la rétine. Cette procédure permet de détecter certaines maladies qui affectent la rétine, comme la dégénérescence maculaire liée

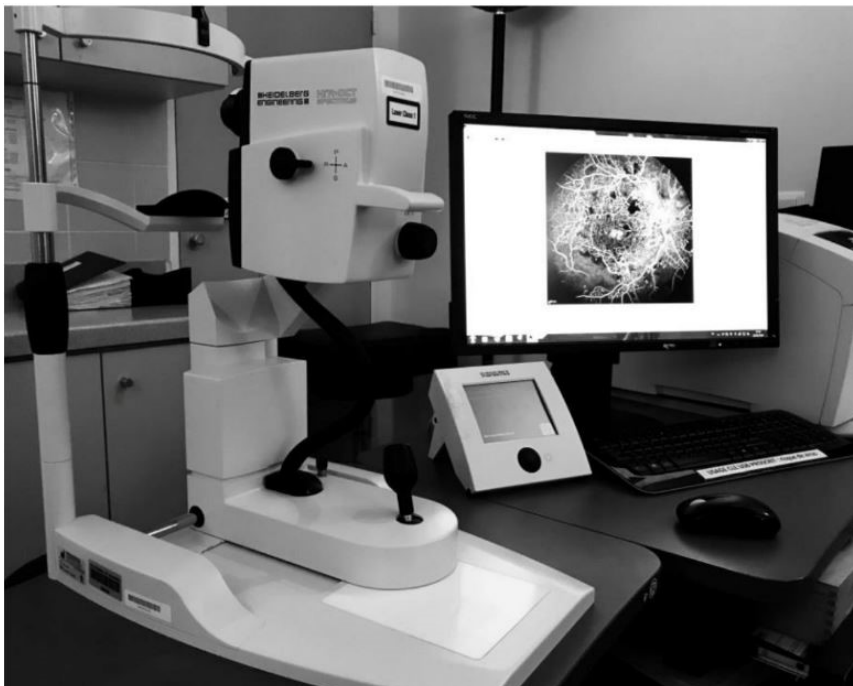


FIGURE 1.5 – La technique de l'angiographie à la fluorescence.

à l'âge, la rétinopathie diabétique ou la rétinite pigmentaire.

La rétinographie est un examen simple, indolore, utile et sûr. De plus, il requiert uniquement l'application d'un collyre afin de dilater la pupille[15].

5 Conclusion :

Dans ce premier chapitre, nous avons décrit l'anatomie de l'oeil, précisément la rétine et ses éléments principaux, puis nous avons présenté la Rétinopathie Diabétique et les autres pathologies rétiniennes, ensuite nous avons présenté les techniques qui permettent l'acquisition des images angiographique et les images couleur.

Dans le chapitre suivant nous allons parler des images numériques, des différentes techniques de traitement d'images numériques, les systèmes de traitements et la classification des images.



FIGURE 1.6 – Diagnostique de la RD avec l'angiographie à la fluorescence (fond d'oeil).



FIGURE 1.7 – L'appareil de rétinographie non mydriatique.

CHAPITRE

2

TRAITEMENT D'IMAGE MÉDICALE

1 Introduction

traitement d'image (image processing) consiste à appliquer des transformations mathématiques sur des images dans le but d'améliorer leur qualité ou d'en extraire une information. On peut citer par exemple : la création d'effets spéciaux au cinéma, l'amélioration de la qualité des photos en corrigeant les conditions de prise de vue, la détection de lettres ou de visages, l'identification de zones cancéreuses en imagerie médicale, la compression JPEG, etc.[16]

2 Qu'est-ce qu'une image ?

Une image est une représentation visuelle que l'on se fait grâce à la vue. Elle répond à plusieurs critères : le relief (la profondeur, les plans), la lumière, ainsi que l'angle de la prise de vue. Tout cela influe sur la perception d'une image.

- Une image numérique est une image acquise, traitée et stockée en bits.
- Une image numérique est un tableau de pixel : chaque pixel est codé par un nombre binaire pour un niveau de gris, ou par trois nombres binaires qui correspondent à une nuance de rouge, de vert et de bleu (codage RVB). [17]

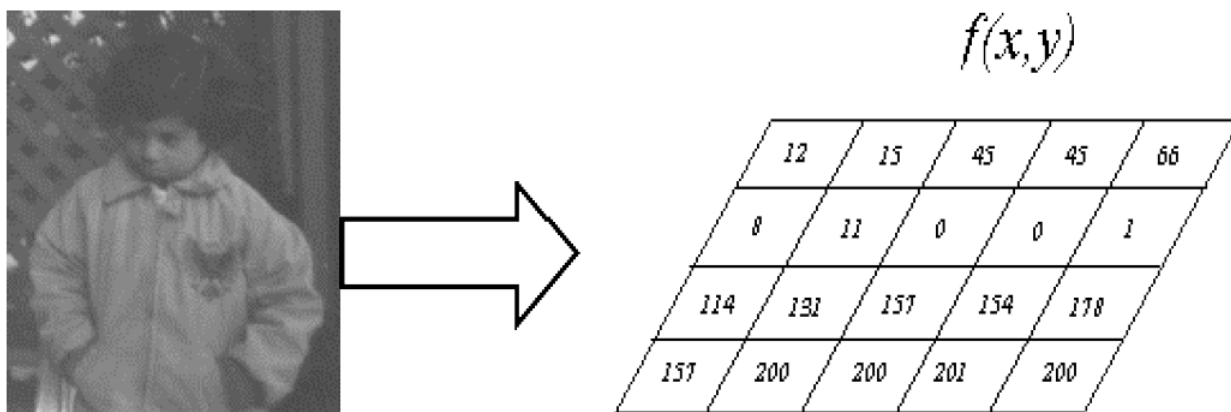


FIGURE 2.1 – Image noir et blanc (gauche) et sa représentation matricielle (droite).

3 Acquisition d'une image :

L'acquisition d'image dans le traitement d'image peut être définie au sens large comme l'action de récupérer une image à partir d'une source, généralement une source matérielle, afin qu'elle puisse être transmise à tous les processus nécessaires par la suite. L'acquisition d'images dans le traitement d'image est toujours la première étape de la

séquence de flux de travail car, sans image, aucun traitement n'est possible.

Selon le domaine de travail, un facteur majeur impliqué dans l'acquisition d'images dans le traitement d'image est parfois la configuration initiale et la maintenance à long terme du matériel utilisé pour capturer les images. Le périphérique matériel réel peut être n'importe quoi, d'un appareil photo numérique, d'un scanner de bureau à un télescope optique massif. Si le matériel n'est pas correctement configuré et aligné, des artefacts visuels peuvent être produits qui peuvent compliquer le traitement de l'image. Un matériel mal configuré peut également fournir des images d'une qualité si médiocre qu'elles ne peuvent pas être récupérées même avec un traitement intensif. Tous ces éléments sont vitaux pour certains domaines, comme le traitement comparatif des images, qui recherche des différences spécifiques entre les ensembles d'images. [18]

4 Les types d'images :

Les images produites et traitées par les ordinateurs sont de deux types : les images Matricielles "bitmap" et les images "vectorielles".



FIGURE 2.2 – Image de type Matricielles "bitmap" .

4.1 Les images Matricielles :

Les images matricielles (ou image en mode point, en anglais « bitmap » ou « raster ») sont celles que nous utilisons généralement pour restituer des photos numériques. Elles reposent sur une grille de plusieurs pixels formant une image avec une définition bien précise. Lorsqu'on les agrandi trop, on perd de la qualité (« pixelisation »).

4.2 Les images Vectorielles :

Ce sont des images dont la particularité est que chaque forme qui la compose est décrite mathématiquement à partir de points et de tangentes. Elle ne peuvent pas décrire



FIGURE 2.3 – Image de type Matricielles "bitmap" .

une image trop complexe comme une photographie, mais sont tout à fait adaptées au rendu typographiques, aux logos et autres formes composées de tracés simples.



FIGURE 2.4 – Image de type vectoriel.

5 Les différents modes colorimétriques :

5.1 Mode bitmap (noir et blanc) :

Avec ce mode, il est possible d'afficher uniquement des images en deux couleurs : noir et blanc. Il utilise une seule couche.

- Codage en 1 bit par pixel (bpp) : $2^1 = 2$ possibilités : [0,1]
- Chaque pixel peut donc avoir 2 couleurs possibles : soit noir ou soit blanc.

5.2 Mode niveau de gris :

il permet d'obtenir différentes valeurs de gris, afin d'afficher des images nuancées. Il utilise qu'une seule couche.

- Codage en 8 bits par pixel (bpp) :

1	1	1	1	1	1	1	1	1	1
1	1	1	0	0	0	0	1	1	1
1	1	0	1	1	1	1	0	1	1
1	0	1	1	1	1	1	1	0	1
1	0	1	0	1	1	0	1	0	1
1	0	1	1	1	1	1	1	0	1
1	0	1	0	1	1	0	1	0	1
1	0	1	1	0	0	1	1	0	1
1	1	0	1	1	1	1	0	1	1
1	1	1	0	0	0	0	1	1	1
1	1	1	1	1	1	1	1	1	1

FIGURE 2.5 – Affichez une image bitmap (noir et blanc).

$2^8 = 256$ possibilités Chaque pixel peut avoir 256 nuances de gris possibles.

— **Codage en 16 bits par pixel (bpp) :**

$2^{16} = 65536$ possibilités Chaque pixel peut avoir 65536 nuances de gris possibles.

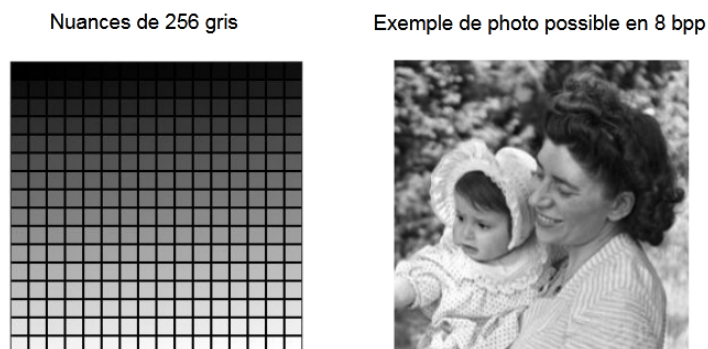


FIGURE 2.6 – exemple de photo possible en 8 bpp .

5.3 Mode couleurs indexées :

permet d'obtenir jusque 256 couleurs fixes, définies à l'avance dans une palette. Il utilise qu'une seule couche.

— **Codage en 8 bits par pixel (bpp) :**

$2^8 = 256$ possibilités Chaque pixel peut avoir jusque 256 couleurs fixes possibles.

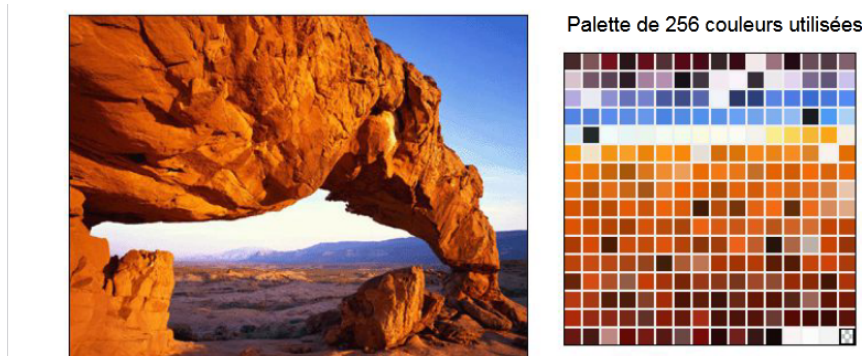


FIGURE 2.7 – exemple de photo en Mode couleurs indexées .

5.4 Mode couleur RVB :

grâce au mélange des 3 couches de couleur, il est possible de reproduire un plus grand nombre de nuances qu'avec une palette en mode couleurs indexées.

*** Avec un codage en RVB 8 bits PAR COUCHE :**

Chaque couche utilise 8bit (1 octet), soit 256 nuances possibles : 8Bits pour le Rouge, 8bit pour le Vert et 8bits pour le Bleu. Donc utilisation de $3 \times 8\text{bits} = 24$ bits utilisées au total.

— $256 \times 256 \times 256 = 2^{24} = 16,7$ millions Chaque pixel peut prendre 16,7 Millions de couleurs possibles.

*** Avec un codage en RVB 16 bits PAR COUCHE :**

Chaque couche utilise le double, soit 16bits ! (65535 nuances). $3 \times 16 = 48$ bits utilisées au total.

— $65535 \times 65535 \times 65535 = 2^{48} = 4,4$ milliards de nuances de couleurs sont possibles.
[19]



FIGURE 2.8 – comparaison couleur indexées 8bit et RVB 24bits .

6 Caractéristiques d'une image Médicale :

6.1 La pixellisation :

Une image numérique est constituée d'un ensemble de points appelés pixels (abréviation de PICture Element) pour former une image. Le pixel représente ainsi le plus petit élément constitutif d'une image numérique. L'ensemble de ces pixels est contenu dans un tableau à deux dimensions constituant l'image .

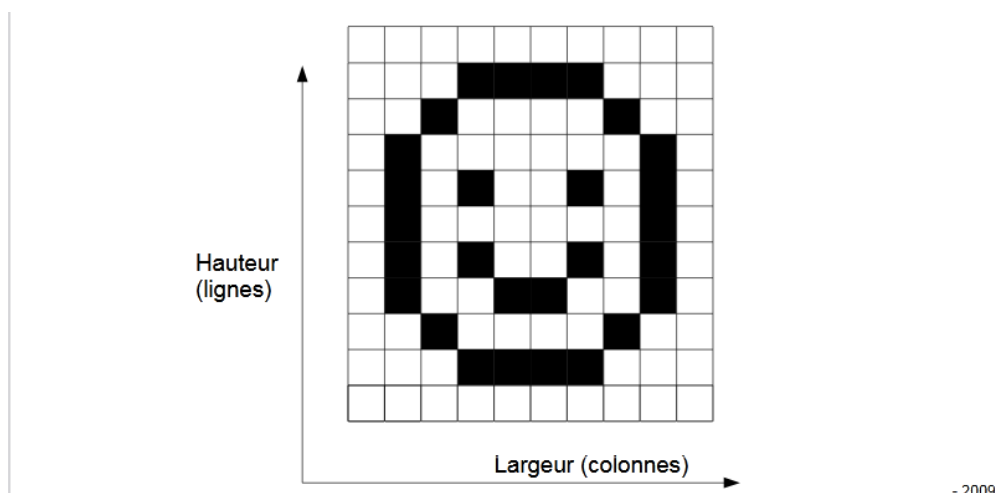


FIGURE 2.9 – exemple de la pixellisation .

6.2 La résolution :

C'est le nombre de points contenu dans une longueur donnée(en pouce). Elle est exprimée en points par pouce (PPP, en anglais : DPI pour Dots Per Inch). Un pouce mesure 2.54 cm, c'est une unité de mesure britannique.- La résolution permet ainsi d'établir le rapport entre la définition en pixels d'une image et la dimension réelle de sa représentation sur un support physique (affichage écran, impression papier...).[3]

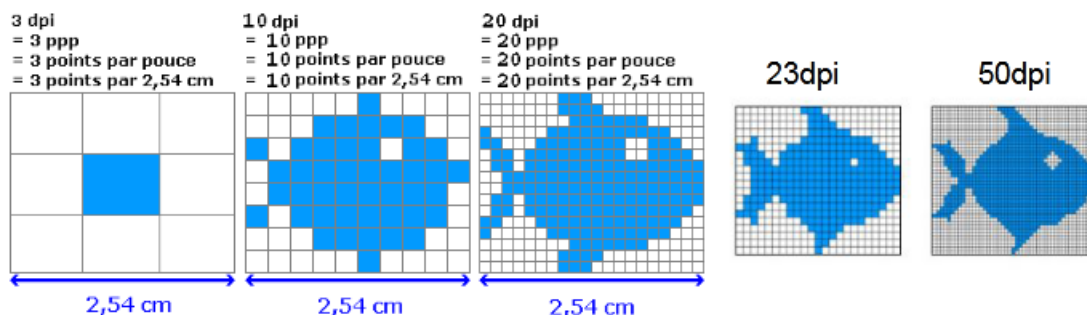


FIGURE 2.10 – exemple de La résolution .

6.3 La dimension :

Les dimensions de l'image sont la longueur et la largeur d'une image numérique. Il est généralement mesuré en pixels, mais certains programmes graphiques vous permettent de visualiser et de travailler avec votre image en pouces ou en centimètres équivalents. En fonction de ce que vous prévoyez d'utiliser votre image, vous souhaitez peut-être modifier la taille de l'image. Par exemple, si vous utilisez une photographie numérique haute résolution, vous souhaitez peut-être réduire les dimensions de l'image pour la publication sur une page Web. Lorsque vous utilisez un programme d'édition graphique ou d'image, vous aurez généralement deux options pour modifier les dimensions de l'image : redimensionner ou rééchantillonner. [20]

6.4 Le bruit :

En photographie numérique, le bruit de l'image peut être comparé au grain du film pour les appareils photo analogiques. Le bruit de l'image se manifeste généralement par des taches aléatoires sur une surface lisse et peut sérieusement affecter la qualité de l'image.

Parfois, cependant, il peut être utile d'augmenter la netteté apparente d'une image numérique.

Le niveau de bruit augmente généralement en fonction de la durée d'exposition, de la température physique et du réglage de sensibilité de l'appareil photo. La quantité de certains types de bruit d'image présente à un réglage donné varie selon les modèles de caméra et est liée à la technologie du capteur.



FIGURE 2.11 – a- Image sans bruit (gauche), b- Image avec bruit (droite).

6.4.1 Trois types de bruit d'image :

- Les principaux types de bruit d'image sont le bruit aléatoire, le bruit à motif fixe et le bruit de bande.
- Le bruit aléatoire est indiqué par la fluctuation des couleurs au-dessus de l'intensité réelle de l'image.
- Un bruit de motif fixe apparaît après de longues expositions et des températures élevées.
- Enfin, le bruit de bande est introduit lorsque la caméra lit les données du capteur et il est directement lié aux facteurs technologiques de la caméra.[21]

6.5 Histogramme :

Un histogramme est une représentation graphique d'un ensemble de données (continu) qui vous permet de découvrir et d'afficher la distribution de fréquence sous-jacente (forme) des données.

Dans le traitement d'image, nous utilisons des histogrammes d'intensité pour décrire la distribution tonale dans une image numérique. En regardant l'histogramme pour une image spécifique, un spectateur pourra juger la distribution tonale entière en un coup d'œil. L'axe horizontal du graphique représente les variations tonales, tandis que l'axe vertical représente le nombre de pixels dans ce ton particulier.

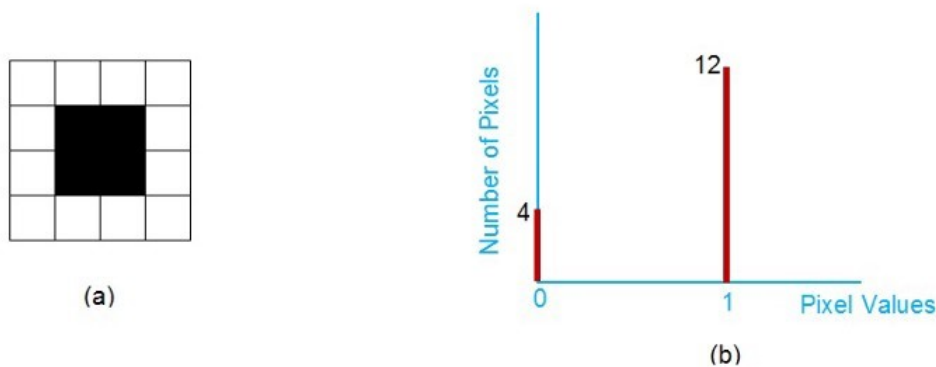


FIGURE 2.12 – pixel avec ses histogrammes.

Dans la représentation 8 bits, il y a 2^8 niveaux de gris, soit 256 niveaux de gris (0,1,..., 255). Nous utilisons le niveau de gris 0 pour représenter le noir et le niveau de gris 255 pour représenter le blanc.

Nous passons à chacun des pixels d'une image numérique pour trouver les niveaux de gris, puis nous incrémentons la fréquence appropriée pour créer l'histogramme. [22]

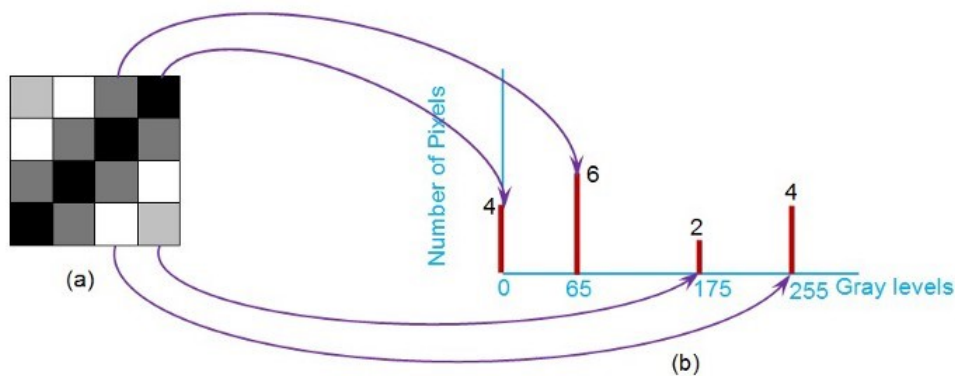


FIGURE 2.13 – Affichage par pixel de l'image en niveaux de gris.

L'image suivante décrit la transformation d'une image en niveaux de gris en histogramme.

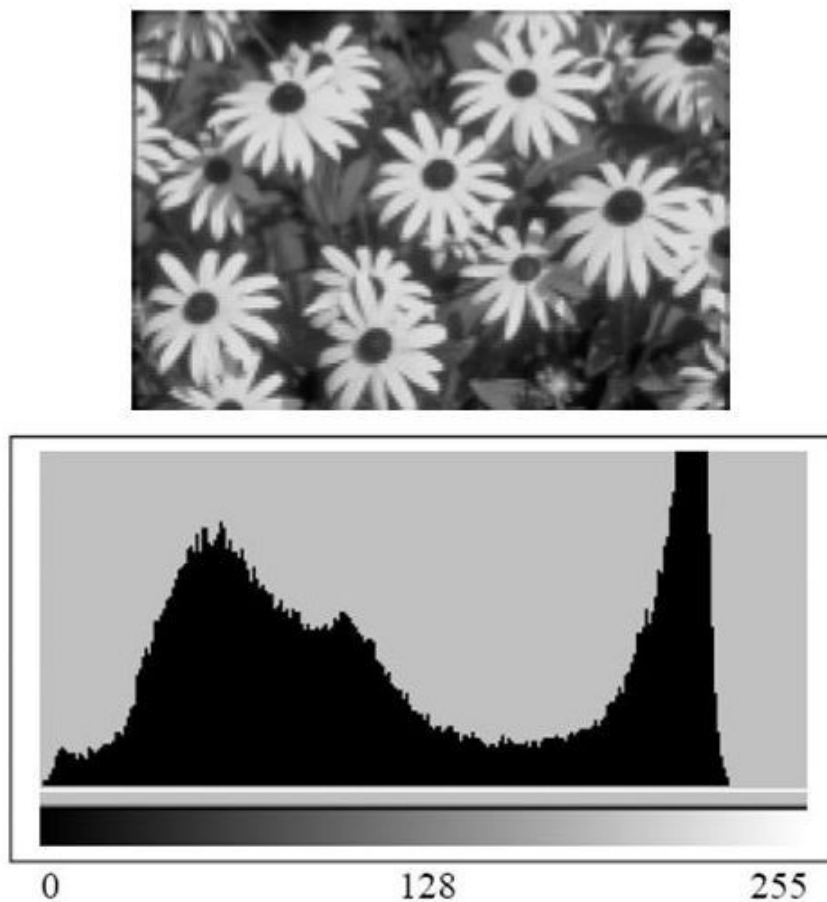


FIGURE 2.14 – Transformation d'une image en niveaux de gris en histogramme.

6.6 Le contraste :

Le contraste est l'une des notions les plus importantes en matière de photographie. Elle indique la différence de densité entre les tons les plus clairs et ceux les plus foncés d'une image. Pour simplifier, une image très contrastée comportera presque que du blanc et du noir, alors qu'une image peu contrastée contiendra énormément de valeurs intermédiaires mais pas nécessairement de véritables noirs et blancs. [23]

6.7 La luminance :

La luminance désigne le signal qui détermine les valeurs de contraste d'une image, du noir le plus profond jusqu'au blanc le plus pur. La couleur correspond à l'autre partie du signal, appelée chrominance. Cette grandeur photométrique fondamentale caractérise

l'émission intrinsèque d'une source, indépendamment de sa géométrie. La luminance est le flux émis par unité de surface apparente et par unité d'angle solide. [24]

6.8 Taille d'une image -compression d'images :

La résolution d'une image est la mesure de base de la taille de l'image et ses dimensions dans un navigateur Web. La résolution de l'image est déterminée par le nombre total de pixels situés dans l'image noté comme pixels de large x pixels de haut. Par exemple, 1024×760 . Ainsi, si la taille de l'écran de votre ordinateur est de 1024×760 , une image de cette taille couvrirait toute la zone de votre écran.

La taille du fichier d'une image est la mesure de base de son poids, qui affecte le temps de téléchargement. Les images plus grandes ou les images de résolution plus élevée ont généralement des fichiers de plus grande taille. Cependant, deux images de même taille ou résolution n'ont pas toujours la même taille de fichier. Le nombre de couleurs et la variation de ces couleurs ont un impact sur la taille du fichier.

Par exemple, vous trouverez ci-dessous deux images de 800×504 pixels. Ils ont la même taille en dimensions, mais la taille de leurs fichiers est très différente, car l'image de gauche contient beaucoup plus d'informations ou de variation de pixels que le logo de droite.



Photo as JPEG

Dimensions/Image Size: 800×504 px
Resolution: 72 ppi
File Size: 346 KB



Illustration as PNG

Dimensions/Image Size: 800×504 px
Resolution: 72 ppi
File Size: 36 KB

FIGURE 2.15 – exemple avec la Taille des images .

La compression photo prend une image et la compresse, ce qui réduit la taille du fichier. Il existe plusieurs méthodes de compression, cependant, la compression générale tente de mélanger des couleurs similaires (un processus appelé «quantification»). Ce processus mélange des pixels similaires pour former un seul pixel, réduisant ainsi la taille de

résolution de l'image (ainsi que le temps de chargement, économisant la bande passante). La compression réduit la taille du fichier d'une image de 60 à 70% en moyenne.

Les deux images ci-dessous sont évidemment la même image. Celui de gauche n'a pas été compressé, tandis que l'image de droite a été compressée. Notez qu'il n'y a aucune différence dans la qualité de l'image. Ils se ressemblent. Cependant, la taille du fichier de l'image de gauche est 630% plus grande. Cela fait une grande différence dans la vitesse de chargement des pages. [25]

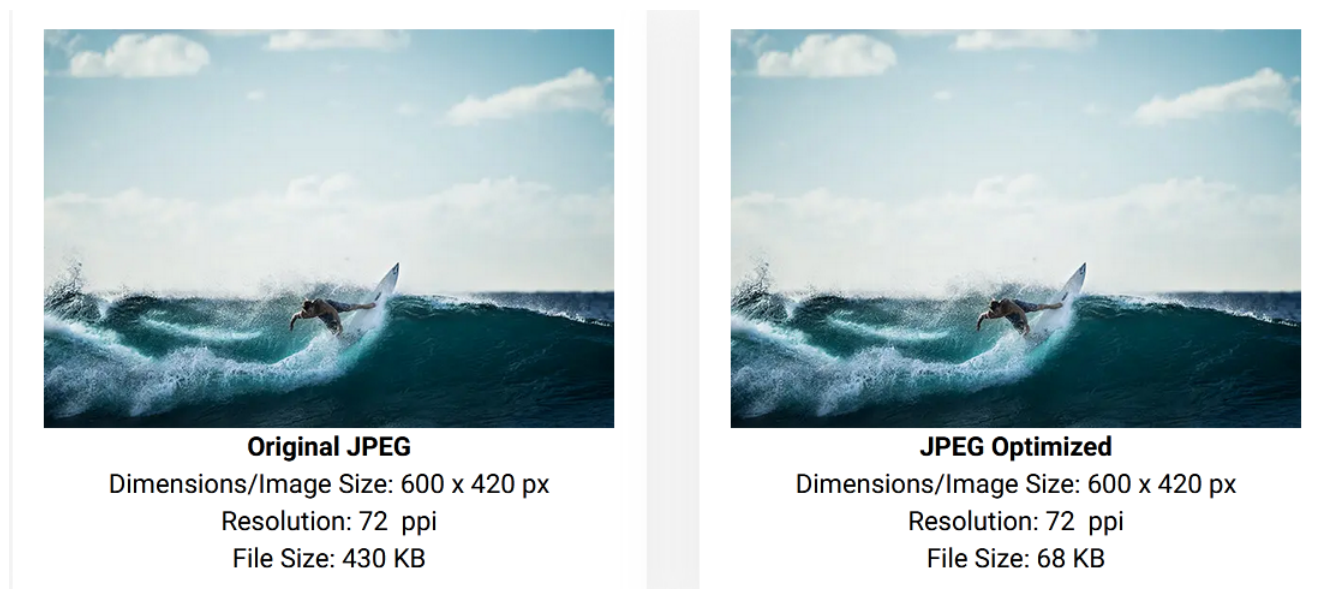


FIGURE 2.16 – exemple avec la compression d'image.

6.9 Représentation d'une image :

Représentation d'image : En informatique, nous pouvons représenter une image sous diverses formes. La plupart du temps, cela fait référence à la manière dont les informations sont transmises, telles que le codage numérique de la couleur, et à la manière dont l'image est stockée, c'est-à-dire à la structure d'un fichier image.

Plusieurs normes ouvertes ont été recommandées pour créer, manipuler, stocker et échanger des images numériques. Les règles décrivaient le format des fichiers image, les algorithmes de codage d'image, la forme d'informations supplémentaires souvent appelées métadonnées.

Une image numérique est la composition de pixels individuels ou d'éléments d'image. Les pixels sont disposés sous forme de ligne et de colonne pour former une zone d'image. Le nombre de pixels dans une image est fonction de la taille de l'image et du nombre

de pixels par unité de longueur (par exemple, pouce) dans la direction horizontale et verticale.

6.10 Format des fichiers image :

Il existe différents types d'images qui sont mentionnés comme :

6.10.1 JPEG (Joint Photographic Experts Group) :

il est utilisé pour les images numériques, en particulier pour les images composées de photographies numériques. L'extension de nom de fichier '.jpeg' est utilisée pour enregistrer le fichier.

6.10.2 PNG (Portable Network Graphics) :

ces fichiers sont couramment utilisés pour stocker les graphiques des images Web. PNG a été développé pour améliorer le remplacement non enregistré du Graphics Interchange Format. L'extension de nom de fichier '.png' est utilisée pour enregistrer le fichier.

6.10.3 GIF (Graphics Interchange Format) :

C'est un format de fichier pour stocker des images graphiques jusqu'à 256 couleurs. PNG est basé sur une méthode de compression sans perte, qui permet une sortie de meilleure qualité. PNG a été créé comme une option plus puissante pour le format de fichier GIF. L'extension de nom de fichier «.gif» est utilisée pour enregistrer le fichier.

6.10.4 TIFF / TIF (Tagged Image File) :

ces fichiers peuvent être enregistrés dans une collection de formats de couleur et de nombreuses formes de compression. Le fichier TIFF est utilisé pour maintenir l'intégrité et la clarté de l'image. Il est souvent utilisé pour la photographie professionnelle. L'extension de nom de fichier «.tif» est utilisée pour enregistrer le fichier.

6.10.5 PSD (Photoshop Document) :

il s'agit d'un fichier image en couches utilisé dans Adobe PhotoShop. Il s'agit d'un format par défaut utilisé par PhotoShop pour enregistrer les données. PSD est un fichier de garde qui permet à l'utilisateur de travailler avec les calques séparés des images même après que le fichier a été enregistré. L'extension de nom de fichier «.psd» est utilisée pour enregistrer le fichier.

6.10.6 PDF (Portable Document Format) :

il est utilisé pour partager les documents entre les ordinateurs et entre les plateformes du système d'exploitation lorsque l'utilisateur a besoin d'enregistrer des fichiers qui ne peuvent pas être modifiés. L'extension de nom de fichier «.pdf» est utilisée pour enregistrer le fichier.

6.10.7 EPS (Encapsulated Postscript) :

Il s'agit d'un format de fichier graphique utilisé dans les images vectorielles. Sous Windows, l'utilisateur a besoin d'un logiciel graphique pour ouvrir le fichier EPS (c'est-à-dire Adobe Illustrator, Coral Draw). L'extension de nom de fichier «.eps» est utilisée pour enregistrer le fichier.

6.10.8 AI (Adobe Illustrator Document) :

C'est un format de fichier développé par Adobe system. Il est utilisé pour représenter des dessins vectoriels d'une seule page aux formats EPS ou PDF. L'extension de nom de fichier «.ai» est utilisée pour enregistrer le fichier. [26]

6.11 Codage de la couleur :

Il existe plusieurs modes de codage des couleurs d'une image numérique, le plus utilisé pour le maniement des images est l'espace colorimétrique Rouge, Vert, Bleu (RVB ou RVG : Red Green Blue) par synthèse additive.

Une image RVB est composée de la somme des trois rayonnements lumineux Rouge, Vert, Bleu dont les faisceaux sont superposés. A l'intensité maximale ils produisent une lumière blanche. La gamme des couleurs reproductibles par ce mode, quoique conditionnée par la qualité du matériel employé, est très étendue, et reproduit bien les couleurs saturées. En contrepartie, elle convient mal à la restitution des nuances délicates des lumières intenses et des tons pastels.

Le codage de la couleur est réalisé sur 3 octets dont les valeurs codent la couleur dans l'espace RVB. Chaque octet représente la valeur d'une composante couleur par un entier de 0 à 255. Le nombre de couleurs différentes est de $256 \times 256 \times 256 = 16,8$ Millions. Une image numérique RVB est représentée par 3 tableaux à 2 dimensions dont la taille dépend du nombre de pixels contenus dans l'image.

Dans ce type d'image seul le niveau de l'intensité est codé sur un octet (256 valeurs). Par convention, la valeur 0 représente le noir (intensité lumineuse nulle) et la valeur 255



FIGURE 2.17 – exemple a rgb colors.

<i>R</i>	<i>V</i>	<i>B</i>	Couleur
0	0	0	Noir
0	0	1	Nuance de Noir
0	0	255	Bleu
0	255	0	Vert
255	0	0	Rouge
128	128	128	Gris
255	128	0	Orange
128	0	128	Violet
255	255	255	Blanc

FIGURE 2.18 – tableau de représentation d'Une image numérique en niveau de gris.

le blanc (intensité lumineuse maximale) :

Pour passer d'une image couleur à une image en niveau de gris, on utilise la formule :

$$G = 0,299 \times R + 0,587 \times G + 0,144 \times B$$

Une image numérique en niveau de gris est représentée par 1 tableau à 2 dimensions dont la taille dépend du nombre de pixels contenus dans l'image. [27] Les différents systèmes de colorimétrie :

0	8	16	32	56	72	90	104	112	128
136	144	160	176	192	208	224	244	248	255

FIGURE 2.19 – table de codage des couleurs .

<i>Mode</i>	<i>Nb de bits par pixels</i>	<i>Nombre de couleurs</i>	<i>Remarques</i>
Monochrome ou Noir et Blanc	1	2	Système utilisé pour scanner les textes pour faire de la reconnaissance de texte (OCR)
Niveaux de gris	8	256	Nuance de gris
Mode 4 bits ou 16 couleurs	4	16	Palette de couleurs peu étendue réservée aux dessins simples sans couleurs nuancée
Mode 8 bits ou 256 couleurs	8	256	Palette de 256 couleurs qui permet de conserver une taille raisonnable
Mode 16 bits	16	65536	Palette de 64536 couleurs qui convient pour la plupart des usages
Mode 24 bits ou Couleurs RVB	24	16,7 millions	Mode utilisé par défaut par de nombreux logiciels
Couleurs CMJN	32	4,3 Milliards	4 couleurs primaires : Cyan, Magenta, Jaune et Noir (256 teintes). Utilisé par les imprimantes

FIGURE 2.20 – table représente Les différents systèmes de colorimétrie .

7 Système de traitement d'image :

Dans le contexte de la vision artificielle, le traitement d'images se place après les étapes d'acquisition et de numérisation, assurant les transformations d'images et la partie de calcul permettant d'aller vers une interprétation des images traitées. Cette phase d'interprétation est d'ailleurs, de plus en plus intégrée dans le traitement d'images, en faisant appel notamment à l'intelligence artificielle pour manipuler des connaissances, principalement sur les informations dont on dispose à propos de ce que représentent les images traitées (connaissance du domaine). Un système de traitement d'image est généralement composé des unités suivantes :

- Un système d'acquisition et de numérisation qui permet d'effectuer l'échantillonnage et la quantification d'une image.
- Une mémoire de masse pour stocker les images numérisées.
- Un système de visualisation.
- Une unité centrale permettant d'effectuer les différentes opérations de traitement d'images.

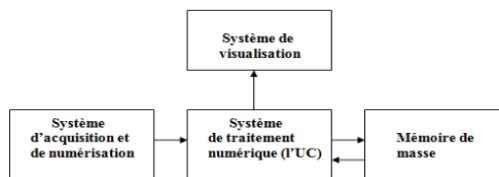


FIGURE 2.21 – Composition d'un système de traitement numérique.

8 Techniques de traitement d'images :

Dans le processus d'analyse d'une image, la phase de prétraitement a pour but d'améliorer l'homogénéité des régions qui la compose, tout en conservant, voire en renforçant le contraste entre régions adjacentes.

Il existe déjà plusieurs techniques qui utilisent l'image sous différents aspects :

8.1 Filtrage Numérique :

Les images brutes permettent rarement de parvenir à une extraction directe des objets à analyser :

- Soit parce que l'éclairage de l'objet n'est pas uniforme.
- Soit parce que l'objet est perçu à travers un bruit assez important : les images contiennent donc un signal et du bruit (dont on veut éliminer la plus grande partie possible).
- Soit encore parce que le contraste n'est pas suffisant.

8.2 Filtres linéaire :

L'utilisation des filtres linéaires provient de l'extension des méthodes mises au point pour le traitement du signal, au traitement des images numériques. Nous pouvons, en effet, considérer une image numérique comme un signal bidimensionnel numérisé. Un filtre est linéaire s'il peut s'écrire de la manière suivante :

Soit F le filtre utilisé, $I1$ et $I2$ représentant deux images alors :

$$F * (I1 + I2) = F * I1 + F * I2$$

L'opération $*$ représente le produit de convolution.

Les filtres linéaires les plus connus sont les filtres passe-bas, passe-haut. Ces filtres sont représentés sous forme de masque auquel est associé un coefficient diviseur qui permet

de ramener les valeurs calculées entre 0 et 255 pour une image à 256 niveaux de gris.

8.3 Filtre passe-bas (lissage) :

L'intensité du pixel considéré est remplacée par la moyenne des pixels de son voisinage. La taille de la zone (fenêtre) entourant le pixel est un paramètre important, plus cette dimension est grande, plus sa sensibilité au bruit diminue, et le lissage devient important (le flou s'accroît). Le filtre passe-bas laisse passer les basses fréquences (les faibles changements d'intensité de l'image) et atténue les hautes fréquences (variations rapides).

$$1/9 * \begin{array}{|c|c|c|} \hline \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \hline \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \hline \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \hline \end{array} \quad \textit{Masque du filtre}$$

FIGURE 2.22 – Masque de filtre passe-bas.

8.4 Filtre passe-haut :

Il est utilisé pour amplifier les détails de hautes fréquences. Il peut permettre par exemple de restaurer des images qui ont été défocalisées et d'accroître les contours en faisant ressortir les pixels compris entre des zones homogènes. Il met en évidence les changements rapides de l'intensité de l'image (les hautes fréquences) et laisse les zones uniformes inchangées (basses fréquences).

$$\mathbf{H=} \begin{array}{|c|c|c|} \hline \mathbf{-1} & \mathbf{-1} & \mathbf{-1} \\ \hline \mathbf{-1} & \mathbf{9} & \mathbf{-1} \\ \hline \mathbf{-1} & \mathbf{-1} & \mathbf{-1} \\ \hline \end{array} \quad \textit{Masque du filtre}$$

FIGURE 2.23 – Masque de filtre passe-haut.

8.5 Filtre gaussien :

L'intérêt de ce filtre est que l'on contrôle facilement le degré de filtrage à travers le paramètre K . Le filtre gaussien donne plus d'importance aux pixels proches du pixel central, et diminue cette importance au fur et à mesure que l'on s'éloigne de celui-ci, il a les mêmes inconvénients que le filtre passe-bas, c'est-à-dire il dégrade les contours. La discrétisation de ce filtre pour un K égale à 0.6 donne le masque suivant :

$$1/16 * \begin{array}{|c|c|c|} \hline 1 & 2 & 1 \\ \hline 2 & 4 & 2 \\ \hline 1 & 2 & 1 \\ \hline \end{array} \quad \text{Masque du filtre}$$

FIGURE 2.24 – Masque de filtre passe-haut.

8.6 Filtres Non linéaire :

Ils sont conçus pour régler les problèmes des filtres linéaires, surtout pour ce qui concerne la mauvaise conservation des contours. Leur principe est le même que celui des filtres linéaires, il s'agit toujours de remplacer la valeur de chaque pixel par la valeur d'une fonction calculée dans son voisinage. La différence majeure, est que cette fonction n'est plus linéaire mais une fonction quelconque (elle peut inclure des opérateurs de comparaisons ou de classification). Les filtres non linéaires les plus connus sont :

- Filtre médian.
- Filtre maximum.
- Filtre minimum.

8.7 Filtre médian :

Ce filtre est très utilisé pour éliminer le bruit sur une image qui peut être de différentes origines (poussières, petits nuages, baisse momentanée de l'intensité électrique sur les capteurs, ...) et qui se traduit par des taches de faible dimension dont la distribution sur l'image est aléatoire. L'avantage de ce filtre réside dans le fait qu'il conserve les

contours alors que les autres types de filtres ont tendance à les adoucir. [Devaux, 97] Sur un voisinage à huit, le nouveau niveau de gris du pixel central est choisi comme étant la valeur médiane de tous les pixels de la fenêtre d'analyse, centrée sur ce dernier. Son avantage est qu'il garde la netteté des éléments qui constituent l'image sans étaler les transitions. L'algorithme de filtre médian est le suivant :

1. Trier les valeurs par ordre croissant.
2. Remplacer la valeur du pixel centrale par la valeur située au milieu de la triée.
3. Répéter cette opération pour tous les pixels de l'image.

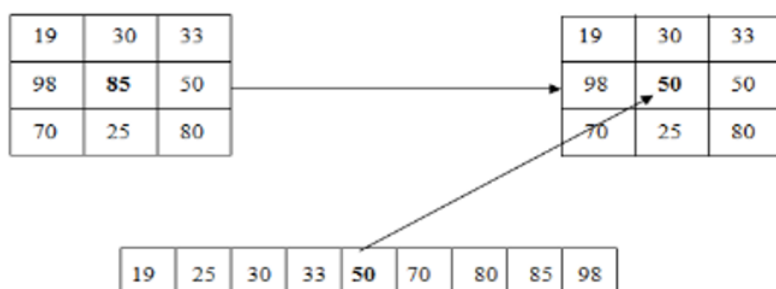


FIGURE 2.25 – Principe du filtre médiane.

8.8 Filtre Maximum :

On applique le même traitement que celui du filtre médian mais la valeur du pixel du centre va être changée par le maximum.

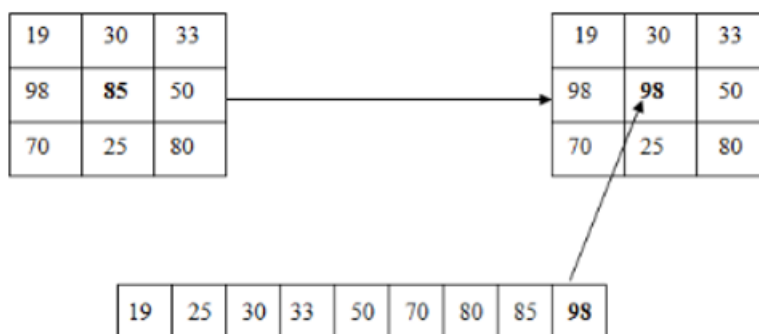


FIGURE 2.26 – Principe du filtre Maximum.

8.9 Filtre Minimum :

On applique le même traitement que celui du filtre maximum mais, cette fois, la valeur du pixel du centre comme la montre la figure va être remplacée par le minimum. [28]

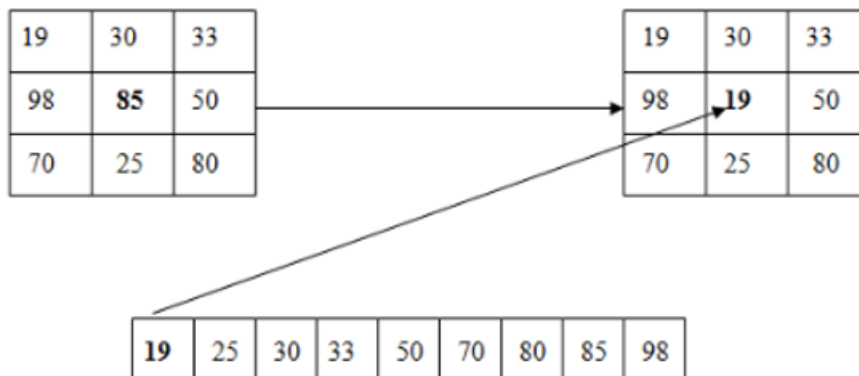


FIGURE 2.27 – Principe du filtre Minimum.

9 Classification des images numériques :

9.0.1 Classification d'image numérique :

Un analyste qui tente de classer les caractéristiques d'une image, utilise les éléments de l'interprétation visuelle pour identifier des groupes homogènes de pixels qui représentent des classes intéressantes de surfaces. La classification numérique des images utilise l'information spectrale contenue dans les valeurs d'une ou de plusieurs bandes spectrales pour classifier chaque pixel individuellement. Ce type de classification est appelé reconnaissance de regroupements spectraux. Les deux façons de procéder (manuelle ou automatique) ont pour but d'assigner une classe particulière ou thème (par exemple : eau, forêt de conifères, maïs, blé, etc.) à chacun des pixels d'une image. La "nouvelle" image qui représente la classification est composée d'une mosaïque de pixels qui appartiennent chacun à un thème particulier. Cette image est essentiellement une représentation thématique de l'image originale.

Lorsqu'on parle de classes, il faut faire la distinction entre des classes d'information et des classes spectrales. Les classes d'information sont des catégories d'intérêt que l'analyste tente d'identifier dans les images, comme différents types de cultures, de forêts ou d'espèce d'arbres, différents types de caractéristiques géologiques ou de roches, etc.

Les classes spectrales sont des groupes de pixels qui ont les mêmes caractéristiques (ou presque) en ce qui a trait à leur valeur d'intensité dans les différentes bandes spectrales des données. L'objectif ultime de la classification est de faire la correspondance entre les classes spectrales et les classes d'information. Il est rare qu'une correspondance directe soit possible entre ces deux types de classes. Des classes spectrales bien définies peuvent apparaître parfois sans qu'elles correspondent nécessairement à des classes d'information intéressantes pour l'analyse. D'un autre côté, une classe d'information très large (par exemple la forêt) peut contenir plusieurs sous-classes spectrales avec des variations spectrales définies. En utilisant l'exemple de la forêt, les sous-classes spectrales peuvent être causées par des variations dans l'âge, l'espèce, la densité des arbres ou simplement par les effets d'ombrage ou des variations dans l'illumination. L'analyste a le rôle de déterminer de l'utilité des différentes classes spectrales et de valider leur correspondance à des classes d'informations utiles.

Les méthodes de classification les plus communes peuvent être séparées en deux grandes catégories : les méthodes de classification supervisée et les méthodes de classification non supervisée.

9.1 Les méthodes de classification supervisée :

Lors de l'utilisation d'une méthode de classification supervisée, l'analyste identifie des échantillons assez homogènes de l'image qui sont représentatifs de différents types de surfaces (classes d'information). Ces échantillons forment un ensemble de données-tests. La sélection de ces données-tests est basée sur les connaissances de l'analyste, sa familiarité avec les régions géographiques et les types de surfaces présents dans l'image. L'analyste supervise donc la classification d'un ensemble spécifique de classes. Les informations numériques pour chacune des bandes et pour chaque pixel de ces ensembles sont utilisées pour que l'ordinateur puisse définir les classes et ensuite reconnaître des régions aux propriétés similaires à chaque classe. L'ordinateur utilise un programme spécial ou algorithme afin de déterminer la "signature" numérique de chacune des classes. Plusieurs algorithmes différents sont possibles. Une fois que l'ordinateur a établi la signature spectrale de chaque classe à la classe avec laquelle il a le plus d'affinités. Une classification supervisée commence donc par l'identification des classes d'information qui sont ensuite utilisées pour définir les classes spectrales qui les représentent.

9.2 Les méthodes de classification non supervisée :

La classification non supervisée procède de la façon contraire. Les classes spectrales sont formées en premier, basées sur l'information numérique des données seulement. Ces classes sont ensuite associées, par un analyste, à des classes d'information utile (si pos-

sible). Des programmes appelés algorithmes de classification sont utilisés pour déterminer les groupes statistiques naturels ou les structures des données.

Habituellement, l'analyste spécifie le nombre de groupes ou classes qui seront formés avec les données. De plus, l'analyste peut spécifier certains paramètres relatifs à la distance entre les classes et la variance à l'intérieur même d'une classe. Le résultat final de ce processus de classification itératif peut créer des classes que l'analyste voudra combiner, ou des classes qui devraient être séparées de nouveau. Chacune de ces étapes nécessite une nouvelle application de l'algorithme. L'intervention humaine n'est donc pas totalement exempte de la classification non supervisée. Cependant, cette méthode ne commence pas avec un ensemble prédéterminé de classes comme pour la classification supervisée. [29]

10 Domaines d'application :

Étant donné que le traitement d'image numérique a des applications très larges et que presque tous les domaines techniques sont touchés par ce dernière, nous allons simplement voir de certaines des principales applications de traitement d'image numérique. Le traitement d'image numérique ne se limite pas seulement à ajuster la résolution spatiale des images quotidiennes capturées par l'appareil photo. Il ne s'agit pas seulement d'augmenter la luminosité de la photo, c'est bien plus que cela.

11 Applications du traitement d'image numérique :

Certains des principaux domaines dans lesquels le traitement d'image numérique est largement utilisé sont mentionnés ci-dessous :

- Renforcement et restauration d'image.
- Domaine médical.
- Télédétection.
- Transmission et encodage.
- Vision machine / robot .
- Traitement des couleurs.
- La reconnaissance de formes.
- Traitement vidéo.
- Imagerie microscopique. [30]

12 Conclusion :

Dans ce chapitre, nous avons parlé du traitement et de la classification des images médical en détail, ce qui est essentiel pour comprendre nos recherches. Le prochain chapitre sera consacré sur l'apprentissage en profondeur(deep Learning) qui sont largement utilisés dans les problèmes de traitement d'image.

CHAPITRE

3

L'APPRENTISSAGE EN
PROFONDEUR(DEEP LEARNING)

1 Introduction

Avec l'avènement de l'intelligence artificielle, l'essor des objets connectés et l'augmentation des capacités technologiques, nombreuses sont les recherches tentant de simuler le comportement humain et également de reproduire ses capacités cognitives. Dans ce contexte le Deep Learning, littéralement "l'apprentissage profond", est actuellement un des axes de recherche les plus explorés.[31]

Ce chapitre va couvrir qu'est-ce donc que le Deep Learning, qu'est-ce qu'un réseau de neurones, quel lien avec notre cerveau humain, comment cela fonctionne ,les bases des CNNs en incluant la description des différentes couches utilisées.

2 Qu'est ce que le deep learning ?

L'apprentissage profond (deep learning en anglais) est un ensemble de méthodes d'apprentissage automatique (machine learning en anglais) qui tentent de modéliser des données avec un haut niveau d'abstraction. Il s'agit d'automatiser une tâche simple, répétitive et calculatoire pour laquelle on a de nombreux exemples. Un tel algorithme consiste en quelque sorte à compresser de l'information contenue dans le signal d'entrée que ce soit une image ou un son dans une classe d'images (présence d'un visage ou absence d'un visage par exemple).

Actuellement les algorithmes les plus importants issus de l'apprentissage profond sont les réseaux de neurones profonds (deep neural networks) car ils ont obtenu des résultats impressionnants ces dernières années. Ce sont l'état de l'art pour la classification d'images avec des centaines de classes complexes à différencier, pour la reconnaissance de sons, la compréhension du langage naturel, la conduite des voitures autonomes, la classification de galaxies, détection de piétons, l'aide au diagnostic médicale (aide à la décision pour la lecture d'imagerie médicale).[32]

3 Fonctionnement du deep Learning

Le deep Learning s'appuie sur un réseau de neurones artificiels s'inspirant du cerveau humain. Ce réseau est composé de dizaines voire de centaines de "couches" de neurones, chacune recevant et interprétant les informations de la couche précédente. Le système apprendra par exemple à reconnaître les lettres avant de s'attaquer aux mots dans un texte, ou détermine s'il y a un visage sur une photo avant de découvrir de quelle personne il s'agit. [33]

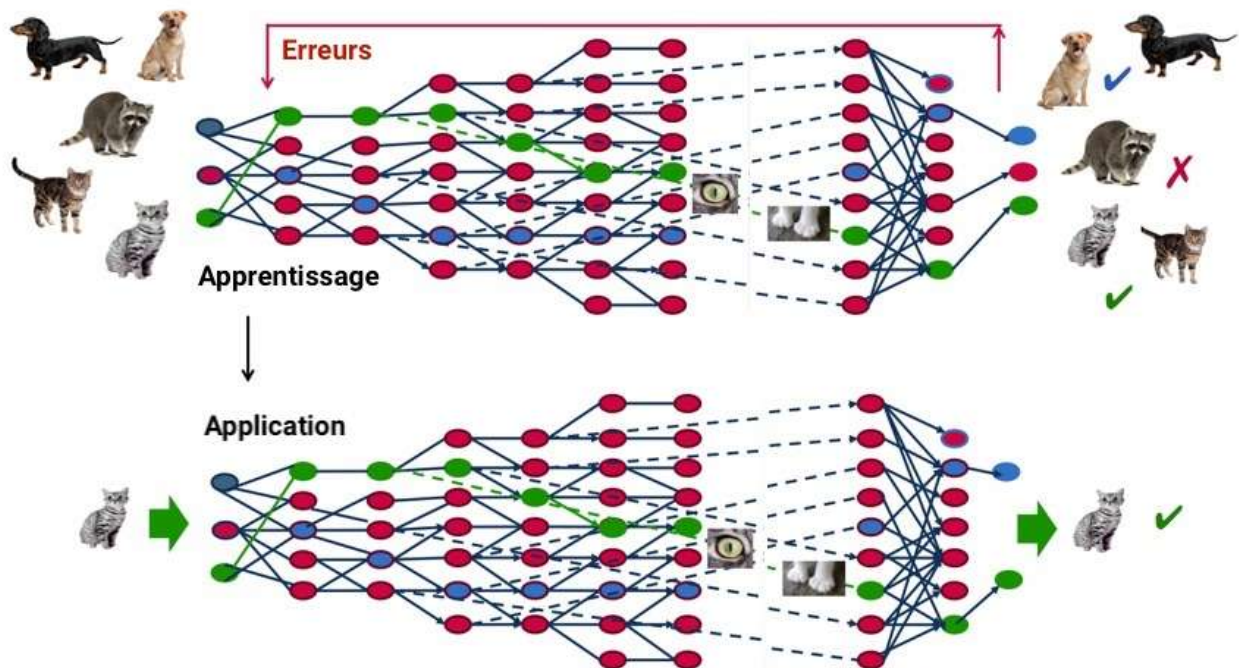


FIGURE 3.1 – Schéma de fonctionnement de Deep Learning .

4 Applications du deep Learning

Le deep Learning est utilisé dans de nombreux domaines :

- reconnaissance d'image,
- traduction automatique,
- voiture autonome,
- diagnostic médical,
- recommandations personnalisées,
- modération automatique des réseaux sociaux,
- prédiction financière et trading automatisé,
- identification de pièces défectueuses,
- détection de malwares ou de fraudes,
- chatbots (agents conversationnels),
- exploration spatiale,
- robots intelligents.

C'est aussi grâce au deep Learning que l'intelligence artificielle de Google Alpha Go a réussi à battre les meilleurs champions de Go en 2016. Le moteur de recherche du géant américain est lui-même de plus en plus basé sur l'apprentissage par deep Learning plutôt

que sur des règles écrites.

Aujourd'hui le deep Learning est même capable de "créer" tout seul des tableaux de Van Gogh ou de Rembrandt, d'inventer un langage totalement nouveau pour communiquer entre deux machines. [33]

5 Exemples d'applications en histoire de l'art

Nous allons maintenant examiner certaines initiatives de recherche intéressantes mêlant apprentissage profond et oeuvres d'art. Ce type d'initiatives sont destinées à être de plus en plus nombreuses en raison de la numérisation croissante des collections d'oeuvres d'art.

Tout d'abord, l'équipe du Visual Geometry Group de l'université d'Oxford (Crowley and Zisserman), effectue de la recherche en ligne d'objets dans une base de données de plus de 200.000 images ("Yours Paintings") . N'importe quel mot peut être cherché à travers cette base d'images comme on fait une recherche sur Google Images. De manière schématique, lorsque l'utilisateur soumet une requête textuelle (par exemple oiseau, chien, train), des images contenant cet objet sont téléchargées depuis Google Images. Celles-ci sont utilisées pour apprendre un classificateur à la volée (modifier les paramètres de l'algorithme). Cet algorithme est ensuite appliqué à toutes les images de la base de peintures pour récupérer des images contenant cet objet. La démonstration est disponible sur le site de l'équipe : http://www.robots.ox.ac.uk/~vgg/research/art_search/.

Une équipe britannique à l'Université de Bristol (<https://research-information.bristol.ac.uk/files/117029431>) a quant à elle modifié un algorithme d'apprentissage profond destiné à détecter des personnes dans les photos pour l'utiliser pour détecter des personnes dans des oeuvres d'art (dans des peintures tout particulièrement). Dans ce cas, on va chercher à utiliser un algorithme plutôt performant pour une tâche donnée : détecter des personnes dans des images en donnant sa position exacte dans la photographie (et pour laquelle on a de très nombreux exemples : plusieurs milliers) pour effectuer une nouvelle tâche : détecter une personne dans une photographie de peinture (pour laquelle on a très peu d'exemple moins d'une centaine). Ces travaux mettent en lumière qu'il s'agit d'une tâche ardue : l'algorithme n'est capable de délimiter correctement les personnes que dans 60% des cas. Cela est dû au fait que les styles, les poses sont très différentes entre les oeuvres d'art. Cependant les pistes de recherche sont nombreuses.

Enfin, l'équipe de Frédéric Kaplan à l'EPFL de Lausanne travaille sur le projet Replica qui a pour objectif ambitieux de détecter automatiquement les ressemblances visuelles entre oeuvres d'art. Cela permettrait alors d'effectuer des recherches par ressem-

blances visuelles dans les bases de données en sélectionnant qu'un élément d'une image (sélectionné avec l'aide de la souris). Pour ce faire, ils utilisent un algorithme d'apprentissage profond pour mesurer une distance entre deux images. Une base de données très finement annotées par une historienne de l'art a évidemment été nécessaire. L'équipe a tout d'abord travaillé sur la base de données Web Gallery of Art mais un passage à l'échelle sur la base de la Fondation Cini sera faite dans le futur.

Vous pouvez visualiser une démonstration de ce moteur de recherche ici : <https://seguinbe.github.io/2016/12/19/replica-project-status-and-roadmap.html> La plateforme de leur projet permet aussi d'effectuer des requêtes plus complexes avec par exemple la possibilité de chercher des images qui présentent une ressemblance avec l'image A tout en évitant de retourner celle ressemblant à l'image B. -> Voir les articles sur le blog de l'APAHAU : retrouver les motifs similaires et CNN et recherche de schéma visuel.

Tous ces travaux sont des exemples d'usages d'outils issus des mathématiques et de l'informatique pour l'histoire de l'art. Il s'agit de projets académiques souvent nés des sciences dures. Il y a une véritable volonté de développer des outils et des méthodes pour d'autres domaines de recherche. Il ne s'agit pas de remplacer un historien ou un biologiste mais de les aider dans leur travaux et de valoriser un patrimoine culturel en permettant une navigation bien plus fluide à travers ces immenses bases de données.[32]

6 la relation entre Intelligence Artificielle, Machine Learning, Deep Learning :

Les domaines d'application et usages potentiels d'une Intelligence Artificielle sont de plus en plus divers : compréhension du langage naturel, reconnaissance visuelle, robotique, système autonome, Machine Learning ...

6.1 Le Machine Learning et le Deep Learning sont de l'Intelligence Artificielle

L'Intelligence Artificielle (IA) telle que nous la connaissons est une Intelligence Artificielle faible, par opposition à l'IA forte, qui n'existe pas encore. Aujourd'hui, les machines sont capables de reproduire un comportement humain, mais sans conscience. Plus tard, leurs capacités pourraient croître au point de se transformer en machines dotées de conscience, de sensibilité et d'esprit.

Si le Machine Learning (ML) et le Deep Learning (DL) sont des Intelligences Artificielles, l'inverse n'est pas vrai. Par exemple, les graphiques de connaissances ou les moteurs de règles sont des Intelligences Artificielles mais ne relèvent pas du ML ni du DL. Le Deep Learning est, quant à lui, une branche du Machine Learning.

L'IA a beaucoup évolué grâce notamment à l'émergence du Cloud Computing et du Big Data, soit d'une puissance de calcul peu coûteuse et de l'accessibilité à un grand nombre de données. Ainsi, les machines ne sont plus programmées ; elles apprennent.[34]

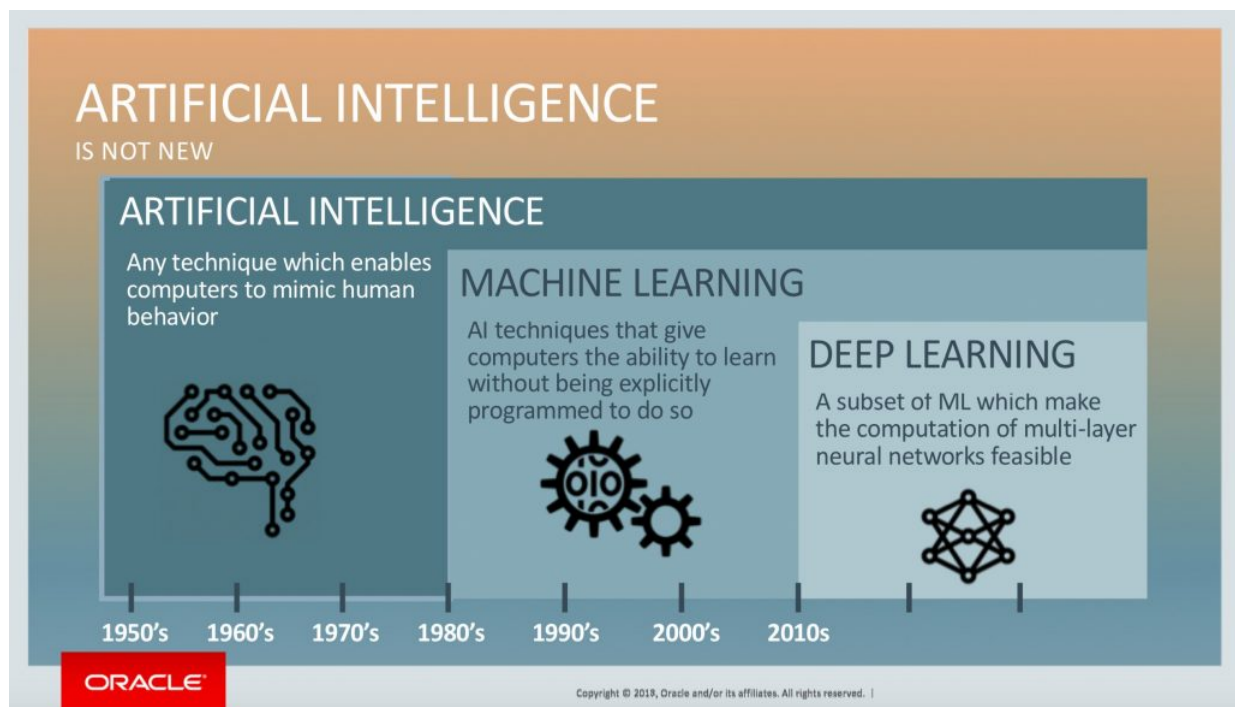


FIGURE 3.2 – Intelligence Artificielle, Machine Learning, Deep Learning

7 C'est quoi un réseau de neurones ?

est un système d'apprentissage informatique qui utilise un réseau de fonctions pour comprendre et traduire une entrée de données d'une forme en une sortie souhaitée, généralement sous une autre forme.[35]

Le concept du réseau neuronal artificiel est un simplification de notre outil le plus puissant, le cerveau. Il utilise des neurones (les cercles dans l'image ci-dessous) qui sont tous connectés les uns aux autres par des poids (les lignes dans l'image ci-dessous).[36]

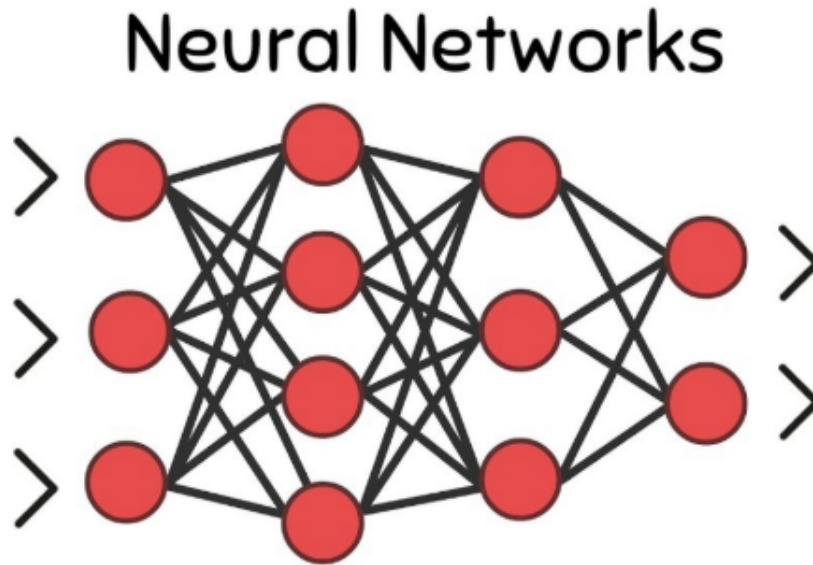


FIGURE 3.3 – le réseau de neurones

8 C'est quoi un CNN

Un réseau neuronal convolutif (CNN) est un type de réseau neuronal artificiel utilisé dans la reconnaissance et le traitement d'images qui est spécifiquement conçu pour traiter les données de pixels. Les CNN sont un puissant traitement d'image, une intelligence artificielle (AI) qui utilise l'apprentissage en profondeur pour effectuer des tâches à la fois génératives et descriptives tâches[37].

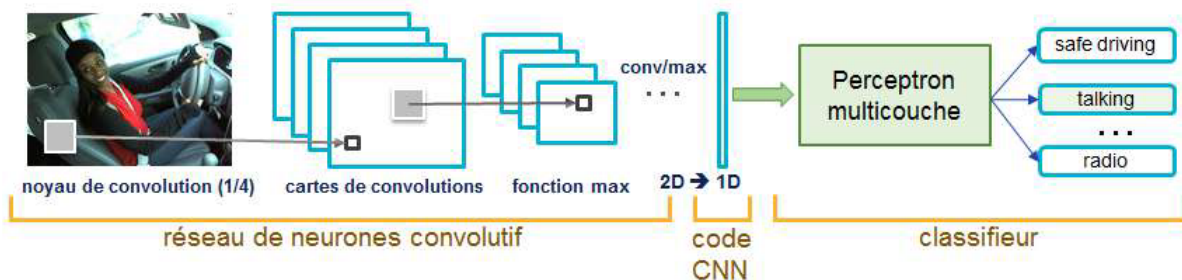


FIGURE 3.4 – Les réseaux de neurones convolutifs.

9 L'architecture d'un CNN

Une architecture de réseau de neurones convolutifs est formée par un empilement de couches de traitement :

1. la couche de convolution (CONV) qui traite les données d'un champ récepteur
2. la couche de pooling (POOL), qui permet de compresser l'information en réduisant la taille de l'image intermédiaire (souvent par sous-échantillonnage).
3. la couche de correction (ReLU), souvent appelée par abus " ReLU " en référence à la fonction d'activation (Unité de rectification linéaire).
4. la couche " entièrement connectée " (FC), qui est une couche de type perceptron.
5. la couche de perte (LOSS).

10 Les couche de CNN

10.1 Les couches convolutives (CONV)

La couche de convolution est le bloc de construction de base d'un CNN. Trois paramètres permettent de dimensionner le volume de la couche de convolution la profondeur, le pas et la marge.

10.1.1 Profondeur de la couche :

nombre de noyaux de convolution (ou nombre de neurones associés à un même champ récepteur).

10.1.2 Le pas :

contrôle le chevauchement des champs récepteurs. Plus le pas est petit, plus les champs récepteurs se chevauchent et plus le volume de sortie sera grand.

10.1.3 La marge (à 0) ou zero padding :

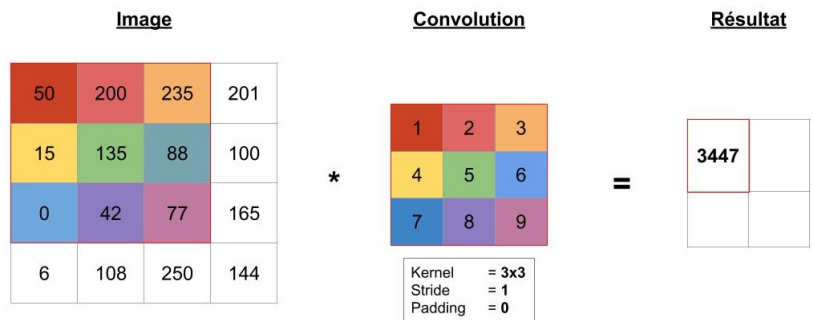
parfois, il est commode de mettre des zéros à la frontière du volume d'entrée. La taille de ce 'zero-padding' est le troisième hyper paramètre. Cette marge permet de contrôler la dimension spatiale du volume de sortie. En particulier, il est parfois souhaitable de conserver la même surface que celle du volume d'entrée.[38]

La taille spatiale du volume de sortie peut être calculée en fonction de la taille du volume d'entrée W_i la surface de traitement K (nombre de champs récepteurs), le pas S avec lequel ils sont appliqués, et la taille de la marge P La formule pour calculer le nombre de neurones du volume de sortie est $W_0 = \frac{w_i - k + 2p}{s} + 1$ Si W_0 n'est pas entier, les neurones périphériques n'auront pas autant d'entrée que les autres. Il faudra donc augmenter la taille de la marge (pour recréer des entrées virtuelles).

Souvent, on considère un pas $S = 1$, on calcule donc la marge de la manière suivante :

$$P = \frac{k - 1}{2}$$

si on souhaite un volume de sortie de même taille que le volume d'entrée. Dans ce cas particulier la couche est dite "connectée localement".



Pour calculer, on fait :
pixel1 de l'image x pixel1 de la convolution + pixel2 de l'image x pixel2 de la convolution + ...
 Ici, cela donne : $50*1 + 200*2 + 235*3 + 15*4 + 135*5 + 88*6 + 0*7 + 42*8 + 77*9 = 3447$

FIGURE 3.5 – Calcul d’une étape de la convolution.

Dans le réseau de neurones convolutifs, nous utilisons un «kernel» pour extraire certaines «fonctionnalités» d’une image d’entrée.

Laisse-moi expliquer. Un noyau est une matrice qui est glissée sur l’image et multipliée par l’entrée de telle sorte que la sortie est améliorée d’une certaine manière souhaitable(voir figure 1.4).

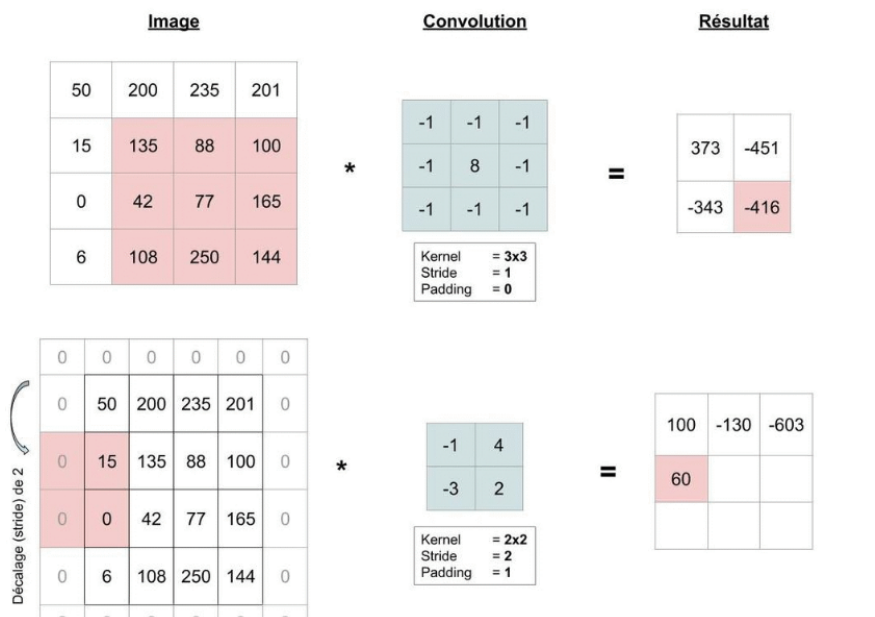


FIGURE 3.6 – Calcul de deux convolutions sur deux images, avec des kernels.

10.2 Les couches de regroupement /sous-échenillage (pooling/sub-sampling)

Le pooling est une opération simple qui consiste à remplacer un carré de pixels (généralement 2×2 ou 3×3) par une valeur unique. De cette manière, l'image diminue en taille et se retrouve simplifiée (lissée).

Pour appliquer le pooling, on commence par sélectionner un carré de pixels de taille 2×2 (pour un pooling de 2×2) puis on calcule la valeur qui va venir remplacer ce carré (cf type de pooling ci-après). Ensuite, on décale ce carré vers la droite de 1 cases si le stride (= pas) vaut 1 par exemple (généralement, il vaut 1 ou 2).

Plusieurs stratégies sont envisageables si la sélection sort de l'image : on peut ne rien calculer, ou calculer malgré tout en ignorant les pixels manquants. Mais en général, le stride est étudié pour que la sélection rentre parfaitement partout.

Une fois arrivé au bout à droite, on recommence tout à gauche en décalant une fois vers le bas (d'un pas égal au stride) et en allant de nouveau vers la droite.

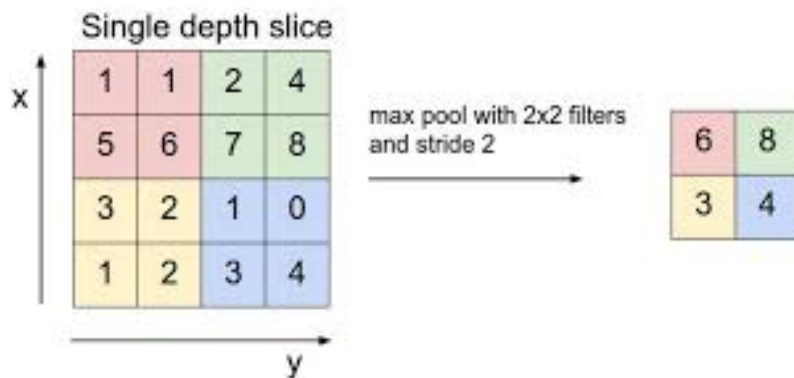


FIGURE 3.7 – Pooling avec un filtre 2x2 et un pas de 2.

Il existe plusieurs types de pooling :

1. Le « **max pooling** » , qui revient à prendre la valeur maximale de la sélection. C'est le type le plus utilisé car il est rapide à calculer (immédiat), et permet de simplifier efficacement l'image.
2. Le « **mean pooling** » (ou **average pooling**), soit la moyenne des pixels de la sélection : on calcule la somme de toutes les valeurs et on divise par le nombre de valeurs. On obtient ainsi une valeur intermédiaire pour représenter ce lot de pixels.
3. Le « **sum pooling** » , c'est la moyenne sans avoir divisé par le nombre de valeurs (on ne calcule que leur somme)[39].

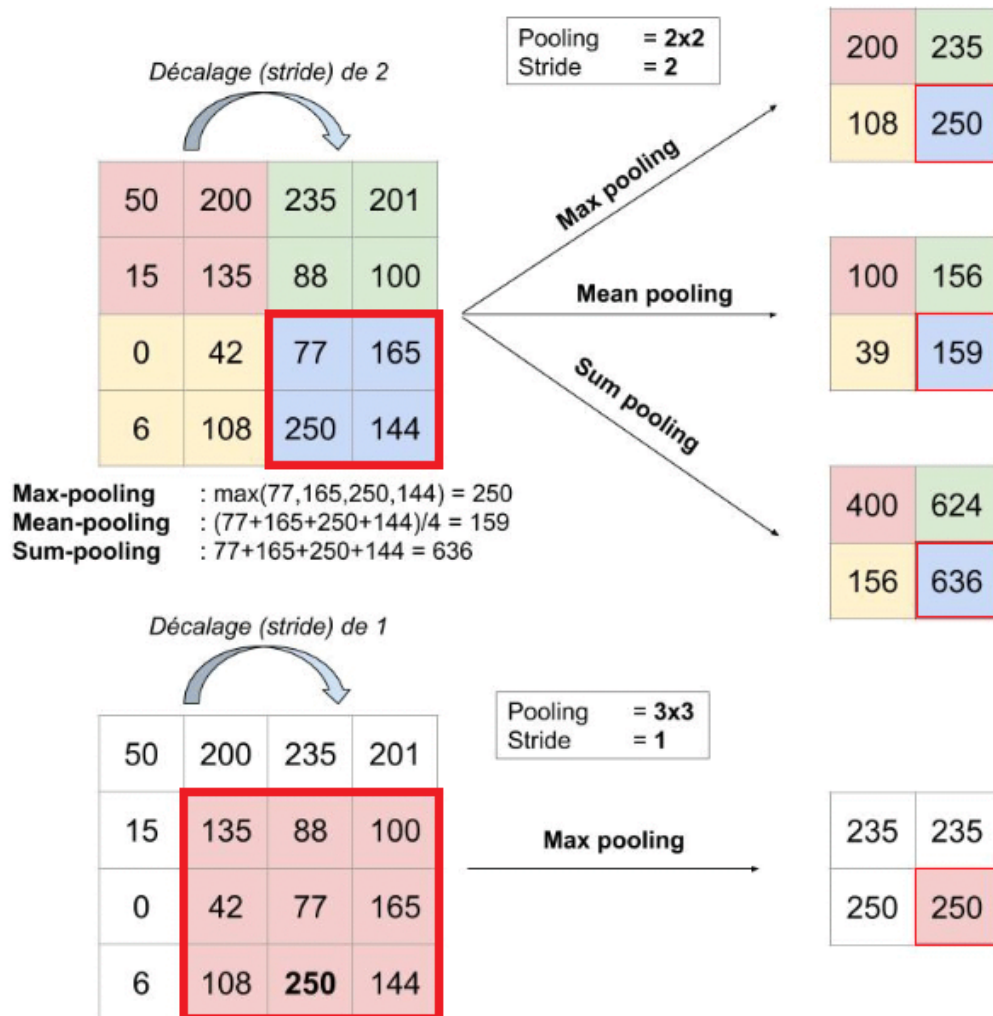


FIGURE 3.8 – Calcul du pooling sur une image 4x4.

10.3 Les couches non-linéaires :

Les réseaux de neurones en général et les CNN en particulier reposent sur une fonction «déclencheur» non linéaire pour signaler une identification distincte des caractéristiques probables sur chaque couche masquée.

10.3.1 ReLU (Rectified Linear Units) :

Après chaque couche conv, il est conventionnel d'appliquer une couche non linéaire (ou couche d'activation) immédiatement après .

Le but de cette couche est d'introduire la non-linéarité dans un système qui, fondamentalement, vient de calculer des opérations linéaires pendant les couches conv. Des fonctions non linéaires comme **tanh** et **sigmoïde** ont été utilisées, mais les chercheurs ont découvert que les couches ReLU fonctionnent beaucoup mieux parce que le réseau est capable de s'entraîner beaucoup plus rapidement sans faire de différence significative dans la précision.

La couche ReLU applique la fonction $f(x) = \max(0, x)$ à toutes les valeurs du volume d'entrée. En termes de base, cette couche modifie toutes les activations négatives à 0(voir figure 1.7). Cette couche augmente les propriétés non linéaires du modèle et du réseau global sans affecter les champs récepteurs de la couche conv[40].

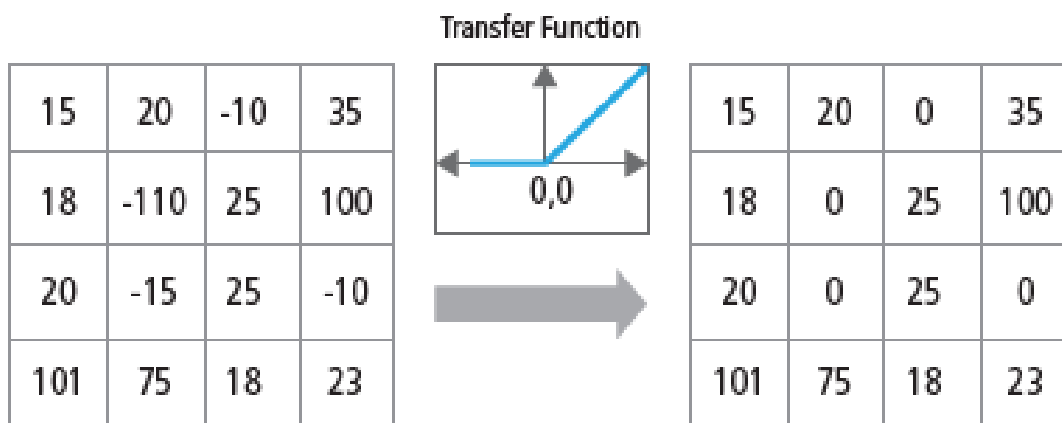


FIGURE 3.9 – La fonctionnalité ReLU.

10.3.2 Les fonctions de déclenchement continues (non linéaires) :

La couche non linéaire opère élément par élément dans chaque entité. Une fonction de déclenchement continue peut être hyperbolique tangente (figure 1.8), absolue de tangente hyperbolique (figure 1.9) ou sigmoïde (figure 1.10). La figure 1.11 montre comment la non-linéarité est appliquée élément par élément[7].

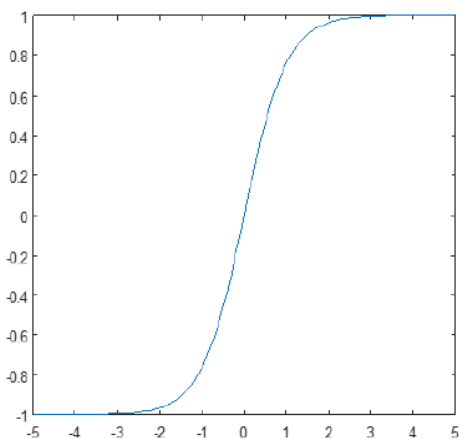


FIGURE 3.10 – Graphe de la fonction tangente hyperbolique.

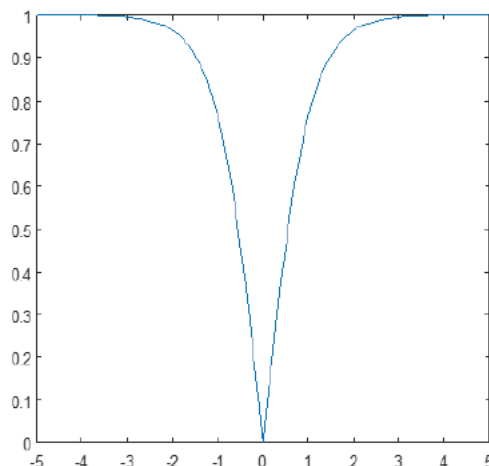


FIGURE 3.11 – Graphe de l'absolu de la fonction tangente hyperbolique.

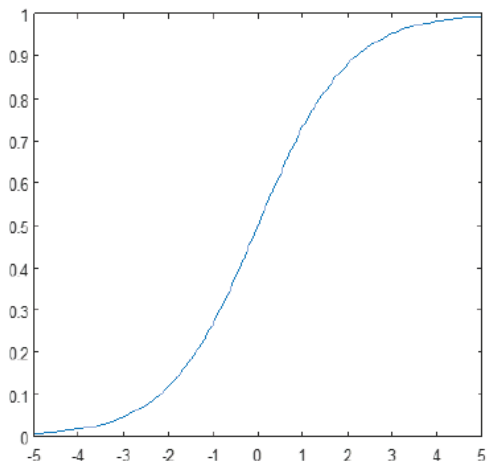


FIGURE 3.12 – Graphe de la fonction sigmoïde .

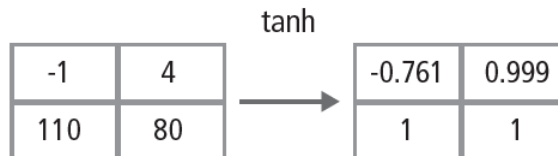


FIGURE 3.13 – Représentation du traitement de tanh .

10.3.3 Les couches entièrement connectées (FC) ou «dense» :

Après plusieurs couches de convolution et de max-pooling, le raisonnement de haut niveau dans le réseau neuronal se fait via des couches entièrement connectées.

Les neurones dans une couche entièrement connectée ont des connexions vers toutes les sorties de la couche précédente. Leurs fonctions d'activations peuvent donc être calculées avec une multiplication matricielle suivie d'un décalage de polarisation.[42].

La figure 1.12 explique la couche L entièrement connectée. La couche L-1 a deux caractéristiques, chacune étant 2x2, c'est-à-dire quatre éléments. La couche L possède deux caractéristiques, chacune ayant un seul élément[41].

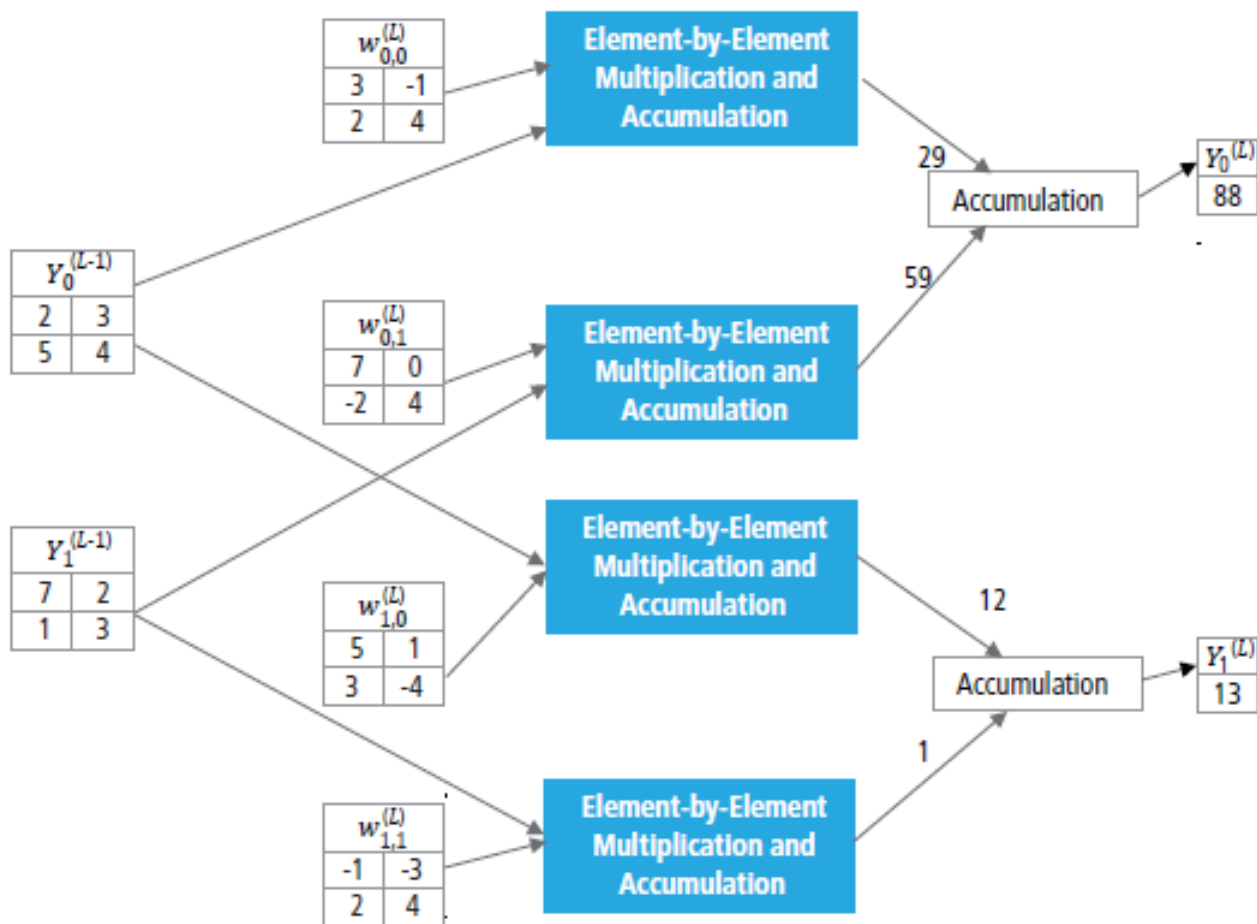


FIGURE 3.14 – Traitement d’une couche entièrement connectée.

10.3.4 La Couche de perte (LOSS) :

Pour apprendre un CNN, sa sortie doit finalement être convertie à un score scalaire E qui est ensuite optimisé. Normalement, ce score est le résultat de la composition du CNN avec une fonction de perte (), qui compare la sortie CNN y à une valeur de sortie de vérité terrain c .

La couche de perte spécifie comment l'entraînement du réseau pénalise l'écart entre le signal prévu et réel. Elle est normalement la dernière couche dans le réseau.

Diverses fonctions de perte adaptées à différentes tâches peuvent y être utilisées. On distingue trois types de fonctions de perte :

1. **La perte « Softmax »(multi-class logistic loss)**

Elle est utilisée pour prédire une seule classe parmi K classes mutuellement exclusives.

$$\ell(\mathbf{x}, \mathbf{c}) = \frac{1}{N} \sum_{n=0}^{N-1} \left[-x_{n,c_n} + \log \sum_{k=0}^{C-1} e^{x_{nk}} \right]$$

2. **La perte par entropie croisée sigmoïde (cross-entropy loss)**

Elle est utilisée pour prédire K valeurs de probabilité indépendante dans l'intervalle $[0,1]$. La formule générale de ce cas est la suivante dont N est la taille du tenseur(vecteur) :

$$\ell(\mathbf{y}, \mathbf{c}) = -\frac{1}{N} \sum_{n=0}^{N-1} \log y_{n,c_n}$$

3. **La perte euclidienne**

est utilisée pour régresser vers des valeurs réelles.[42]

11 Augmentation de données :

L'augmentation des données d'image est une technique qui peut être utilisée pour augmenter artificiellement la taille d'un ensemble de données d'entraînement en créant des versions modifiées des images dans l'ensemble de données.

L'entraînement des modèles de réseaux neuronaux d'apprentissage en profondeur sur plus de données peut aboutir à des modèles plus habiles, et les techniques d'augmentation peuvent créer des variations des images qui peuvent améliorer la capacité des modèles d'ajustement à généraliser ce qu'ils ont appris à de nouvelles images.

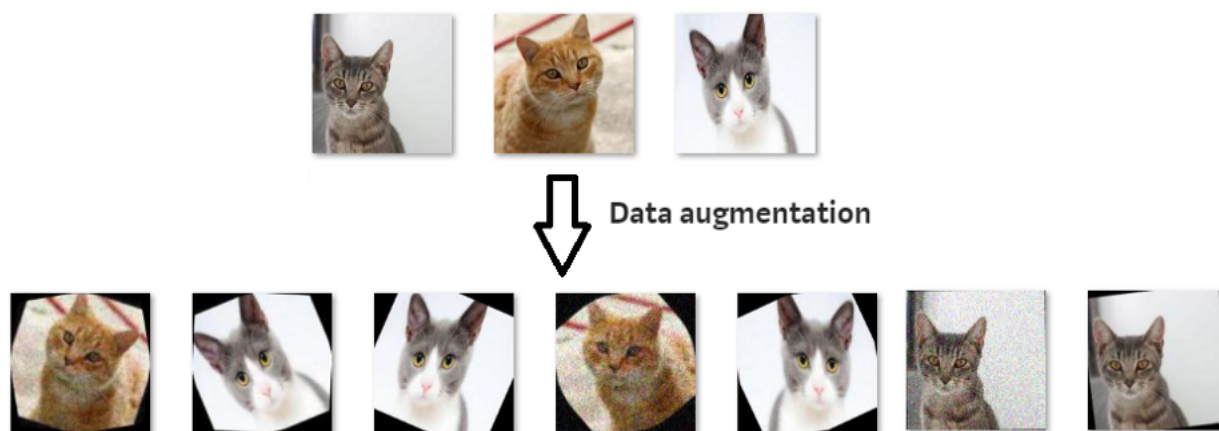


FIGURE 3.15 – exemple sur l'augmentation de données.

11.1 Rotation

Les augmentations de rotation se font en faisant pivoter l'image à droite ou à gauche sur un axe compris entre 1° et 359° . La sécurité des augmentations de rotation est fortement déterminée par le paramètre de degré de rotation. De légères rotations telles qu'entre 1 et 20 ou -1 à -20 pourraient être utiles sur les tâches de reconnaissance de chiffres telles que MNIST, mais à mesure que le degré de rotation augmente, l'étiquette des données n'est plus conservée après la transformation.

11.2 Translation

Le décalage des images vers la gauche, la droite, le haut ou le bas peut être une transformation très utile pour éviter un biais de position dans les données. Par exemple,

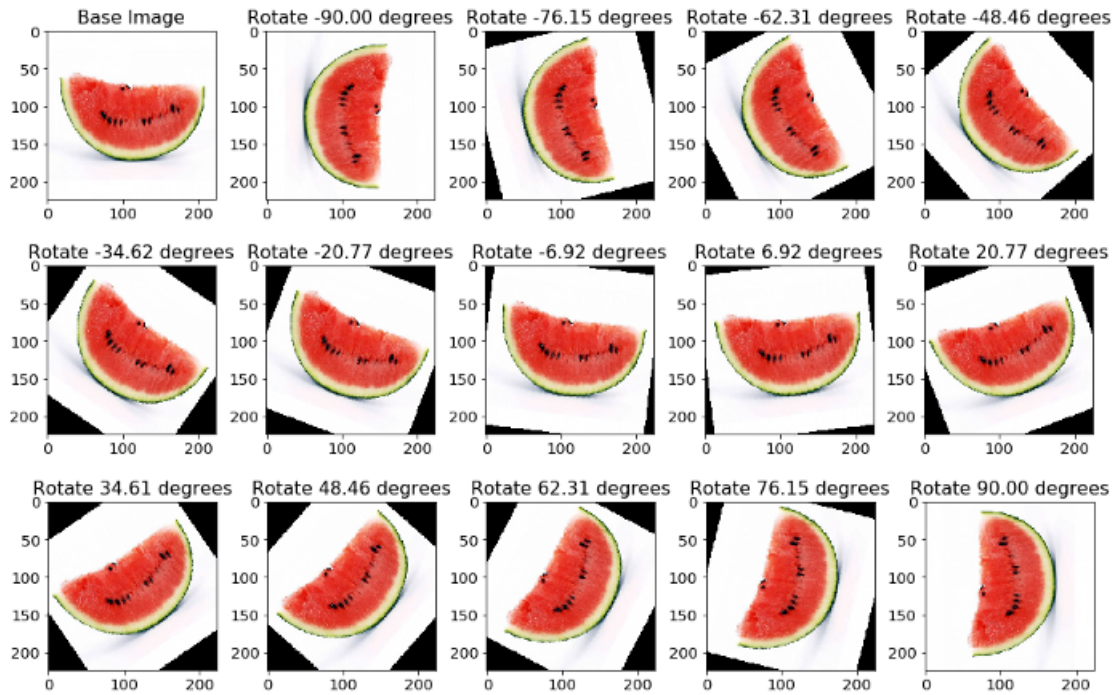


FIGURE 3.16 – Image avec l'effet de rotation.

si toutes les images d'un ensemble de données sont centrées, ce qui est courant dans les ensembles de données de reconnaissance faciale, cela exigerait que le modèle soit également testé sur des images parfaitement centrées.

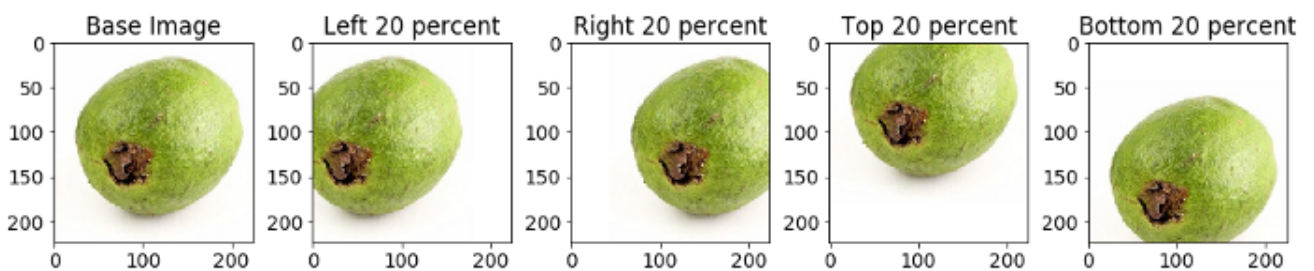


FIGURE 3.17 – Image avec l'effet de translation.

11.3 Retournement

Le retournement de l'axe horizontal est beaucoup plus courant que le retournement de l'axe vertical. Cette augmentation est l'une des plus faciles à mettre en œuvre et

s'est avérée utile sur des ensembles de données tels que CIFAR-10 et ImageNet. Sur les ensembles de données impliquant la reconnaissance de texte tels que MNIST ou SVHN, il ne s'agit pas d'une transformation préservant les étiquettes.

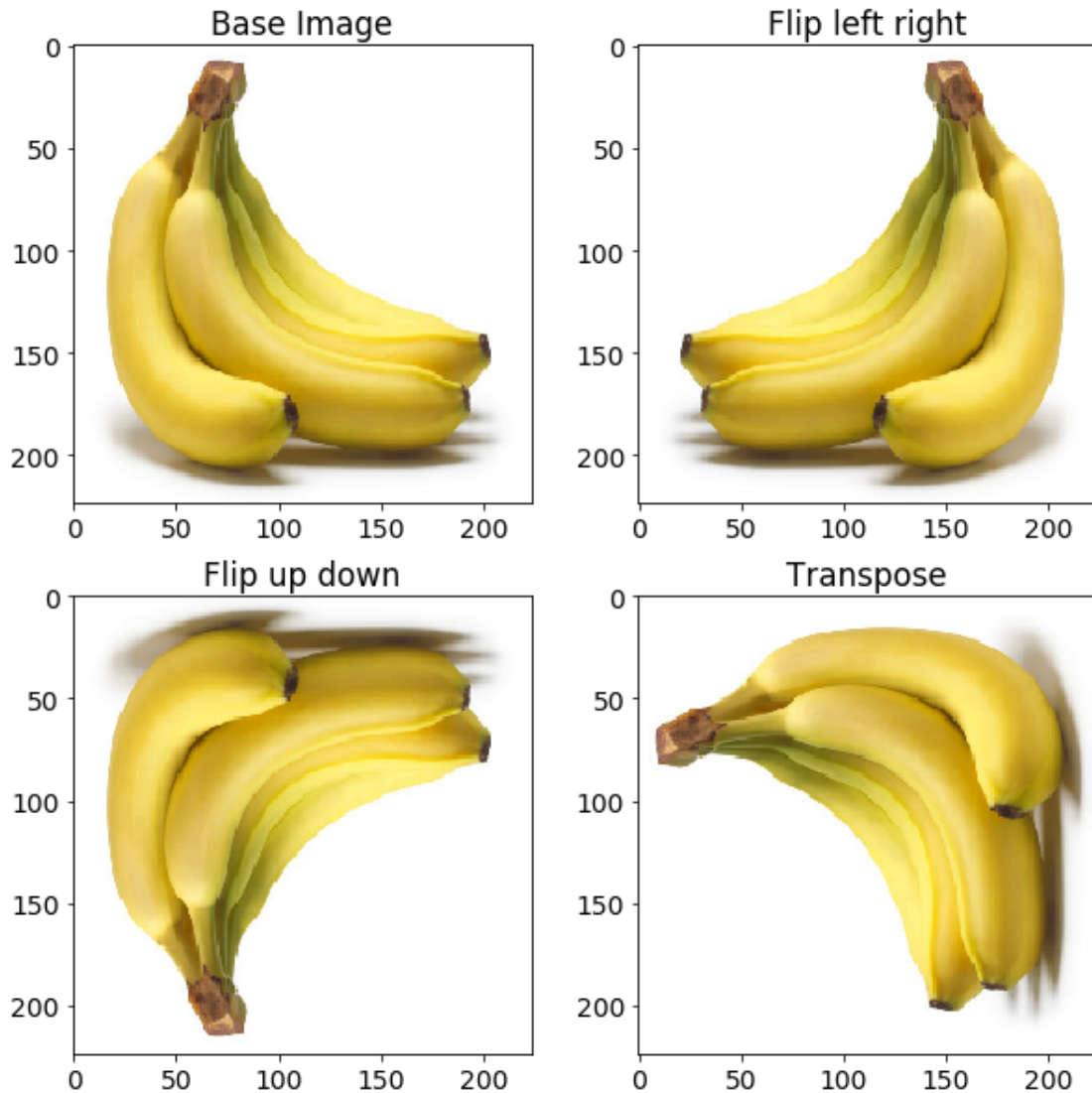


FIGURE 3.18 – Image avec l'effet de Retournement.

11.4 Echelle

L'image peut être redimensionnée vers l'extérieur ou vers l'intérieur. Lors de la mise à l'échelle vers l'extérieur, la taille de l'image finale sera plus grande que la taille de l'image d'origine. La plupart des cadres d'image découpent une section de la nouvelle image, avec une taille égale à l'image d'origine.

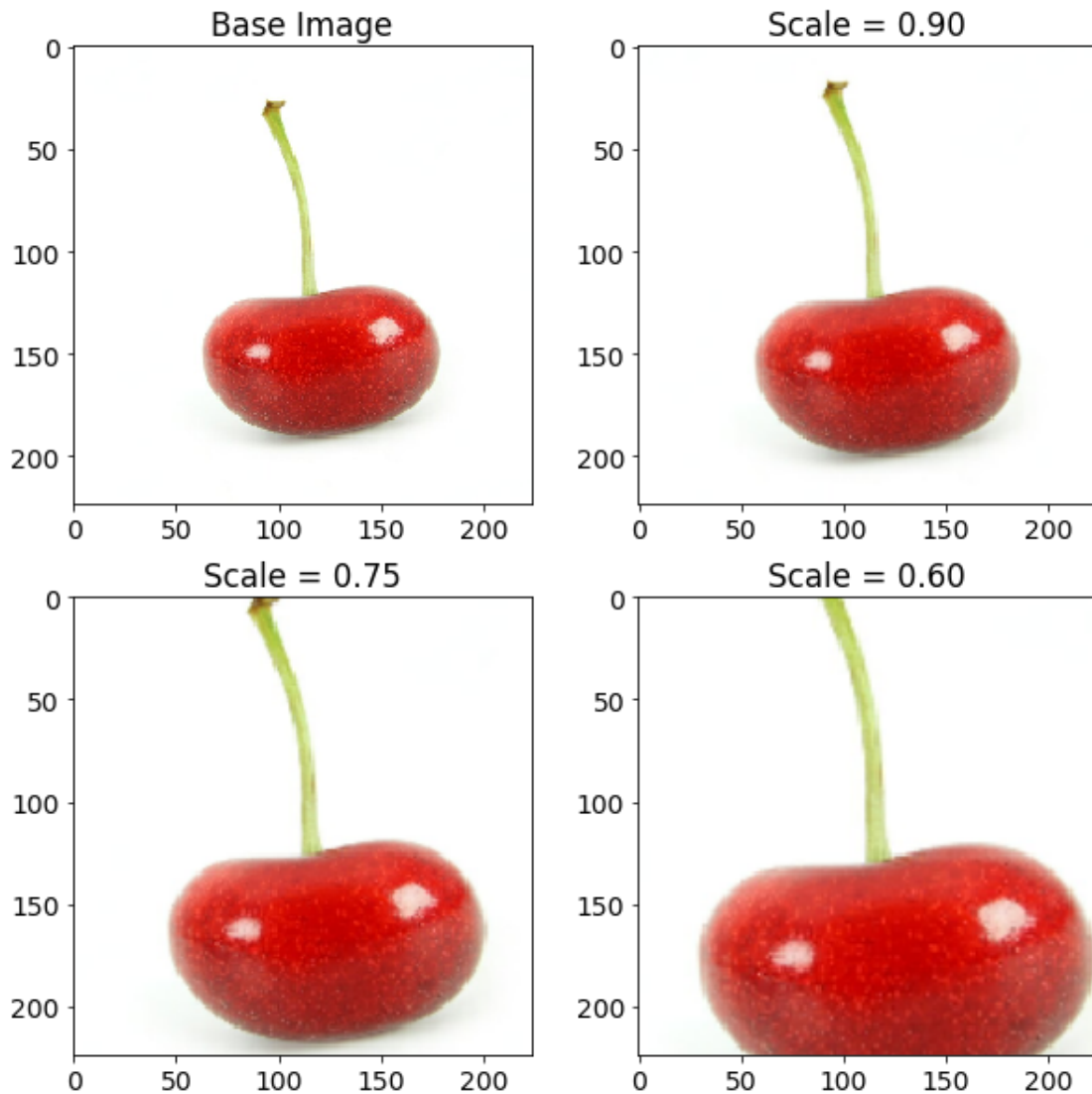


FIGURE 3.19 – Image avec l'effet de mise en echelle.

11.5 Injection de bruit

L'injection de bruit consiste à injecter une matrice de valeurs aléatoires généralement tirées d'une distribution gaussienne. L'injection de bruit est testée par Moreno-Barea et al. [57] sur neuf ensembles de données du référentiel UCI [58]. L'ajout de bruit aux images peut aider les CNN à acquérir des fonctionnalités plus robustes.

11.6 Recadrage (cropping) :

Le recadrage d'images peut être utilisé comme étape de traitement pratique pour les données d'image avec des dimensions mixtes de hauteur et de largeur en recadrant un patch central de chaque image.

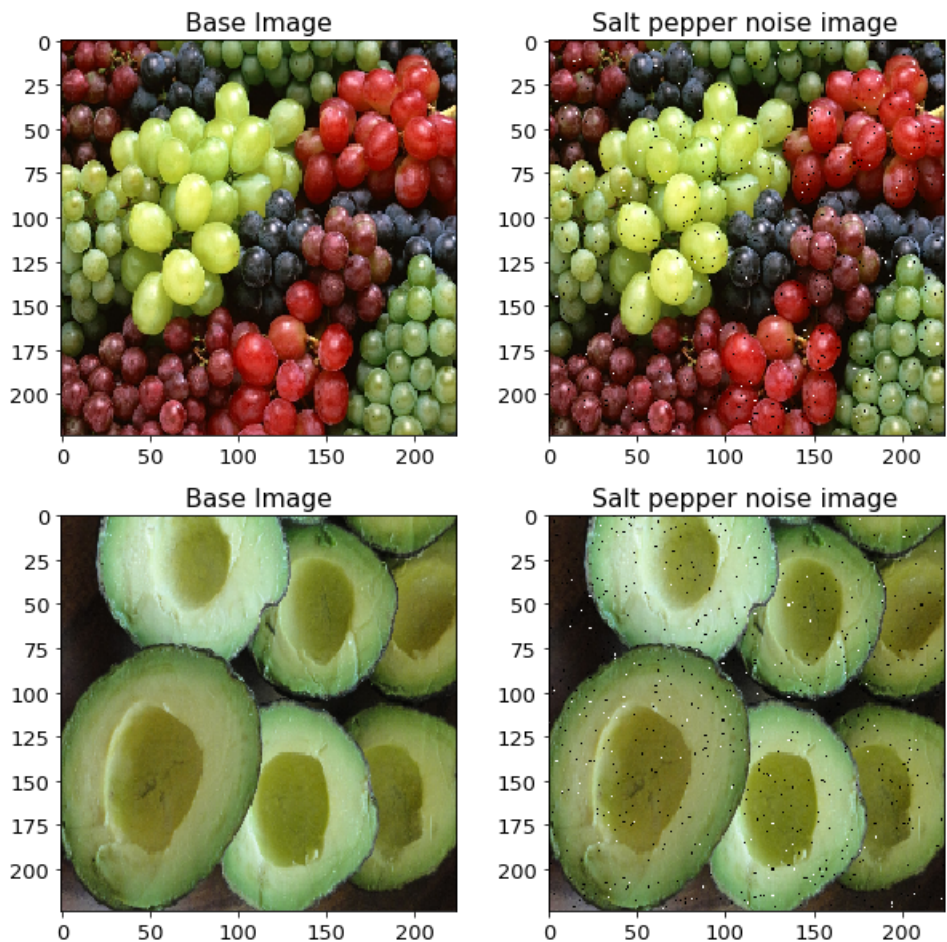


FIGURE 3.20 – En partant de la gauche, nous avons l'image d'origine, image avec Injection de bruit ajouté.

De plus, le recadrage aléatoire peut également être utilisé pour fournir un effet très similaire aux traductions. Le contraste entre le recadrage aléatoire et les traductions est que le recadrage réduira la taille de l'entrée telle que $(256,256) \rightarrow (224, 224)$, tandis que les traductions préservent les dimensions spatiales de l'image. Selon le seuil de réduction choisi pour le recadrage, il se peut qu'il ne s'agisse pas d'une transformation préservant l'étiquette.[43]

11.7 Condition d'éclairage :

La luminosité de l'image peut être augmentée soit en assombrissant les images de manière aléatoire, soit en éclaircissant les images, ou les deux.[43]



FIGURE 3.21 – Image originale(à gauche), une section carrée recadrée en haut (au centre), puis une section carrée recadrée en bas à droite.



FIGURE 3.22 – Images générées par un condition d'éclairage sur une image.

12 Métriques d'évaluation d'un RN :

L'évaluation de votre algorithme d'apprentissage automatique est une partie essentielle de tout projet. Votre modèle peut vous donner des résultats satisfaisants lorsqu'il est évalué à l'aide d'une métrique, par exemple, `precision_score`, mais peut donner de mauvais résultats lorsqu'il est évalué par rapport à d'autres métriques telles que `logarithmic_loss` ou toute autre métrique de ce type.[44]

La plupart du temps, nous utilisons la précision de classification pour mesurer les performances de notre modèle, mais cela ne suffit pas pour vraiment juger notre modèle.

12.1 Précision de classification (ACCURACY) :

La précision est une métrique pour évaluer les modèles de classification. De manière informelle, la précision est la fraction des prédictions que notre modèle a obtenues. Formellement, la précision a la définition suivante :

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

Pour la classification binaire, la précision peut également être calculée en termes de positifs et de négatifs comme suit :

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Où TP = vrais positifs, TN = vrais négatifs, FP = faux positifs et FN = faux négatifs.

Essayons de calculer la précision pour le modèle suivant qui a classé 100 tumeurs comme malignes (la classe positive) ou bénignes (la classe négative) :

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{1 + 90}{1 + 90 + 1 + 8} = 0.91$$

La précision est de 0,91, soit 91% (91 prédictions correctes sur 100 exemples au total). Cela signifie que notre classificateur de tumeurs fait un excellent travail pour identifier les tumeurs malignes, non ?[45]

12.2 Logarithmic loss : (perte logarithmique)

Logarithmic Loss ou Log Loss, fonctionne en pénalisant les fausses classifications. Cela fonctionne bien pour la classification multi-classes. Lorsque vous travaillez avec Log

<p>Vrai Positif (VP) : Realite : Malin Model ML predict : Malin Nombre de resultats VP :1</p>	<p>Faux Positive (FP) : Realite : Benin Model ML predict : Malin Nombre de resultats FP :1</p>
<p>Faux Negatif (FN) : Realite : Malin Model ML predict : Benin Nombre de resultats FN : 8</p>	<p>Vrai Negatif (VN) : Realite : Benin Model ML predict : Benin Nombre de resultats VN :90</p>

TABLE 3.1 – Illustration de l'exemple (classes positives et classes négatives).

Loss, le classificateur doit attribuer une probabilité à chaque classe pour tous les échantillons. Supposons qu'il y ait N échantillons appartenant à des classes M, puis la perte de journal est calculée comme ci-dessous :

$$LogarithmicLoss = \frac{-1}{N} \sum_{i=1}^N \sum_{j=1}^M y_{ij} * \log(p_{ij})$$

où,

$y_{i,j}$, indique si l'échantillon i appartient à la classe j ou non

$p_{i,j}$, indique la probabilité que l'échantillon i appartienne à la classe j

La perte de journal n'a pas de limite supérieure et elle existe sur la plage $[0, \infty)$. Une perte de journal plus proche de 0 indique une précision plus élevée, tandis que si la perte de journal est éloignée de 0, cela indique une précision inférieure. En général, minimiser la perte de journal donne une plus grande précision pour le classificateur.

12.3 Matrice de confusion :

La matrice de confusion comme son nom l'indique nous donne une matrice en sortie et décrit les performances complètes du modèle. Supposons que nous ayons un problème de classification binaire.

Nous avons quelques échantillons appartenant à deux classes : OUI ou NON. De plus, nous avons notre propre classificateur qui prédit une classe pour un échantillon d'entrée donné. En testant notre modèle sur 165 échantillons, nous obtenons le résultat suivant.

n=165	Predicted: NO	Predicted: YES
Actual: NO	50	10
Actual: YES	5	100

FIGURE 3.23 – Résultats de matrice de confusion.

Il y a 4 termes importants :

- **Vrais positifs** : Les cas dans lesquels nous avons prédit OUI et la sortie réelle était également OUI.
- **Vrais négatifs** : Les cas dans lesquels nous avons prédit NON et la sortie réelle était NON.
- **Faux positifs** : Les cas dans lesquels nous avons prédit OUI et la sortie réelle était NON.
- **Faux négatifs** : Les cas dans lesquels nous avons prédit NON et la sortie réelle était OUI.

12.4 F1 score :

Le score F1 est la moyenne harmonique entre la précision et le rappel. La plage du score F1 est $[0, 1]$. Il vous indique la précision de votre classificateur (combien d'instances il classe correctement), ainsi que sa robustesse (il ne manque pas un nombre significatif d'instances).

Haute précision mais rappel plus faible, vous donne un résultat extrêmement précis, mais il manque alors un grand nombre d'instances difficiles à classer. Plus le score F1 est élevé, meilleures sont les performances de notre modèle. Mathématiquement, il peut être

exprimé comme :

$$F1 = 2 * \frac{1}{\frac{1}{Precision} + \frac{1}{Rappel}}$$

F1 Score essaie de trouver l'équilibre entre précision et rappel.

12.4.1 Précision

Précision : C'est le nombre de résultats positifs corrects divisé par le nombre de résultats positifs prédits par le classificateur.

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives}$$

12.4.2 Rappel

Rappel : il s'agit du nombre de résultats positifs corrects divisé par le nombre de tous les échantillons pertinents (tous les échantillons qui auraient dû être identifiés comme positifs).

$$Precision = \frac{TruePositives}{TruePositives + FalseNegatives}$$

12.5 Mean Absolute Error (MAE) :

L'erreur absolue moyenne est la moyenne de la différence entre les valeurs d'origine et les valeurs prédites. Cela nous donne la mesure de l'écart entre les prévisions et la sortie réelle. Cependant, ils ne nous donnent aucune idée de la direction de l'erreur, c'est-à-dire si nous sous-prédisons les données ou sur-prédisons les données. Mathématiquement, il est représenté par :

$$MAE = \frac{1}{N} \sum_{j=1}^N |y_i - \hat{y}_j|$$

12.6 Mean Absolute Error (MAE) :

L'erreur quadratique moyenne (MSE) est assez similaire à l'erreur absolue moyenne, la seule différence étant que MSE prend la moyenne du carré de la différence entre les valeurs d'origine et les valeurs prédites.

L'avantage de MSE est qu'il est plus facile de calculer le gradient, alors que l'erreur absolue moyenne nécessite des outils de programmation linéaire compliqués pour calculer le gradient. Au fur et à mesure que nous considérons l'erreur au carré, l'effet des erreurs plus importantes devient plus prononcé que celui des erreurs plus petites.[44]

$$MSE = \frac{1}{N} \sum_{j=1}^N (y_i - \hat{y}_j)^2|$$

13 Conclusion :

Dans ce chapitre on a parlé des réseaux de neurones convolutifs avec leurs différentes couches et le rôle de chaque couche . Dans le chapitre suivant nous allons aborder les techniques et les outils de la classification automatique des images numériques en parlant de l'intelligence artificielle.

CHAPITRE

4

IMPLÉMENTATION ,RÉSULTATS ET
ÉVALUATION DU TRAVAIL

1 Introduction

Dans ce chapitre, nous allons présenter les outils qu'on a utilisés, les bases d'images de l'entraînement, de validation et de test .Ensuite on va présenter nos architectures des deux modèles. A la fin on discutera et évaluera les résultats de ces modèles.

2 Ressources matérielles et logicielles

Dans cette section nous allons mentionner et décrire les outils et l'environnement utilisés dans notre implémentation :

2.1 Ressources logicielles

2.1.1 Python :

Nous avons utilisé le langage de programmation Python la version 3.7(stable).

Python est un langage de programmation open source créé par le programmeur Guido van Rossum en 1991.

Il s'agit d'un langage de programmation interprété, qui ne nécessite donc pas d'être compilé pour fonctionner. Un programme "interpréteur" permet d'exécuter le code Python sur n'importe quel ordinateur. Ceci permet de voir rapidement les résultats d'un changement dans le code. En revanche, ceci rend ce langage plus lent qu'un langage compilé comme le C.

ce langage est aussi utilisé dans le domaine de la science des données et du Machine Learning. Avec l'essor de l'analyse de données dans toutes les industries, c'est d'ailleurs devenu l'un de ses principaux cas d'usage.[47]

2.1.2 TensorFlow :

TensorFlow est une bibliothèque de logiciels open source pour le calcul numérique utilisant des graphiques de flux de données. Il a été initialement développé par l'équipe Google Brain en 2011 au sein de l'organisation de recherche Machine Intelligence de Google pour machine learning et la recherche sur les réseaux de neurones profonds.

TensorFlow est multiplateforme. Il fonctionne sur presque tout : les GPU et CPU - y compris les plates-formes mobiles et embarquées - et même les unités de traitement des tenseurs (TPU), qui sont du matériel spécialisé pour faire des calculs de tenseurs, mais ils ne sont pas encore largement disponibles.[48]

2.1.3 Keras :

Keras est une bibliothèque de réseaux neuronaux open source écrite en Python. Il est capable de fonctionner sur TensorFlow, Microsoft Cognitive Toolkit, R, Theano ou PlaidML. Conçu pour permettre une expérimentation rapide avec des réseaux de neurones profonds, il se concentre sur être convivial, modulaire et extensible. Il a été développé dans le cadre de l'effort de recherche du projet ONEIROS (Open-end Neuro-Electronic Intelligent Robot Operating System), et son principal auteur et mainteneur est François Chollet, un ingénieur de Google. Chollet est également l'auteur du modèle de réseau neuronal profond Xception. [49]

2.2 Ressources matérielles

La configuration du matériel utilisé dans notre implémentation est :

- Un PC portable HP i3 CPU 1.7 GHZ
- Carte graphique Intel(R) HD Graphics Family Version DirectX installée : 11.0
Fréquence du GPU : 600 MHz
- RAM de taille 4 GO
- Disque dur HDD de taille 500 Go
- Système d'exploitation Windows 10 Famille
- Ainsi que Google Colab .

2.2.1 Google Colab :

Colaboratory, ou «Colab» en abrégé, est un produit de Google Research. Colab permet à quiconque d'écrire et d'exécuter du code Python arbitraire via le navigateur, et est particulièrement bien adapté à l'apprentissage automatique, à l'analyse de données et à l'éducation. Plus techniquement, Colab est un service de notebook Jupyter hébergé qui ne nécessite aucune configuration à utiliser, tout en offrant un accès gratuit aux ressources informatiques, y compris les GPU.[50]

2.2.2 Kaggle :

Kaggle est une plateforme web organisant des compétitions en science des données. Sur cette plateforme, les entreprises proposent des problèmes en science des données et offrent un prix aux datalogistes obtenant les meilleures performances. L'entreprise a été créée en 2010 par Anthony Goldbloom C'est une filiale Google qui offre aux utilisateurs une interface pour exécuter leurs projets via le cloud .

3 Les bases d'images :

Dans le but de tester le taux de réussite , nous avons utiliseons kaggle dataset "Diabetic Retinopathy Kaggle competition dataset " comme une ensemble de donnée "dataset" .

3.1 APTOS 2019 Blindness Detection :

Grand ensemble d'images de la rétine prises à l'aide de la photographie du fond d'œil dans diverses conditions d'imagerie, contient 5590 images couleur du fond d'oeil(3662 images de train at 1928 images de test). Un clinicien a évalué chaque image en fonction de la gravité de la rétinopathie diabétique sur une échelle de 0 à 4 :

- 0 - Pas de DR
- 1 - Doux
- 2 - Modéré
- 3 - Sévère
- 4 - DR proliférative

Bases d'images	train	test	Total
nombre des images	3662	1928	5590

TABLE 4.1 – Nombre dimages APTOS 2019 Blindness Detection

3.2 Diabetic Retinopathy Kaggle competition dataset :

DRKCD est un grand ensemble d'images rétinienne à haute résolution prises dans diverses conditions d'imagerie, contient 35126 images couleur du fond d'oeil. Un champ gauche et droit est fourni pour chaque sujet. Les images sont étiquetées avec un identifiant de sujet ainsi que soit à gauche ou à droite (par exemple, 1_left.jpeg est l'oeil gauche du patient id 1).

Un clinicien a évalué la présence de rétinopathie diabétique dans chaque image sur une échelle de 0 à 4, selon l'échelle suivante :

- 0 - Pas de DR
- 1 - Doux
- 2 - Modéré

- 3 - Sévère
- 4 - DR proliférative[51]

3.2.1 Base d'image :

Dans le tableau suivant, nous montrerons notre base d'images :

Bases d'images	Pas de DR	Doux	Modéré	Sévère	DR proliférative	Total
apprentissage	15492	1464	3174	524	426	21080
validation	5164	487	1057	174	141	7023
test	5164	487	1057	174	141	7023
Total	25820	2438	5288	872	708	35126

TABLE 4.2 – Nombre d'images dans chaque base.

4 Traitements préliminaires :

4.1 Filtrage :

Les données que nous avons utilisées ont été traitées à l'aide d'une ensemble de techniques de traitement d'images ,tel que la mise au niveau du gris , le réglage de la luminosité , le filtrage et d'autres .

4.2 Redimensionnement :

Le modèle sélectionné (dans la section suivante) impose la contrainte selon laquelle chaque image doit avoir la même taille et le même rapport d'aspect. La taille sélectionnée des images était de 205 x 205 pixels. Pour cela, un prétraitement était effectué aux images en utilisant la bibliothèque OpenCV.

5 L'apprentissage par transfert(Transfer Learning)

L'apprentissage par transfert(transfer learning en anglais) est l'un des champs de recherche de l'apprentissage automatique qui vise à transférer des connaissances d'une ou plusieurs tâches sources vers une ou plusieurs tâches cibles. Il peut être vu comme la capacité d'un système à reconnaître et appliquer des connaissances et des compétences, apprises à partir de tâches antérieures, sur de nouvelles tâches ou domaines partageant des similitudes.[55]

En vision par ordinateur, l'apprentissage par transfert s'exprime généralement à travers l'utilisation de modèles pré-entraînés(pre-trained).

5.1 modèle pré-entraîné(pre-trained models) :

Un modèle pré-entraîné est un modèle qui a été formé sur un grand ensemble de données de référence pour résoudre un problème similaire à celui que nous voulons résoudre. En conséquence, en raison du coût de calcul de la formation de tels modèles, il est courant d'importer et d'utiliser des modèles à partir de la littérature publiée (par exemple VGG, Inception, ResNet).

6 Architecture de notre réseau :

Au cours de nos expérimentations, nous avons créé deux modèles modèle 1 et modèle 2 avec une différence entre les couches . Dans ce qui suit on présente l'architecture des deux modèles :

6.1 EfficientNet

EfficientNet-B0 est le réseau de base développé par AutoML MNAS, tandis que Efficient-B1 à B7 sont obtenus en augmentant le réseau de base. En particulier, EfficientNet-B7 atteint une nouvelle précision de pointe de 84,4% dans le top 1 / 97,1% dans le top 5, tout en étant 8,4 fois plus petit que le meilleur CNN existant.

6.2 Architecture EfficientNet

L'efficacité de la mise à l'échelle du modèle dépend également fortement du réseau de base. Ainsi, pour améliorer encore les performances, Ils ont également développé un nouveau réseau de base en effectuant une recherche d'architecture neuronale à l'aide du framework AutoML MNAS, qui optimise à la fois la précision et l'efficacité (FLOPS). L'architecture résultante utilise la convolution de goulot d'étranglement inversé mobile (MBConv), similaire à MobileNetV2 et MnasNet, mais est légèrement plus grande en raison d'un budget FLOP accru. Nous étendons ensuite le réseau de base pour obtenir une famille de modèles, appelée EfficientNets.[54]

6.3 Architecture du modèle 1 :

Dans le premier modèle, nous avons utilisé un modèle pré-entraîné (EfficientNet) Pour obtenir le meilleur résultat possible.

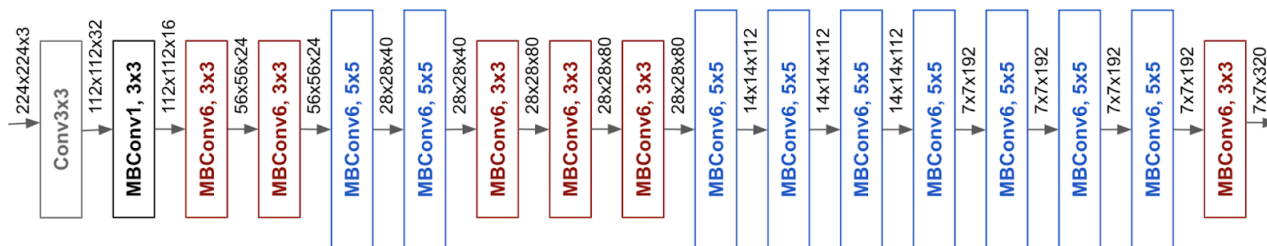


FIGURE 4.1 – Architecture EfficientNet.

Nous avons préparé ce modèle pour nous conformer à notre dataset par supprimé la couche d’entrée pour donner notre data et la dernière couche(couche de fully connected) pour donner notre production.

Le premier modèle que nous présentons dans la Figure 4.2 est composé de tous les couche de EfficientNet-b3 et en plus une couches de average_pooling et une couche de fully connected De notre problème. Ce modèle était entraîné avec la base d’images AP-TOS 2019 Blindness Detection. Nous avons utilisé une couches dropout avant la couche de fully connected (dernière couche du réseau) .L’image en entrée est de taille 108*108 en niveau de gris (channel=1).

La dernière couche utilisée la fonction d’activation softmax qui permet de calculer la distribution de probabilité des 5 classes (Pas de DR,Doux,Modéré,Sévère,DR proliférative).Et nous avons optimisé par Adam .

```

=====
Layer (type)                Output Shape                Param #
=====
efficientnet-b3 (Model)      (None, 4, 4, 1536)         10783528
global_average_pooling2d_1  (None, 1536)                0
dropout_1 (Dropout)         (None, 1536)                0
dense_1 (Dense)              (None, 5)                   7685
=====
Total params: 10,791,213
Trainable params: 10,703,917
Non-trainable params: 87,296
=====
None
    
```

FIGURE 4.2 – Configuration du model 1.

6.4 Architecture du modèle 2 :

Le deuxième modèle que nous présentons dans la Figure 4.3 est composé de quatre couches de convolution et quatre couches de maxpooling et deux couche de fully connected. Ce modèle était entraîné avec la base d'images APTOS 2019 Blindness Detection. Nous avons utilisé deux couches dropout avant la couche de fully connected (dernière couche du réseau) .

L'image en entrée est de taille 108×108 en couleur (channel=3) ,l'image passe d'abord à la première couche de convolution.

Cette couche est composée de 32 filtres de taille 108×108 , la fonction d'activation ReLU est utilisé, cette fonction d'activation force les neurones à retourner des valeurs positives, après cette convolution 16 features maps seront créés de taille 3×3 .

Le Maxpooling est appliqué après pour réduire la taille de l'image. À la sortie de cette couche, nous aurons 16 feature maps de taille 54×54 .

Ensuite, les feature maps qui sont obtenus ils sont données en entrée de la deuxième couche de convolution qui est composée aussi de 32 filtres de taille 54×54 . La fonction d'activation ReLU est appliquée sur cette couche. Le Maxpooling est appliqué après pour réduire la taille de l'image. À la sortie de cette couche, nous aurons 64 feature maps de taille 27×27 . On répète la même chose avec les couches de convolutions trois et quatre, la fonction d'activation ReLU est appliquée toujours sur chaque convolution. Une couche de Maxpooling est appliqué après les couches de convolution trois et quatre.

La dernière couche utilisée la fonction d'activation softmax qui permet de calculer la distribution de probabilité des 5 classes (Pas de DR,Doux,Modéré,Sévère,DR proliférative).Et nous avons optimisé par Adam .

Layer (type)	Output Shape	Param #
conv2d_105 (Conv2D)	(None, 108, 108, 16)	448
max_pooling2d_1 (MaxPooling2D)	(None, 54, 54, 16)	0
conv2d_106 (Conv2D)	(None, 54, 54, 64)	9280
max_pooling2d_2 (MaxPooling2D)	(None, 27, 27, 64)	0
conv2d_107 (Conv2D)	(None, 27, 27, 215)	124055
max_pooling2d_3 (MaxPooling2D)	(None, 13, 13, 215)	0
conv2d_108 (Conv2D)	(None, 13, 13, 512)	991232
max_pooling2d_4 (MaxPooling2D)	(None, 6, 6, 512)	0
flatten_1 (Flatten)	(None, 18432)	0
dense_1 (Dense)	(None, 512)	9437696
dropout_1 (Dropout)	(None, 512)	0
dense_2 (Dense)	(None, 512)	262656
dropout_2 (Dropout)	(None, 512)	0
dense_3 (Dense)	(None, 5)	2565
Total params: 10,827,932		
Trainable params: 10,827,932		
Non-trainable params: 0		

FIGURE 4.3 – Configuration du model 2.

7 Résultats obtenus et discussion :

7.1 L'évaluation de notre réseau :

Pour évaluer notre réseau nous avons choisi la métrique de l'accuracy, vu qu'elle est la plus efficace et la plus utilisée. Nos résultats sont illustrés donc en terme précision-perte pour chaque modèle .

7.2 Résultats obtenu pour modèle 1 :

Le modèle était entraîné sur 50 époques :



FIGURE 4.4 – Graphes de précision (accuracy) à gauche – et de la perte (loss) à droite résultants du modèle 1.

D’après la figure ci-dessus la précision de l’apprentissage augmente avec le nombre d’époque, jusqu’à elle est stable ceci reflète qu’à chaque époque le modèle apprend plus d’information, par contre validation dont les valeurs atteignent parfois 0.4.

De même l’erreur d’apprentissage est diminuée, d’autre part la validation augmente avec le nombre d’époque. Au fur et à mesure que l’apprentissage progresse, l’erreur de généralisation peut diminuer au minimum, puis augmenter de nouveau à mesure que le réseau s’adapte aux particularités des données de l’apprentissage.

Le problème du sur-apprentissage est que plus le modèle est spécialisé dans le training data, moins il est capable de généraliser à de nouvelles données, ce qui entraîne une augmentation des erreurs de généralisation. Cette augmentation de l’erreur de généralisation peut être mesurée par la performance du modèle sur l’ensemble de données de validation.

Ce que nous remarquons, c’est que le modèle comprend bien ce que signifie reconnaître facilement toutes les données. Parce que le nombre de données est assez moyen. Nous avons donc augmenté le nombre d’images dans la base de données en utilisant data agumentation en appliquant un ensemble de techniques aléatoirement .

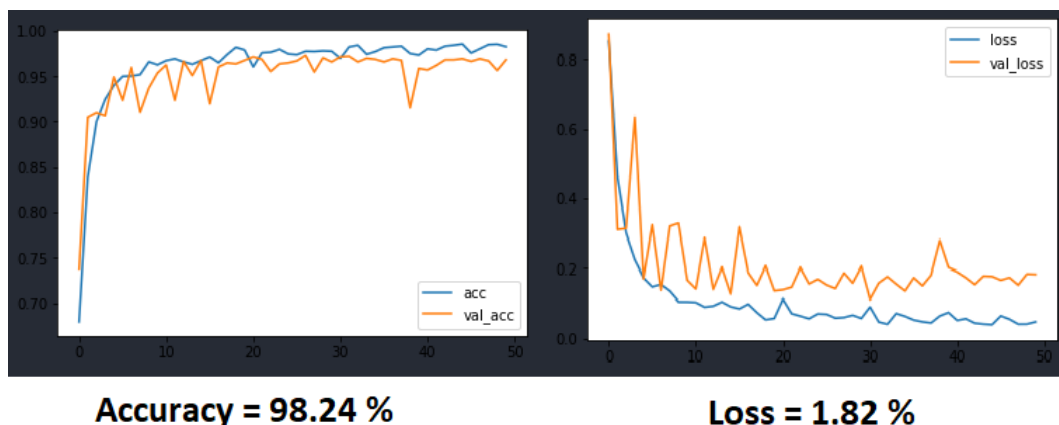


FIGURE 4.5 – Graphes de précision (accuracy) à gauche – et de la perte (loss) à droite résultants du modèle 1(après augmentation de données).

7.3 Résultats obtenu pour modèle 2 :

Le modèle était entraîné sur 10 epochs -16 steps par epoch .

Le graph de l'accuracy , on observe que le taux de validation est supérieur à celui de training, ce qui signifie que le réseau n'a pas bien retenu le dataset , ou bien il a su valider les cas normaux . Le graph de loss est acceptable par rapport celui de loss , observons que le training et la validation se diminuent après chaque epoch .



FIGURE 4.6 – Graphes de précision (accuracy) à gauche – et de la perte (loss) à droite résultants du modèle 2(100 epochs).

7.4 Comparaison des deux modèles :

Comparaison des deux modèles :

	Modèle1	Modèle2
Nombre des Époques	50	50-100
Accuracy	98.24%	90.19%
Loss	1.82%	4.42

TABLE 4.3 – Comparaison des résultats des deux modèles.

Dans le tableau 4.3 on va à faire la comparaisent entre les deux modèles proposés. D'après ce tableau nous remarquons que le modèle 1 nous donne des meilleurs résultats par apport au modèle2.

7.5 Résultat final :

Comme résultat final le meilleur model qui donne des bons résultats c'est le model 1 qui été le modèle pré-formé (EfficientNet) avec dropout de 0.5, nombre des époques 50 -100 step par epoch et des données augmentés.

7.6 L'explication de Code source :

1. `keras.layers.Conv2D(filters=96, kernel_size=(11,11), strides=(4,4), activation='relu', input_shape=(118,118,3)),`

Cette commande permet de créer 96 features maps on utilisant un filtre de taille 11 par 11 pixels et avec un mode de bordure égal à la taille de l'image précédente et une fonction d'activation de type RELU.

2. `model.add(layers.MaxPooling2D((2, 2)))`

Cette ligne de commande permet de réduire la taille de l'image, La méthode Max Pooling est utilisée et la taille de l'image sera diviséesur 2.

3. `model.add(layers.Flatten())`

Cette commande permet de créer un seul vecteur 1D puis connecter avec la première couche cachée pour commencer la classification.

4. `keras.layers.Dense(4096, activation='relu')`,

Cette commande permet de créer une couche cachée avec une taille de 1024 neurones, la fonction RELU est utilisé comme fonction d'activation.

5. `keras.layers.Dropout(0.5)`,

Pour ne pas tomber dans le problème de sur apprentissage il faut utiliser dropout elle est très efficace pour les réseaux de neurones pour le régulariser.

6. `keras.layers.Dense(5, activation='softmax')`

Cette commande permet de créer une couche de sortie composée de 5 neurones (nombre de classes) la fonction softmax est utilisé pour calculer la probabilité de chaque classe.

7. `model.compile(loss='categorical_crossentropy',optimizer='adam',metrics=['accuracy'])`

Cette commande permet de compiler notre modèle, elle prend deux paramètres la fonction loss et Opimizer. On a choisi la fonction `categorical_crossentropy` comme fonction loss et `adam` comme optimizer .

8. `model.fit(x_train, y_train, batch_size=batch_size, nb_epoch=num_epochs, verbose=1, validation_split=0.2)`

Cette commande permet de lancer l'apprentissage puis la validation .

9. `model.evaluate(x_test, y_test, verbose=0) print('Test score :', score[0]) print('Test accuracy :', score[1])`

Cette commande permet d'évaluer notre modèle sur la base de test.

8 Conclusion :

Nous avons présenté dans ce chapitre une approche de classification basée sur les réseaux de neurones, pour cela on a utilisé deux modèles avec différents architectures et

on a montré les différents résultats. La comparaison des résultats trouvés a montré que le model EfficientNet (model 1) a donné les meilleurs résultats de précision et de perte .

CONCLUSION GÉNÉRALE

Conclusion générale

Dans ce projet , nous avons proposé une approche de classification d'image avec deep learning (plus précisément réseaux des neurones convolutifs) pour classier les images rétinienne dans le cas de la rétinopathie diabétique . Tout d'abord nous avons présenté domaine médical ,dont on a identifié l'anatomie de l'oeil et la maladie de la rétinopathie diabétique , ensuite on a passé au traitement d'images numériques qui sont les entrées nécessaires pour tester notre classifieur .

La troisième partie de notre recherche est couvrir l'état de l'art dans le domaine du deep learning et les projets actuels dans la classification d'images médicales .Puis on a focalisé sur le type de réseaux des neurones qui est dédié à la classification d'images .

Le dernier pas était l'implémentation et la mise en oeuvre de notre classifieur et la discussion des résultats obtenus en comparant deux modèles .

Durant la réalisation de ce projet nous avons essayé, d'augmenter l'efficacité et la performance du model de notre classifieur , dans le but développer une nouvelle approche de classification des images médicale qui sera plus performante que les approches existantes.

Par ailleurs ce projet de fin d'étude était pour nous une occasion pour acquérir des atouts et des connaissances techniques sur l'analyse de données et le domaine du deep learning .

Finalement, nous avons obtenu un classifieur convenable qui a donné un bon résultat . Mais, cela ne c'est pas fait sans peine car nous avons été obligés de modifier plusieurs fois les mêmes choses pour atteindre les résultats souhaités .Surtout que nous avons fait face à un problème avec le matériel vu que le nombre immense de données nécessite d'être traités sur une machine puissante (GPU, RAM 25GB etc) . ce problème a été réglé en entraînant notre réseau sur les plateformes en ligne tel que Google Colab .

L'apprentissage profond nécessite une quantité considérable de données de formation, car la précision de la classification de l'apprentissage profond dépend en grande partie de la qualité et de la taille du jeu de données.

Ce projet a été très bénéfique pour nous en tant qu'une expérience personnelle et professionnelle non seulement en informatique mais aussi dans le domaine médical . Notre mission nous a beaucoup intéressés et nous a permis de réaliser un classifieur de réseau des neurones convolutif (CNN) .

Perspectives :

- On propose aussi d'utiliser une base d'images plus large pour avoir les bons résultats , une bonne annotation (ou un étiquetage) de données est un facteur d'augmentation de la précision .

BIBLIOGRAPHIE

- [1] Consulté le 06/08/2020
<https://www.sante-sur-le-net.com/maladies/ophtalmologie/retinopathie-diabetique-sy>
- [2] Consulté le 06/08/2020
<https://www.futura-sciences.com/sante/definitions/corps-humain-%C5%93il-14131/>
- [3] Consulté le 06/08/2020
<https://www.gatinel.com/recherche-formation/la-cornee/>
- [4] Consulté le 06/08/2020
<https://www.gatinel.com/recherche-formation/le-cristallin>
- [5] Consulté le 06/08/2020
<https://www.futura-sciences.com/planete/definitions/botanique-iris-13342/>
- [6] Consulté le 06/08/2020
<https://www.guide-vue.fr/glossaire/humeur-aqueuse>
- [7] Consulté le 06/08/2020
<https://www.teteamodeler.com/sante/vue/vue5.asp>
- [8] Consulté le 06/08/2020
<https://www.futura-sciences.com/sante/definitions/corps-humain-retine-4322/>
- [9] Consulté le 06/08/2020
<https://www.guide-vue.fr/glossaire/macula>
- [10] Consulté le 06/08/2020
<https://www.medicinenet.com/script/main/art.asp?articlekey=4653>
- [11] Consulté le 06/08/2020
http://campus.cerimes.fr/ophtalmologie/enseignement/ophtalmo16/site/html/3_33_1.htm

BIBLIOGRAPHIE

- [12] Consulté le 06/08/2020
<https://www.sciencesetavenir.fr/sante/ophtalmo/retinopathie-diabetique-definition-s#origine>.
- [13] Consulté le 06/08/2020
http://campus.cerimes.fr/ophtalmologie/enseignement/ophtalmo16/site/html/3_33_1.htm
- [14] Ophtalmologie paris ,Dr Assouline Michaël chirurgien ophtalmologiste . *Les Examens Complémentaires en Ophtalmologie// Angiographie rétinienne*,. 2003.
- [15] Consulté le 06/08/2020
<http://www.institutmacula.com/fr/procedimiento/retinografia/> ,
Consulté le 30/05/2019.
- [16] M2 IRIV. *Outils fondamentaux en traitements d'images*..2019–2020.
- [17] Définition image numérique,
<https://www.schoolmouv.fr/definitions/image-numerique/definition>Consulté le 10/10/2020
- [18] Eugene P. *What Is Image Acquisition in Image Processing?*.Last Modified Date : September 17, 2020.
- [19] Raphaël Isdant. *Traitement numérique de l'image*.2009.
- [20] Vangie Beal. *image dimensions*.Consulté le 10/10/2020.
- [21] Julia Kuzmenko McKim, RA Team . *Understanding Image Noise*.Consulté le 10/10/2020.
- [22] Nickson Joram. *Intensity Histogram*.Jul 28, 2019.Consulté le 10/10/2020.
- [23] CONTRASTE : PRISE DE VUES ET DÉVELOPPEMENT,
<https://www.infotographiste.fr/contraste-prise-de-vues-developpement/>Consulté le 10/10/2020
- [24] Luminance,
<https://www.futura-sciences.com/sciences/definitions/physique-luminance-2510/>Consulté le 10/10/2020
- [25] Website Image Compression : Is Bigger Really Better?,
<https://marketwithfirefly.com/image-compression/>Consulté le 10/10/2020
- [26] Image Representation in Computer Graphics,
<https://www.tutorialandexample.com/image-representation-in-computer-graphics>Consulté le 10/10/2020
- [27] ISN – Informatique et Sciences du Numérique. *TP : CODAGE DES IMAGES NUMÉRIQUES*.Jul 28, 2019.Consulté le 10/10/2020.
- [28] CHIKH Mohammed Tahar. *Mémoire de fin d'études pour l'obtention du diplôme de Master en Informatique, Amélioration des images par un modèle de réseau de neurones (Comparaison avec les filtres de base)*.Présenté le 28 Septembre 2011 devant la commission composée de MM.2010/2011.

BIBLIOGRAPHIE

- [29] Classification et analyse des images,
<https://www.rncan.gc.ca/cartes-outils-publications/imagerie-satellitaire-photos-aer>
le 10/10/2020
- [30] image processing,
https://www.tutorialspoint.com/dip/applications_and_usage.htm Consulté le
10/10/2020
- [31] Mehdi MOUADIL. *Introduction au Deep Learning : les réseaux de neurones*. 23 février 2018.
- [32] NICOLAS GONTHIER, NICOLAS GONTHIER. *Introduction au Deep Learning*. POSTÉ PAR OLIVIER BONFAIT, LE 1 AVRIL 2019.
- [33] Céline Deluzarche. *Deep Learning*. POSTÉ PAR OLIVIER BONFAIT, LE 1 AVRIL 2019.
- [34] Intelligence Artificielle, Machine Learning, Deep Learning : une histoire de poupées russes,
<https://www.oracle.com/fr/artificial-intelligence/deep-learning-machine-learning-in>
le 08/10/2020
- [35] Neural Network,
<https://deeptai.org/machine-learning-glossary-and-terms/neural-network>
- [36] Types of Neural Networks (and what each one does!) Explained,
<https://towardsdatascience.com/types-of-neural-network-and-what-each-one-does-expl>
- [37] convolutional neural network,
<https://searchenterpriseai.techtarget.com/definition/convolutional-neural-network>
- [38] A. Krizhevsky, I. Sutskever et G. E. Hinton. *ImageNet Classification with Deep Convolutional Neural Networks* », *Advances in neural Processing Systems de traitement*. 2012.
- [39] K. Wisaeng, N. Hiransakolwong, and E. Pothiruk, *Automatic detection of exudates in retinal images based on threshold moving average models*. *Biophysics*, vol. 60, no. 2, pp. 288–297, 2015.
- [40] Focus : Le Réseau de Neurones Convolutifs,
<https://penseeartificielle.fr/focus-reseau-neurones-convolutifs/> 11 janvier 2019.
- [41] HIJAZI, Samer, KUMAR, Rishi, et ROWEN, Chris. *Using convolutional neural networks for image recognition*. *Cadence Design Systems Inc.*. San Jose, CA, USA, 2015.
- [42] A Beginner's Guide To Understanding Convolutional Neural Networks Part 2,
<https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-N>
July 29, 2016.

BIBLIOGRAPHIE

- [43] Boughaba et Boukhris , *L'apprentissage profond (Deep Learning) pour la classification et la recherche d'images par le contenu*. Mémoire de pfe Master-Université de Ouargla, 2016-2017.
- [44] A survey on Image Data Augmentation for Deep Learning,
<https://link.springer.com/article/10.1186/s40537-019-0197-0> Consulté le 07/08/2020
- [45] Metrics to Evaluate your Machine Learning Algorithm,
<https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm>.
le 09/09/2020
- [46] Classification : Accuracy,
<https://developers.google.com/machine-learning/crash-course/classification/accuracy>
le 09/09/2020
- [47] python,
: <https://www.lebigdata.fr/python-langage-definition> Consulté le 09/07/2020
- [48] tensorflow,
<https://opensource.com/article/17/11/intro-tensorflow> Consulté le 09/07/2020
- [49] Keras,
https://en.wikipedia.org/wiki/Keras#cite_note-10 Consulté le 16/09/2020
- [50] What is Colaboratory ?,
<https://research.google.com/colaboratory/faq.html> Consulté le 16/09/2020
- [51] kaggle dataset,
<https://www.kaggle.com/c/diabetic-retinopathy-detection/data> Consulté le 16/09/2020
- [52] Durlu, Y. K., Körođlu, Ç., & Tolun, A. (2014). *Novel recessive cone-rod dystrophy caused by POC1B mutation. JAMA ophthalmology*..132(10), 1185-1191.
- [53] Dave Gershgorn (18 June 2018). *The inside story of how AI got good enough to dominate Silicon Valley*. Quartz. Retrieved 5 October 2018.
- [54] Krizhevsky, Alex; Sutskever, Ilya; Hinton, Geoffrey. *ImageNet classification with deep convolutional neural networks*".E. (2017-05-24).
- [55] Sinno Jialin Pan et Qiang Yang, *A Survey on Transfer Learning*, IEEE Transactions on Knowledge and Data Engineering, 10/2010,p. 1345-1359.