



République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieure et de la Recherche
Scientifique



Université KasdiMerbah Ouargla

Faculté des nouvelles technologies de l'information et de la communication
Département d'informatique et technologie de l'information

Mémoire de Master en Informatique

Option : Administration et sécurité des réseaux

Thème

Détection du COVID-19 à l'aide des images médicales

X-ray

Présenté par

Aicha FAKROUN, Souhila MAACHE

Devant le jury composé de :

Président	ICHI Saleh M.A.A	Université de Ouargla
Rapporteur	KAHLESSENANE FaresM.A.A	Université de Ouargla
Examineur	BOUANAN Khadra M.C.A	Université de Ouargla

Promotion 2021/2022

Remerciements

Merci beaucoup à Dieu qui nous a donné la force, la volonté et la patience pour faire ce travail.

Nos remerciements et notre gratitude vont tout d'abord au Professeur KhelsananFares, qui nous a encadrés et guidés durant nos études et la réalisation de ce projet avec un intérêt constant, et qui l'a préféré dans nos travaux.

Nous exprimons également notre respect et notre gratitude aux membres du jury qui nous ont fait l'honneur de juger ce travail et qui, par leur disponibilité, leurs retours et leurs témoignages, nous ont permis d'enrichir notre travail. Nos remerciements et notre respect à tous nos professeurs durant les années scolaires.

Enfin, les mots sont trop faibles pour exprimer notre gratitude à nos chers parents et frères pour leur aide et leurs encouragements durant toutes ces années scolaires.

Sans oublier bien sûr de remercier profondément tous ceux qui ontContribué directement ou indirectement à la réalisation de ce travail.

Dédicaces

Dieu soit loué d'abord - puis pour les efforts de mes proches qui m'ont beaucoup aidé par leurs conseils et leurs efforts. Je dédie ce travail :

À ma chère mère,

À mon père.

À tous mes frères et sœurs.

À tous mes amis proches.

Au professeur superviseur, Khalsanan Faris.

À tous mes collègues.

Résumé

COVID-19, également connu sous le nom de Coronavirus, a partiellement paralysé le monde de presque toutes les manières, le domaine de la vision par ordinateur, de l'apprentissage en profondeur et de la machine... a montré une grande amélioration et est maintenant appliqué dans de nombreux domaines tels que la sécurité et la médecine. La propagation rapide de cette épidémie nous a incités à remuer nos esprits, à trouver des solutions et à profiter de la technologie disponible pour réduire l'impact de l'épidémie. À l'aide d'une radiographie pulmonaire, il fournit de la profondeur à un réseau neuronal convolutif pour aider les cliniciens à prendre les bonnes décisions de diagnostic. Et après plusieurs tests et différentes façons d'utiliser les modèles CNN, nous pouvons dire que nous avons atteint le résultat souhaité à 98.66%.

Mots clés : COVID-19, la vision par ordinateur, apprentissage en profondeur, la machine, un réseau neuronal convolutif (CNN), radiographie pulmonaire.

Abstract

COVID-19, also known as Coronavirus, has partially paralyzed the world in almost every way, the field of computer vision, deep learning and machine... has shown great improvement and is now applied in many fields such as security and medicine. The rapid spread of this epidemic has prompted us to stir our minds, find solutions and take advantage of available technology to reduce the impact of the epidemic. Using a chest x-ray, it provides depth to a convolutional neural network to help clinicians make the right diagnostic decisions. And after several tests and different ways of using CNN models, we can say that we have achieved the desired result at 98.66%.

Key words: COVID-19, computer vision, deep learning, machine, convolutional neural network (CNN), chest X-ray.

المخلص:

تسبب COVID-19 المعروف ايضا باسم Coronavirus في شل العالم جزئيا بكل الطرق تقريبا فقد اظهر مجال رؤية الكمبيوتر و التعلم العميق و التعلم الالي...تحسنا كبيرا و يتم تطبيقه الان في العديد من المجالات مثل الامن و الطب دفعنا الانتشار السريع لهذا الوباء الى اثاره اذهاننا و ايجاد الحلول و الاستفادة من التكنولوجيا المتاحة للحد من تأثير الوباء باستخدام الاشعة السينية للصدر بتوفير شبكة عصبية تلافيفية عميقة (CNN) لمساعدة الاطباء على اتخاذ قرارات التشخيص الصحيحة و بعد عدة اختبارات و طرق مختلفة لاستخدام نماذج يمكننا القول اننا حققنا النتيجة المرجوة بنسبة 98.66%.

الكلمات المفتاحية COVID-19: رؤية الكمبيوتر، التعلم العميق، التعلم الالي، الشبكة العصبية التلافيفية (CNN)، الاشعة السينية على الصدر.

Table des matières

Table des matières	V
Liste des tableaux	IX
Liste des figures	X
Liste des abréviations :	XII
Introduction générale	1
Chapitre 1 : COVID-19	
1.1 Introduction :	3
1.2 Aperçu:	3
1.3 Contexte :	4
1.4 Motivation :	5
1.5 Historique de la pandémie en Algérie :	5
1.6 Les type de coronavirus :	5
1.7 Transmission et propagation de Covid-19 :	6
1.8 Symptômes de la Covid-19 :	7
1.8.1 Symptômes les plus courants:	7
1.8.2 Symptômes moins courants:	7
1.8.3 Symptômes graves:	8
1.9 Conclusion :	8
Chapitre 2 : Imagerie Médicales	
2.1 Introduction :	9
2.2 vision par ordinateur:	9
2.2.1 Objectif de vision par ordinateur :	10
2.3 Une image :	10
2.4 Images Médicales :	11
2.4.1 Définition :	11
2.4.2 Objectifs d'imagerie médicale :	11

2.4.3 Les principaux types d'images médicales:.....	11
2.4.4 Différentes modalités d'imagerie médicale :.....	12
2.4.4.1 Les Rayons X :	12
2.4.4.2 CT scan (tomodensitométrie) :.....	12
2.4.4.3 IRM :	12
2.4.4.4 Ultrason :.....	13
2.5 Traitement d'images médicales :	13
2.6 La relation entre covid-19 et les images chest x-ray :.....	14
2.7 Radiographie pulmonaire :.....	14
2.8 Le big data et le défi de l'imagerie médicale :.....	15
2.9 Apprentissage profond en imagerie médicale :	16
2.10 Conclusion :.....	17

Chapitre 3 : IA (Artificial Intelligence)...

3.1 Introduction :.....	18
3.2 Artificiel Intelligence :.....	19
3.3 Machine Learning et deep learning:	19
3.3.1 Machine Learning:.....	19
3.3.1.1 Pourquoi utiliser machine learning ?	19
3.3.2 deep learning:.....	20
3.4 Domaine de l'apprentissage automatique :.....	20
3.5 Paradigme d'apprentissage automatique:.....	20
3.5.1 Apprentissage supervisé (Supervised Learning):.....	21
3.5.1.1 Classification:.....	21
3.5.1.2 Régression:	21
3.5.2 Apprentissage non supervisé (Unsupervised Learning) :.....	21
3.5.3 Apprentissage Semi-supervisé (Semi-supervised Learning) :.....	22
3.5.4 Apprentissage Renforcement (Reinforcement Learning):	22
3.6 Rseau de Neurones ou ANN:.....	22
3.6.1 Neurone Biologique:	22
3.7 Classification des images :.....	23
3.8 Méthodes de classification des images :.....	24
3.8.1 Réseau neuronal récurrent (RNN):.....	24

3.8.2 Arbre de décision :	24
3.8.3 Les machines à vecteurs de support(SVM):	24
3.9 Conclusion :	24

Chapitre 4 : Reseaux de neurones onvolutifs(CNN)

4.1 Introduction :	25
4.2 Définition de CNN :	25
4.2.1 Convolution Layer :	26
4.2.2 Pooling Layer:	27
4.2.3 Fully-connected Layer (FC Layer):	27
4.3 Fonction d'activation :	28
4.3.1 Fonction ReLU (Rectified Linear Unit):	28
4.3.2 Fonction Sigmoid :	28
4.3.3 Fonction Softmax :	29
4.3.4 Dropout :	29
4.4 Quelques Modèles de CNN :	30
4.4.1 VGG16 :	30
4.4.1.1 Architecture:	30
4.4.2 RESNET 50 :	31
4.4.2.1 Architecture:	32
4.5 Conclusion:	32

Chapitre 5 : Conception et implementation.....

5.1 Introduction :	33
5.2 Architecture de notre modèle :	33
5.2.1 Prétraitement :	34
5.2.1.1 Normalisation :	34
5.2.1.2 Augmentation :	34
5.3 Préparation de l'expérience :	34
5.3.1 Exécution :	34
5.3.2 Base de données :	35
5.4 Extraction de fonctionnalités :	37
5.5 Résultat de l'expérience :	38

5.5.1 Code d'explication :.....	38
5.5.2 Résultats et discussion :	40
5.5.3 Fonction de perte et fonction de précision :.....	40
5.5.4 La matrice de confusion (MC) :	42
5.5.4.1 Termes de la matrice de confusion :.....	42
5.5.4.2 Mesures d'évaluation du classificateur :.....	43
5.5.4.3 Résultat de la matrice de confusion :.....	45
5.5.5 Résultat général:	45
5.5.6 Discussion :.....	46
5.6 Conclusion :	47
Conclusion générale& perspectives	48
Bibliographie	50
Webographie.....	53

Liste des tableaux

TAB. 2 .1 : Les avantages et les inconvénientsd'férence modalité d'imagerie médicale.	13
TAB. 2 .2 : Types et mécanismes des différentes modalités d'imagerie.	15
TAB. 5 .1 : distribution d'ensembles de données.....	36
TAB. 5 .2 : La matrice de confusion.	42
TAB. 5 .3 : Résultats extraits.....	46
TAB. 5 .4 : Accuracy.	46

Liste des figures

FIG-1.1 : Une image montrant la forme du coronavirus.....	3
FIG-2.1 : vision par ordinateur	9
FIG-2.2 : image médicale	11
FIG-2.3 : Radiographie pulmonaire.	15
FIG-3.1 : La relation entre l'IA, le ML et le deep learning	18
FIG-3.2 : Type d'apprentissage en ML.....	20
FIG-3.3 : Réseau neuronal	23
FIG-4.1 : Architecture générale du système.....	26
FIG-4.2:Exemple de principe du filtre convolutionnel	26
FIG-4.3:Exemple de principe du Pooling	27
FIG-4.4:Principe de la couche entièrement connectée (FC)	27
FIG-4.5 : Fonction ReLU	28
FIG-4.6 : Fonction sigmoïde	28
FIG-4.7 : Dropout	29
FIG-4.8: VGG16 Architecture.....	31
FIG-4.9: RESNET 50 Architecture	32
FIG-5.1:Pour une méthode de diagnostic automatisée proposée pour COVID-19.	33
FIG-5.2 : Schéma fonctionnel du processus d'apprentissage du modèle de Classification COVID-19 basé sur CNN.	37
FIG-5.3 : Modèle Resnet pour la segmentation.	38
FIG-5.4 : Génération du mode.....	39
FIG-5.5 : Structure du réseau d'impression.	39
FIG-5.6 : Code de démarrage du train et du test.	39
FIG-5.7 : Code de courbe.....	39
FIG-5.8 : Code de Matrice.	40
FIG-5.9 : Le résultat de la structure du réseau d'impression.	40
FIG-5.10 : Résultats de l'entraînement du modèle avec les images.	41

FIG-5.11 : courbe de loss et validation loss.....	41
FIG-5.12 : courbe d'accuracy et validation accuracy.	42
FIG-5.13: Résultat de la matrice de confusion.	45

Liste des abréviations :

- MERS-COV : le Coronavirus de Syndrome Respiratoire de Moyen-Orient.
- SARS-COV : Server Acute Respiratory Syndrome.
- IRM : Imagerie par ResonanceMagnetique.
- NMR : NuclearMagneticResonance.
- TDM : Tomodensitometrie ou Scanner.
- TEP : Tomographie par Emission de Positrones.
- IA : Intelligence Artificielle.
- ML : Machines Learning.
- DL : Deep Learning.
- ANN : Artificial Neural Network.
- SVM : support-vector networks.
- CM : Confusion Matrix.
- PA : Antérieur-postérieur.
- RT-PCR : Reverse Transcriptase-Polymerase Chain Reaction.

Introduction générale

COVID-19 est actuellement classé comme l'une des principales maladies infectieuses dans le monde, Le coronavirus 2019 a été signalé pour la première fois à Wuhan, puis dans de nombreuses régions du monde. Cette maladie a été signalée presque partout dans le monde et provoque des effets secondaires graves. La détection précoce du COVID-19 est essentielle non seulement pour les soins aux patients mais aussi pour la santé publique en assurant l'isolement des patients et le contrôle de l'épidémie, la RT-PCR a été l'étalon-or pour diagnostiquer le COVID-19 mais a des aspects limités avec certaines caractéristiques qui le rendent difficile pour diagnostiquer la maladie. La RT-PCR est un processus manuel qui prend du temps est complexe, très coûteux et a une faible sensibilité.

Les radiographies sont également connues comme normes pour indiquer l'état de la maladie et les Raisons du diagnostic chez les patients présentant des symptômes de maladie respiratoire Le système à rayons X est largement utilisé en raison de sa production de rayonnement rapide, peu coûteuse, plus fiable et de son accessibilité. Le processus de diagnostic de la maladie Covid-19 sera plus rapide et plus intelligent et contribuera à réduire la charge pesant sur le secteur médical.

Notre objectif est d'essayer de faire une classification automatique des images X-Ray de COVID 19, cette classification nous permet de détecter les maladies atteints du Covid-19 en utilisant une des techniques de l'intelligence artificielle.

Contenu du mémoire :

Chapitre 1:Cela inclut l'émergence de Covid-19, ses types et ses effets.

Chapitre 2: Nous allons parler des images médicales et de leur relation avec la technologie.

Chapitre 3: Dans ce chapitre, nous discuterons de l'intelligence artificielle et des fondements de l'apprentissage automatique et de ses algorithmes.

Chapitre 4: Nous allons aborder le concept de réseau de neurones convolutionnel (CNN), nous introduirons son architecture et les différentes couches et hyper paramètres d'un réseau CNN.

Chapitre 5: Ce chapitre se compose de deux parties dans la première nous allons expliquer l'architecture du système ainsi que les différentes étapes de conception et la deuxième partie illustre l'environnement du travail ainsi que les résultats obtenus.

1

COVID-19

1.1 Introduction :

Les coronavirus sont une grande famille de virus qui causent une variété de maladies chez l'homme, du rhume au syndrome respiratoire du Moyen-Orient (MERS) et au syndrome respiratoire aigu sévère (SRAS). Le nouveau coronavirus (COVID-19) a été découvert à Wuhan, en Chine, en 2019. Il s'agit d'un nouveau type de coronavirus qui n'a pas encore été trouvé chez l'homme. Ce chapitre donne un aperçu de la COVID-19 et des virus respiratoires émergents ; Il est destiné aux professionnels de la santé publique, aux gestionnaires d'incidents, aux coordonnateurs d'urgence et aux collaborateurs travaillant avec les Nations Unies, les organisations internationales et les ONG.

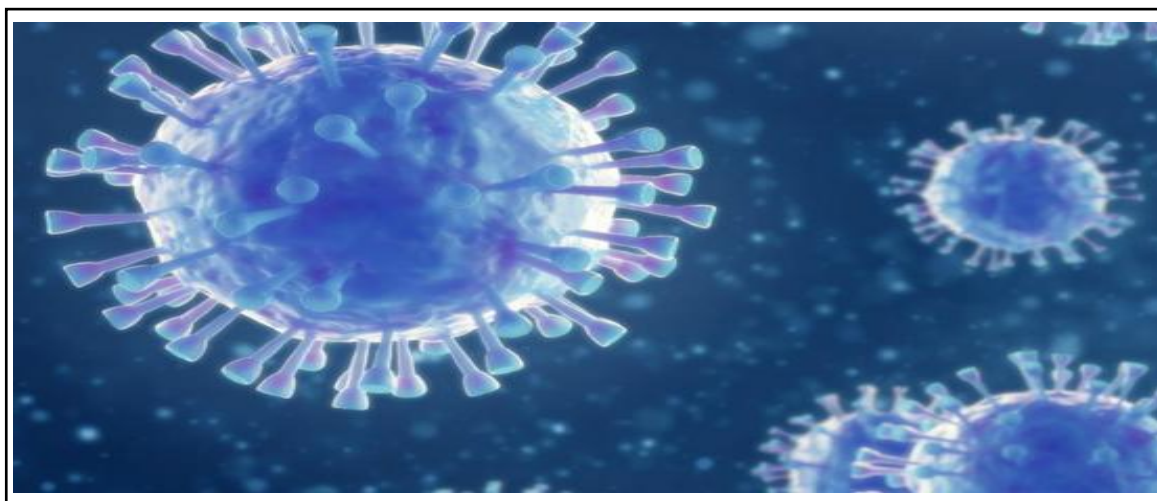


FIG-1.1 : Une image montrant la forme du coronavirus [Web1].

1.2 Aperçu:

La pandémie de virus corona de 2020 s'est propagée en Algérie à partir du 25

Février 2020 et la première personne à avoir été testée positive pour la maladie à virus corona était un citoyen italien. Ensuite, d'autres cas infectés par le Covid-19 ont été révélés, et le total des cas confirmés en Algérie s'élevait à 125 102 cas, dont 3 254 décès. L'Algérie a pris plusieurs mesures afin de limiter sa propagation, à savoir une augmentation des lits de récupération et la préparation d'équipements médicaux tels que des ventilateurs et autres, avec la possibilité d'importer les fournitures nécessaires d'autres pays. Malgré tout cela, il n'a pas été en mesure de contenir cette épidémie, et ce pour plusieurs raisons, dont les fréquentes des visites dans les hôpitaux pour se contrôler, ce qui contribue à la propagation du virus, car cela a conduit à une crise de pénurie de lits de récupération et a promis certains médicaments et d'autres. Afin d'améliorer les conditions de vie et de contribuer à réduire la propagation de ce virus mortel, nous avons décidé d'introduire des technologies modernes pour en bénéficier en créant une application qui permet à une personne de se tester pendant qu'elle est à la maison, ce qui conduit à réduire sa propagation, et nous avons également formé un modèle capable de détecter Corona 19 Sur la base de la radiographie pulmonaire, l'une des raisons qui nous a fait renoncer aux tomodensitogrammes pour le calcul des radiographies est que les tomodensitogrammes sont très coûteux pour le patient et qu'il y a un pénurie de leur équipement, nous avons donc utilisé radiographies au détriment des tomodensitogrammes.

1.3 Contexte :

Le monde a souffert de l'épidémie massive de coronavirus, ou COVID-19, qui a été détectée pour la première fois en décembre 2019 à Wuhan, en Chine. Les principaux signes de la COVID-19 sont une température élevée, une toux sèche, un rhume, un essoufflement, une combinaison de perte d'odorat et d'oppression thoracique. Afin de l'éviter, vous devez prendre quelques précautions simples, telles que l'éloignement physique, le port d'un masque, la ventilation des pièces, l'évitement des foules, le nettoyage des mains et la toux dans un coude ou un mouchoir plié.

L'une des raisons qui nous a fait entrer dans cet aspect et y travailler est :

La propagation massive et rapide de celui-ci Bien que l'on en sache beaucoup sur la propagation du COVID-19, de nombreuses connaissances sont encore en développement.

- ❖ L'urgence de trouver une solution rapide qui limite la propagation du virus.
- ❖ Bénéficier du développement de la technologie ces dernières années et l'exploiter dans cette circonstance.

1.4 Motivation :

Les motivations qui poussent à travailler sur un tel sujet sont d'abord des raisons humanitaires. Parmi les principales raisons figure la détection précoce de la maladie et la réduction de sa propagation. Le deuxième motif est purement scientifique, la grande compétition technologique nous donne envie d'entrer dans la compétition ainsi que la manière d'introduire la technologie dans la plupart des domaines de la vie quotidienne de l'individu et de la société afin de réduire la souffrance et les dangers qui l'entourent. Le domaine est également très intéressant, nous espérons que tout le monde le verra et aura une idée de ce qui se passe dans le monde de la technologie.

1.5 Historique de la pandémie en Algérie :

Le premier cas, un ressortissant italien, a été notifié le 25 février 2020 dans une base de vie Hassi Messaoud dans la wilaya d'Ouargla. A partir du 02 mars 2020 un foyer a été détecté dans la wilaya de Blida suite à une alerte lancée par la France après la confirmation au Covid-19 de deux citoyens Algériens résidant en France ayant séjourné en Algérie. Depuis l'épidémie s'est étendue à l'ensemble du territoire national [Web17].

1.6 Les type de coronavirus :

Il existe quatre sous-groupes principaux de coronavirus, appelés alpha, bêta, gamma et delta. Et sept formes différentes dont quatre communes (moins graves que les autres) [10].

- ❖ 229E (alpha coronavirus).
- ❖ NL63 (alpha coronavirus).
- ❖ OC43 (bêta-coronavirus).
- ❖ HKU1 (bêta- coronavirus).
- ❖ MERS-COV (le coronavirus bêta qui provoque le syndrome respiratoire du Moyen-Orient, ou MERS découvert en 2012 en Arabie Saoudite).
- ❖ SARS-COV (le bêta-coronavirus qui cause le syndrome respiratoire aigu sévère, SARS, identifié en Chine en 2002).Ce coronavirus a beaucoup de similitudes avec celui du SRAS (origine animale, identique génétiquement à 80%, responsable d'infections pulmonaires) mais aussi des différences notables pour les scientifiques au niveau de sa contagion. Il est contagieux dès le début des symptômes voire parfois en l'absence de symptômes alors que le Sars l'était quelques jours après les premiers symptômes. Il existe aussi des formes bénignes et asymptomatiques du Covid-19 alors que le Sars n'entraînait que des formes sévères [11].

1.7 Transmission et propagation de Covid-19 :

La covid-19 se transmet par des personnes porteuses du virus. La maladie se propage principalement d'un homme à l'autre lors de contacts rapprochés par le biais de gouttelettes respiratoires expulsées par le nez ou par la bouche lorsqu'une personne malade tousse, éternue ou parle. Ces gouttelettes sont relativement lourdes, ne parcourent pas de grandes distances et tombent rapidement au sol .Il est possible de contracter le covid-19 en cas d'inhalation de ces gouttelettes [Web18].Toute situation dans laquelle des personnes sont à proximité immédiate les unes des autres pendant de longues périodes augmente le risque de transmission. Les espaces intérieurs, en particulier lorsqu'ils sont mal ventilés, présentent plus de risques que les espaces en plein air. Les activités lors desquelles le volume de particules expulsées par la bouche est plus important, comme le fait de chanter ou de respirer fortement pendant l'exercice physique, augmentent également le risque de transmission [Web19].La transmission est facilitée dans les lieux et situations répondant aux 3 critères suivants :

- ❖ Espaces bondés.

- ❖ Contact étroit, par exemple lorsque des personnes discutent en se tenant très près les unes des autres.
- ❖ Espaces confinés et clos, mal ventilés.
- ❖ Le risque de propagation de la Covid-19 est plus élevé dans les endroits où ces trois critères coïncident. Il est donc important d'appliquer également les règles d'hygiène respiratoire (par exemple, en se couvrant la bouche et le nez avec le pli du coude lorsque l'on tousse) et aussi maintenir à distance les personnes. Pour prévenir l'infection et ralentir la transmission du Covid-19.

1.8 Symptômes de la Covid-19 :

Les manifestations des coronavirus (covid-19) font leur apparition moins de 24 heures après l'infection. Le plus généralement, le virus entraîne des maladies respiratoires légères à modérées comme le rhume avec des symptômes tels que [Web16]:

1.8.1 Symptômes les plus courants:

- ❖ Fièvre.
- ❖ Toux sèche.
- ❖ Fatigue.

1.8.2 Symptômes moins courants:

- ❖ Courbatures et douleurs.
- ❖ Maux de gorge.
- ❖ Diarrhée.
- ❖ Conjonctivite.
- ❖ Maux de tête.
- ❖ Perte du goût ou de l'odorat.
- ❖ Eruption cutanée ou décoloration des doigts ou des orteils.

1.8.3 Symptômes graves:

- ❖ Difficultés respiratoires ou essoufflement.
- ❖ Douleurs ou sensation d'oppression dans la poitrine.
- ❖ Perte de la parole ou difficultés à se déplacer.

1.9 Conclusion :

Dans ce chapitre, nous avons parlé de cette nouvelle pandémie du Corona Virus dites (Covid-19) et expliquer d'où provienne cette maladie, ses causes et ses conséquences sur la santé de l'être humain, aussi ses symptômes et effets .Dans le chapitre suivant, nous allons présenter notre approche de prédiction basée sur les modèles du Machine Learning supervisé particulièrement.

2

Imagerie Médicale

2.1 Introduction :

La segmentation d'images est un problème important dans le domaine de l'analyse d'images. Par exemple, elle est utilisée en imagerie médicale pour analyser et identifier les différentes structures de filtres présentes dans les images. Les méthodes de segmentation peuvent être classées en fonction de l'objectif à atteindre. Cette technique permet visualiser indirectement la physiologie ou l'imagerie fonctionnelle du corps humain, mais il est également utilisé dans divers domaines tels que l'industrie, la chimie, etc.

2.2 vision par ordinateur:

C'est un domaine de l'intelligence artificielle qui entraîne les ordinateurs à interpréter et à comprendre le monde visuel. À l'aide d'images numériques provenant de caméras, de vidéos et de modèles d'apprentissage en profondeur, les machines peuvent identifier et classer les objets avec plus de précision que les humains, puis interagir avec ce qu'elles voient [Web2].



FIG-2.1 : vision par ordinateur [Web2].

2.2.1 Objectif de vision par ordinateur :

Il s'agit de créer des applications intelligentes capables de comprendre le contenu des images telles que les humains les comprennent, car les données d'image peuvent prendre de nombreuses formes.....

- ❖ Par images successives (vidéo).
- ❖ Scènes de plusieurs caméras.
- ❖ Données de plusieurs dimensions issues d'un appareil d'imagerie médicale.

Les images sur les ordinateurs sont souvent stockées sous forme de grands réseaux de pixels. Chaque pixel est stocké par une couleur, qui est stockée sous la forme d'un mélange de trois couleurs primaires ajoutées, rouge, vert et bleu, R G B, et en combinant les différentes intensités de ces trois couleurs produit ce qu'on appelle une valeur de pixel. Si nous pouvons représenter n'importe quelle couleur [Web7].

2.3 Une image :

Une image est une représentation du monde, une représentation d'une personne ou d'un objet par la peinture, la sculpture, le dessin, la photographie, le film, etc. C'est aussi un ensemble structuré d'informations qui, après affichage sur l'écran, ont une signification pour l'œil humain [Web26]. On peut définir une image comme un tableau bidimensionnel dont chaque élément (pixel) représente une surface élémentaire de l'image. La disposition de ces pixels est généralement en ligne et colonne [Web27]. Mathématiquement, Une image est la projection sur un plan d'une scène 3D, Elle peut être définie comme une fonction à deux variables $f(x, y)$ où (x, y) est la position d'un point de l'espace sur le plan de projection et $f(x, y)$ est l'intensité (ou brillance) au point de coordonnées (x, y) [20]. Il existe plusieurs types d'images Image binaire, Image couleur RGB, Image en niveau de gris...etc. L'image a des caractéristiques : Pixel, Dimension, Bruit.

2.4 Images Médicales :

2.4.1 Définition :

L'imagerie médicale est le processus de production d'images visuelles des structures internes du corps et de vues visuelles de la fonction des tissus internes pour une utilisation en médecine, en recherche scientifique et en traitement. Ce processus crée une base de données de la structure et de la fonction régulières des organes pour aider à identifier les anomalies [Web28].

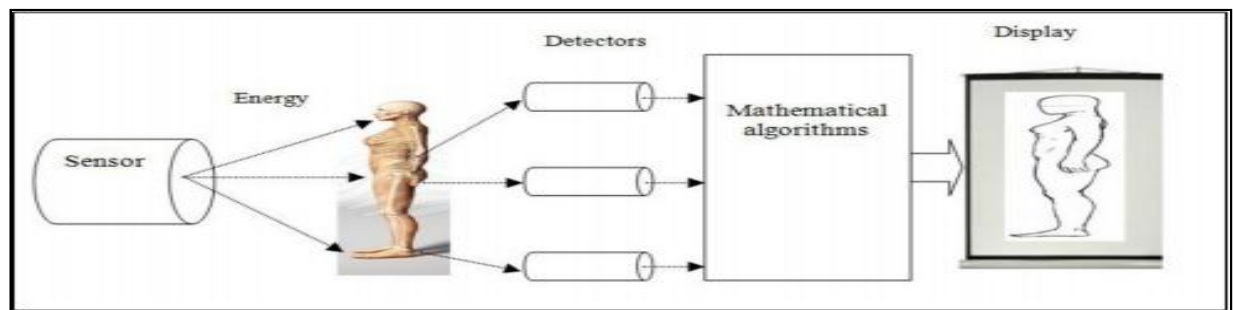


FIG-2.2 : image médicale[Web28].

2.4.2 Objectifs d'imagerie médicale :

- ❖ Aider à la décision, diagnostiquer les maladies, suivre leur évolution, découvrir leur fonctionnement et mieux les soigner [1].
- ❖ Traitement d'une grande quantité de données.
- ❖ Procédures médicales et chirurgicales assistées par ordinateur.
- ❖ Perception scientifique.

2.4.3 Les principaux types d'images médicales:

Les images proviennent principalement de rayons X, de scanners, d'IRM, d'échographie, d'imagerie de médecine nucléaire ou de scintigraphie. Les images médicales fournissent des informations sur la forme et la fonction des organes humains [13].

2.4.4 Différentes modalités d'imagerie médicale :

Les différents outils d'obtention d'images médicales sont nombreux et faciles à obtenir différentes informations selon le processus physique utilisé pour observer les tissus du corps humain. Certaines procédures fournissent des informations anatomiques (imagerie structurelle) tandis que d'autres fournissent détectez la fonctionnalité de l'appareil en fournissant une carte d'activité [Web23].

2.4.4.1 Les Rayons X :

Les rayons X sont des rayons électromagnétiques du même type que la lumière mais non visibles et très animé. Ainsi, ils peuvent pénétrer profondément dans la matière vivante [16].

2.4.4.2 CT scan (tomodensitométrie) :

Le scanner, aussi appelé tomographie dans le jargon scientifique, repose sur le même principe que la radiologie, c'est-à-dire l'utilisation d'une source de rayons X. Il permet d'étudier différentes parties du corps humain, comme le cerveau, la cage thoracique, l'abdomen ou encore les os .Ainsi, il permet de rechercher des anomalies non visibles sur les radiographies [17].

2.4.4.3 IRM :

La résonance magnétique nucléaire est une technologie qui a été développée au cours des 50 dernières années de l'année. Le concept de base pour décrire le phénomène de RMN utilise des lois simples. Le principe de l'imagerie par résonance magnétique est une puissante technologie de diagnostic médical qui fournit des images tridimensionnel et coupé avec une grande précision anatomique. L'imagerie par résonance magnétique est une technique radiologique Moderne, non chirurgical et sans effets secondaires connus, basé sur le phénomène de résonance physique magnétique nucléaire. C'est simplement une question de surveillance par résonance magnétique nucléaire

(RMN). Protons d'eau dans le corps, c'est-à-dire la réponse des noyaux exposés au champ magnétique externe et excitation électromagnétique [18].

2.4.4.4 Ultrason :

L'échographie est un type de radiologie diagnostique utilisé pour détecter des maladies ou des problèmes de santé, dans lequel les ondes sonores sont utilisées à la place du rayonnement pour créer des instantanés ou transmettre des images de structures à l'intérieur du corps.

❖ Les avantages et les inconvénients de la modalité d'imagerie médicale :

	Les avantages	Les inconvénients
Image Médicale	<ul style="list-style-type: none"> ✓ La radiologie est particulièrement adaptée à l'observation des tissus durs, ainsi que des structures dures, des os et des dents. C'est une technique couramment utilisée pour observer les os et le système squelettique, et notamment pour détecter les fractures en cas de traumatisme. ✓ Haute vitesse et précision. ✓ angiographie Etudes. 	<ul style="list-style-type: none"> ⊗ L'imagerie médicale n'est pas une tâche facile et elle doit être utilisée avec précision, lorsque nous faisons des radiographies ou des tomodensitométries, nous sommes exposés aux rayons X et leur accumulation, si nous devons les faire régulièrement, peut entraîner des risques de cancer.

TAB. 2 .1 : Les avantages et les inconvénients de la modalité d'imagerie médicale.

2.5 Traitement d'images médicales :

Le traitement d'images médicales implique le développement de méthodes spécifiques à un problème pour optimiser les données d'images médicales brutes pour une visualisation sélective et une analyse plus approfondie. Il existe de nombreux sujets

en traitement d'images médicales : certains se concentrent sur la théorie appliquée générale, tandis que d'autres se concentrent sur des applications spécifiques. Nous nous concentrons principalement sur la classification des images de radiographie pulmonaire [15].

2.6 La relation entre covid-19 et les images chest x-ray :

Dans les techniques d'IA existantes, toutes les approches détectent la pneumonie causée par COVID-19 en utilisant des images de radiographie pulmonaire (data base)[Web2]. Parmi les défis auxquels l'IA est confrontée en ce qui concerne la détection de la pneumonie, il y a «comment le système réalise que la la détection de la pneumonie dans la radiographie pulmonaire est causée par COVID-19 », de sorte que la pneumonie peut être causée par de nombreux autres virus, Bactriane et par attaque fongique. Dans la pratique courante, tous les chercheurs s'accordent à poser le problème alors que «la majorité des décès dus de COVID-19 sont dus à une pneumonie dans les poumons de patients vulnérables », a déclaré le Dr Tom Naunton Morgan, médecin-chef de voici.ai. «La pneumonie est une maladie potentiellement mortelle causée par plusieurs agents pathogènes, y compris, directement ou indirectement, Infection au COVID-19 ».

2.7 Radiographie pulmonaire :

Les radiographies thoraciques sont souvent prises dans les hôpitaux comme outil d'imagerie diagnostique important pour identifier les anomalies thoraciques. Deux côtés sont souvent capturés : la vue de face et la vue de côté. La figure 2 montre les radiographies thoraciques antérieure et latérale du patient, respectivement. Pour le diagnostic des maladies pulmonaires assistées par ordinateur (CAD), la segmentation des régions pulmonaires à partir des radiographies thoraciques est une partie importante du système. Pour chaque méthode, une connaissance préalable de la radiographie thoracique est nécessaire pour appliquer le tableau. Non seulement la forme du poumon externe diffère selon la vue de l'image, mais les caractéristiques du poumon interne peuvent également différer, comme le montre la FIG-2.3.Par

conséquent, les classificateurs développés pour la classification peuvent devoir être formés différemment. Selon la façon dont l'image est affichée.

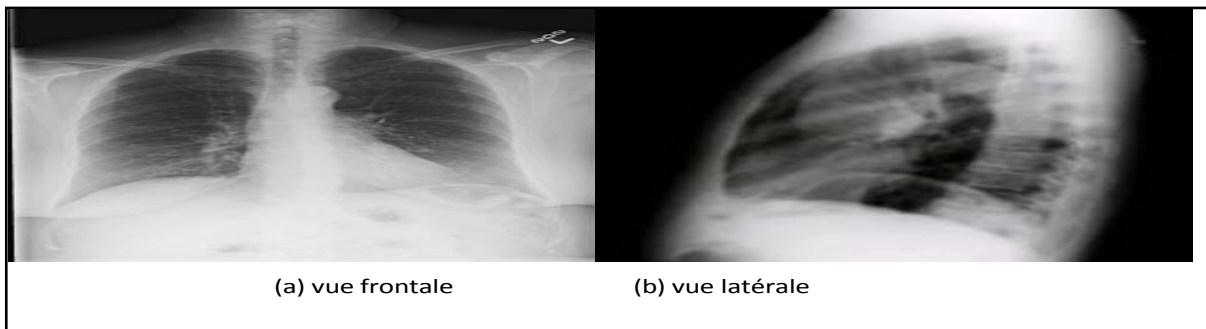


FIG-2.3 : Radiographie pulmonaire.

Type	Mécanisme	Modalité	2D/3D
Morphologique	Absorption des rayons X	Radiographie	Projection 2D
		Angiographie numérisée	
		Angiographie rotationnelle	3D
		TDM (scanner)	
	Réflexion des ultrasons	Échographie	Coupes 2D
		Échographie 3D	3D
Fonctionnelle	Émission de photons gamma	Scintigraphie	Projection 2D
		Tomoscintigraphie	3D
	Émission de positons	TEP	3D
Mixte	Échos de l'aimantation des noyaux	IRM	3D

TAB. 2 .2 : Types et mécanismes des différentes modalités d'imagerie.

2.8 Le big data et le défi de l'imagerie médicale :

❖ Dans n'importe quelle industrie dans le monde, comment pouvez-vous faciliter, analyser et gérer le changement peut le faire. Le domaine le plus prometteur dans lequel les méga données sont utilisées pour faire une grande différence est celui de la santé. En général, pour améliorer la qualité de vie, réduire les coûts de traitement et

prévenir les maladies évitables, prévoir les flambées épidémiques, analyse médicale du potentiel.

❖ Le big data est très utile pour l'imagerie médicale. Cela trouve aussi son utilité. Surveiller, diagnostiquer ou traiter des conditions médicales pour comprendre les différentes techniques utilisées en imagerie médicale du corps humain. L'imagerie médicale est couramment utilisée pour identifier toutes les techniques qui produisent des images de l'intérieur du corps. À des fins cliniques telles que la déficience fonctionnelle, l'examen des blessures, la recherche de détection, la pathologie ou le diagnostic pour créer des images, des procédures et des techniques.

❖ En médecine et en soins de santé, les méga données sont définies comme des données complexes et importantes qui étaient auparavant difficiles à analyser par du matériel ou des logiciels. Analysez et vérifiez les données, surveillez la qualité des données, modélisez, interprétez et intégrez la couverture de données hétérogènes avec l'analyse du Big Data. À partir de la grande quantité de données disponibles, en appliquant l'analyse de données volumineuses, offrant une découverte complète des connaissances grâce aux applications analytiques du Big Data.

2.9 Apprentissage profond en imagerie médicale :

Bien que les modèles d'apprentissage profond aient connu beaucoup de succès dans l'analyse d'images médicales, les ensembles de données médicales de petite taille demeurent le principal obstacle dans ce domaine et continueront d'être un défi pour nous. L'imagerie médicale est l'opération de production d'images visuelles des structures internes de le corps pour l'étude et le traitement médical et scientifique, ainsi qu'une vue visible de la fonction des tissus intérieurs [2]. L'une des tâches de l'apprentissage profond est d'analyser et d'interpréter des images médicales, et d'extraire des informations utiles pour aider à diagnostiquer et planifier le traitement. Ce processus crée une banque de données de la structure stable et fonction des organes pour qu'il soit facile de reconnaître les anomalies. Ce processus identification et gestion des troubles Les algorithmes d'apprentissage profond, en

particulier les réseaux convolutionnels, deviennent rapidement une méthodologie privilégiée pour l'analyse d'images médicales [3.4].

2.10 Conclusion :

Dans ce chapitre, nous avons discuté de l'explication des images médicales. Le traitement d'images médicales concerne le développement d'approches spécifiques à un problème pour améliorer les données d'images médicales brutes à des fins de présentation sélective ainsi que pour une analyse plus approfondie. Nous nous sommes principalement concentrés sur la classification des radiographies.

3

IA (Artificielle Intelligence)

3.1 Introduction :

Actuellement, nous sommes au début d'un nouveau concept technologique C'est une intelligence artificielle qui n'épargnera aucun secteur. La présence d'un grand nombre de Algorithmes et données stockés dans des bases de données sont reliés entre eux grâce à des réseaux de plus en plus étendus. L'intelligence artificielle est devenue un sujet d'une grande importance à notre époque, car ce développement technologique a pu résoudre de nombreux problèmes complexes, ce que d'autres technologies ont échoué. Cette idée est incassable Technologique, c'est plutôt un continuum informatique. L'intelligence artificielle, qui est considérée comme l'intelligence fournie par machine, pour développer et faire progresser de nombreux domaines, y compris Santé, industrie, commerce et agriculture.

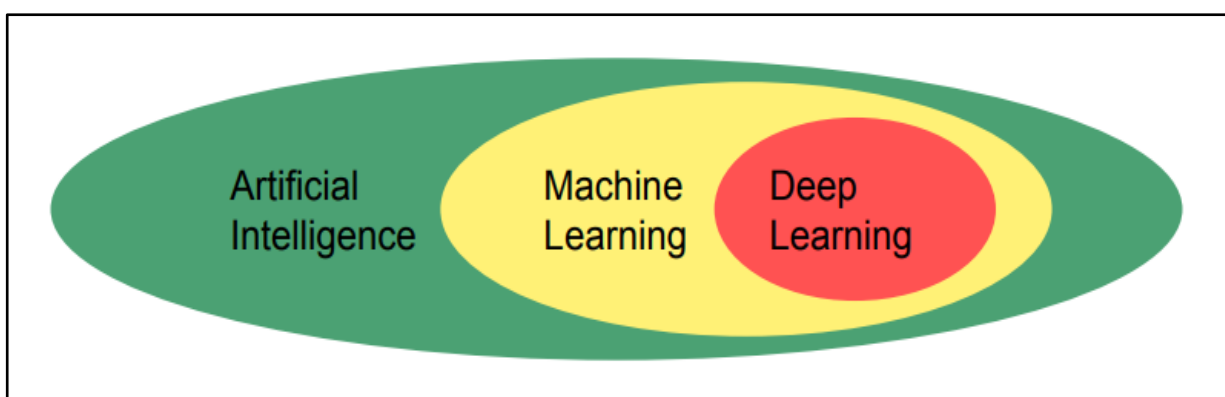


FIG-3.1 :La relation entre l'IA, le ML et le deeplearning[Web29].

3.2 Artificiel Intelligence :

L'intelligence artificielle est la simulation des processus de l'intelligence humaine par des machines, en particulier des systèmes informatiques. Les applications spécifiques de l'IA incluent les systèmes experts, le traitement du langage naturel, la reconnaissance vocale et la vision par ordinateur [Web24].

3.3 Machine Learning et deep learning:

3.3.1 Machine Learning:

Le Machine Learning (ML) est l'un des domaines de l'informatique qui peut donner du sens aux données à l'aide de certains systèmes informatiques afin d'imiter la capacité humaine à le faire. En termes simples, ML est un type d'IA qui extrait des modèles à partir de données brutes à l'aide d'un algorithme ou d'une méthode. L'objectif principal de ML est de permettre aux systèmes informatiques d'apprendre de l'expérience sans intervention humaine [12].

3.3.1.1 Pourquoi utiliser machine learning ?

Machine Learning peut servir à résoudre des problèmes :

- ❖ Que l'on ne sait pas résoudre.
- ❖ Que l'on sait résoudre, mais dont on ne sait formaliser en termes algorithmiques comment nous les résolvons (c'est le cas par exemple de la reconnaissance d'images ou de la compréhension du langage naturel).
- ❖ Que l'on sait résoudre, mais avec des procédures beaucoup trop gourmandes en ressources informatiques (c'est le cas par exemple de la prédiction d'interactions entre molécules de grande taille, pour lesquelles les simulations sont très lourdes).

Le machine Learning est donc utilisé quand les données sont abondantes (relativement), mais les connaissances peu accessibles ou peu développées. Ainsi, la machine Learning peut aussi aider les humains à apprendre : les modèles créés par des algorithmes d'apprentissage peuvent révéler l'importance relative de certaines informations ou la façon dont elles interagissent entre elles pour résoudre un problème particulier.

3.3.2 deeplearning:

Le Deeplearning est une sorte ML dans l'IA qui contient des systèmes permettant aux ordinateurs d'apprendre de l'expérience et de comprendre le monde tel qu'il est réellement. Deeplearning s'inspire des réseaux de neurones artificiels, les réseaux de neurones artificiels appelés ANN sont inspirés des réseaux de neurones biologiques humains [Web30].

3.4 Domaine de l'apprentissage automatique :

L'apprentissage automatique est une branche de l'intelligence artificielle (IA) et de l'informatique qui se concentre sur l'utilisation de données et d'algorithmes pour simuler la façon dont les humains apprennent et améliorer progressivement sa précision .L'apprentissage automatique est une composante importante du domaine en plein essor de la science des données. Grâce à l'utilisation de méthodes statistiques, les algorithmes sont formés pour effectuer des classifications ou des prédictions, et pour révéler des informations clés sur les projets d'exploration de données. Ces informations orientent ensuite les décisions au sein des applications et des entreprises, affectant idéalement les mesures de croissance clés. Alors que le Big Data continue de se développer et de se développer [Web3].

3.5Paradigme d'apprentissage automatique:

Il y a de nombreux sous domaines ou sous sections:

- ❖ Supervised Learning.
- ❖ Unsupervised Learning.
- ❖ Semi-supervised Learning.
- ❖ Reinforcement Learning.

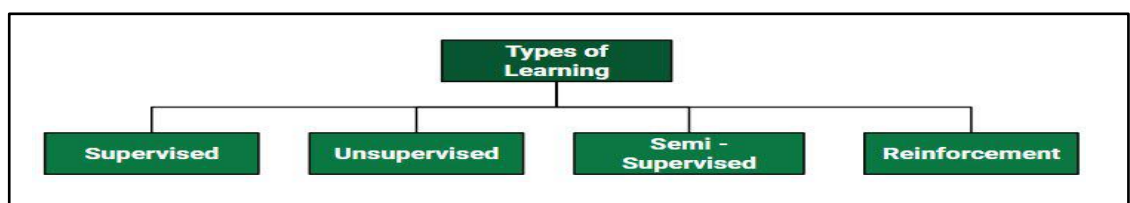


FIG-3.2 :Type d'apprentissageenML.

3.5.1 Apprentissage supervisé (Supervised Learning):

Si les classes cibles sont prédéterminées, le système apprend à classer selon un modèle de classification ou de classement, on parle alors d'apprentissage supervisé. Un expert doit préalablement étiqueter des exemples. Le processus se passe en deux phases. Lors de la première phase (hors ligne, dite d'apprentissage), il s'agit de déterminer un modèle à partir des données étiquetées. La seconde phase (en ligne, dite de test) consiste à prédire l'étiquette d'une nouvelle donnée, connaissant le modèle préalablement appris. Parfois il est préférable d'associer une donnée non pas à une classe unique, mais une probabilité d'appartenance à chacune des classes prédéterminées. Les problèmes dans l'apprentissage supervisé sont 2 types : Classification, Régression [Web4].

3.5.1.1 Classification:

C'est une tâche d'apprentissage supervisé dont la sortie a des étiquettes définies (valeur discrète) ce qu'on appelle classes ou catégories. On peut avoir des classifications Binaire ou alors Multi classe. La classification binaire prédit pour un modèle (0 ou 1) vrai ou faux, dans le cas de multi classes le modèle va prédire plus d'une classe. Le but ici reste le même quel que soit le nombre de classes, prédire si les valeurs discrètes appartiennent à une classe particulière.

3.5.1.2 Régression:

Il s'agit d'une tâche d'apprentissage supervisé dont la sortie à une valeur continue. Le but ici est de prédire une valeur aussi proche de la valeur de sortie réelle puis l'évaluation est effectuée en calculant la valeur d'erreur.

3.5.2 Apprentissage non supervisé (Unsupervised Learning) :

Quand le système ou l'opérateur ne dispose que d'exemples, mais non d'étiquette, et que le nombre de classes et leur nature n'ont pas été prédéterminés, on parle d'apprentissage non supervisé ou clustering en anglais. Aucun expert n'est requis. L'algorithme doit découvrir par lui-même la structure plus ou moins cachée des données. Le partitionnement de données, data clustering en anglais, est un algorithme d'apprentissage non supervisé. Le système doit cibler, dans l'espace de description

(l'ensemble des données), les données selon leurs attributs disponibles, pour les classer en groupes homogènes d'exemples [Web5].

3.5.3 Apprentissage Semi-supervisé (Semi-supervised Learning) :

Nous avons préalablement vu l'apprentissage supervisé et non supervisé, dont la majeure différence réside dans le fait que les données soient étiquetées ou non, et à cela s'ajoute les méthodes adéquates utilisées pour traiter ses données. L'apprentissage semi-supervisé regroupe ses deux principes, il prend un ensemble réduit de données étiquetées avec un autre ensemble de données non étiquetées du même type. L'avantage de ce type d'apprentissage réside principalement dans le processus d'étiquetage des données prend beaucoup de temps et souvent coûteux. Donc paradoxalement le non étiquetage devient bénéfique pour le processus d'apprentissage, et la construction du modèle est moins coûteuse [Web6].

3.5.4 Apprentissage Renforcement (Reinforcement Learning):

C'est une approche de l'intelligence artificielle dans laquelle l'environnement d'apprentissage s'adapte à l'apprenant, car le système adapte ses paramètres en fonction des interactions reçues de l'environnement, qui à son tour réagit aux décisions prises.

3.6 Réseau de Neurones ou ANN:

3.6.1 Neurone Biologique:

Un Neurone ou une Cellule nerveuse est l'unité fonctionnelle de base qui constitue un système nerveux, il assure la transmission d'un signal bioélectrique appelé influx nerveux. Le neurone biologique reçoit en entrées des signaux transmis par d'autres neurones, grâce à une interaction qui est faite entre les dendrites. Le signal reçu est analysé et traité en effectuant une sommation des signaux en entrée, le résultat est par la suite comparé à un seuil d'activation: s'il est supérieur, une décharge est alors envoyée le long de son axone vers d'autres neurones. Un réseau de neurones a toujours:

❖ **Couche d'entrée** : (Input Layer) c'est l'ensemble de neurones qui porte le signal d'entrée du réseau, et par la suite tous les neurones de cette couche sont reliés à la couche suivante.

- ❖ **Couche cachée** : (Hidden layer) elles peuvent être une ou plusieurs, c'est ici où les relations entre les variables vont être mises en exergue. Le choix du nombre de couches et de neurones est intuitif et nécessite de l'expérience venant de l'expert.
- ❖ **Couche de sortie** : (Output Layer) elle représente le résultat du réseau de neurones c'est ce qu'on appelle la prédiction[Web20].

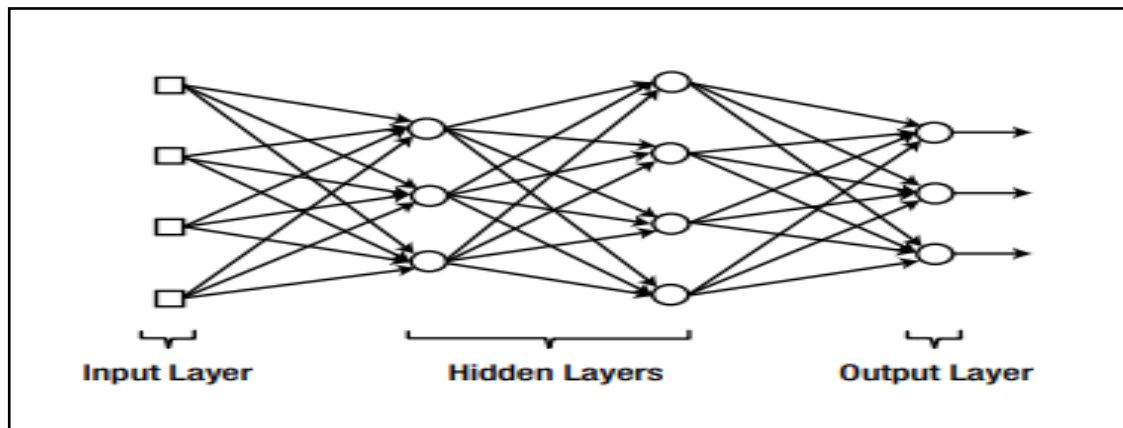


FIG-3.3 :Réseau neuronal [Web25].

3.7 Classification des images :

- ❖ La classification entre les objets est une mission simple pour l'homme, mais elle est plus difficile pour les machines.
- ❖ La classification inclut le prétraitement d'image, les capteurs d'image, l'objet segmentation, la détection d'objets, la classification des objets et l'extraction de fonctionnalités.
- ❖ L'augmentation des ordinateurs de grande capacité, la disponibilité des caméras vidéo et le besoin croissant d'analyse vidéo automatique ont conduit à se concentrer sur les algorithmes de classification d'objets.
- ❖ Un système de classification simple contient d'une caméra haute fixe au-dessus de la zone concernée.
- ❖ La classification des images est une tâche difficile et importante plusieurs domaines d'application, y compris la biométrie, l'imagerie biomédicale, l'inspection visuelle industrielle, la navigation robotisée, la vidéosurveillance et la navigation des véhicules.

❖ Système de classification contiennent de base de données qui se compose de modèles prédéfinis qui comparent avec l'objet détecté à classer dans la catégorie à droite.

3.8 Méthodes de classification des images :

3.8.1 Réseau neuronal récurrent (RNN):

Les réseaux de neurones récurrents RNN , sont un type de réseau de neurones largement utilisé dans le domaine de l'apprentissage en profondeur (Deep Learning) . Les RNN utilisent les sorties précédentes comme entrées supplémentaires et sont parfaitement adaptés au traitement de données séquentielles.

3.8.2 Arbre de décision :

DecisionTree (DT) calcule l'appartenance à la classe en partitionnant à plusieurs reprises un ensemble de données en sous-ensembles uniformes .Un arbre de décision est un outil fort pratique lorsqu'il s'agit de répartir une population en groupes homogènes selon des critères bien précis, les variables de segmentation.

3.8.3 Les machines à vecteurs de support(SVM):

Les SVM sont un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de discrimination et de régression. Les SVM sont une généralisation des classifieurs linéaires [14].

3.9 Conclusion :

L'intelligence artificielle est devenue un sujet d'une grande importance à notre époque car ce développement technologique a pu résoudre de nombreux problèmes. Dans ce chapitre, nous avons abordé les concepts de base de l'intelligence artificielle, nous avons expliqué l'apprentissage automatique et l'apprentissage en profondeur, qui représentent une branche de l'intelligence artificielle et leur importance dans ce domaine, et enfin, nous avons évoqué le concept de réseau de neurones.

4

Réseaux de neurones convolutifs (CNN)

4.1 Introduction :

La vision par ordinateur change rapidement de jour en jour. L'une des raisons est le développement de l'apprentissage en profondeur. Lorsque nous parlons de vision par ordinateur, le terme réseau de neurones convolutifs s'est avéré efficace dans de multiples domaines tels que la reconnaissance et la classification d'images, ce que nous avons utilisé dans notre travail pour concevoir un modèle d'apprentissage en profondeur qui nous permet de catégoriser les images en différentes catégories et puis construisez un modèle pour détecter si les gens ont le COVID-19 [7].

4.2 Définition de CNN :

Un réseau neuronal convolutif est un type particulier de réseau neuronal basé sur processus de torsion .Les réseaux convolutifs sont dérivés des architectures de type Perceptron multicouche (Perceptron multicouche : MLP), mais ils utilisent des poids Commun, lié à la fenêtre de torsion, leur permettant d'extraire implicitement de Caractéristiques locales Les CNN sont particulièrement adaptés reconnaissance d'images .On dit qu'un réseau convolutif reçoit chaque neurone L'information ne provient pas de toute la couche précédente, mais seuls les neurones situés dans leur champ récepteur [19].

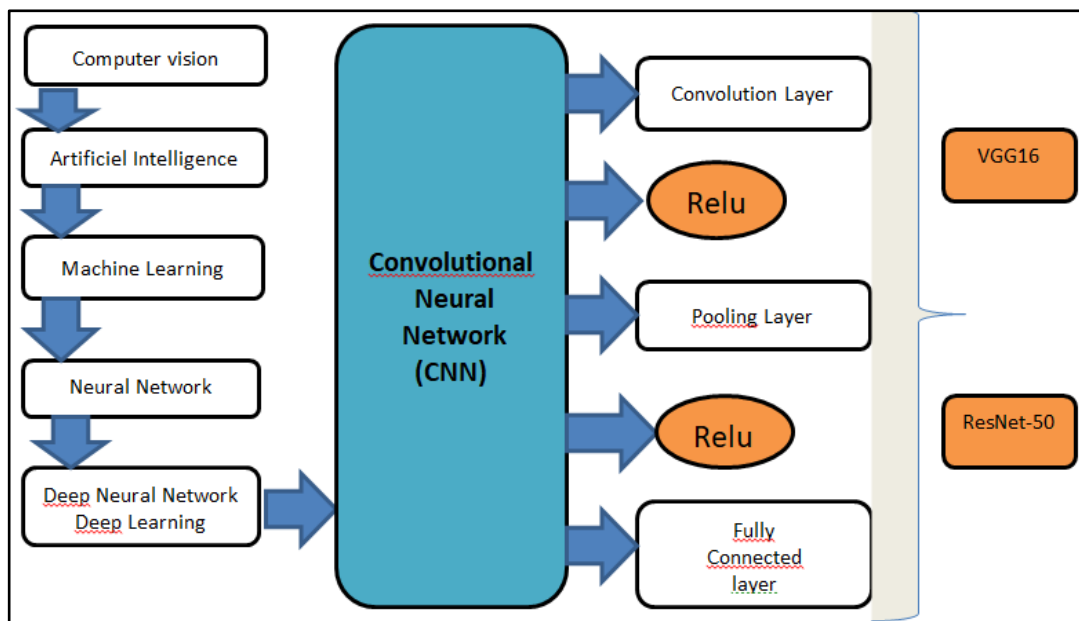


FIG-4.1 : Architecture générale du système.

4.2.1 Convolution Layer :

C'est la couche la plus importante et le cœur des éléments constitutifs du réseau convolutif, et c'est aussi elle qui effectue le plus de calculs lourds, car les caractéristiques de l'image sont extraites dans cette couche. Les couches de convolution sont essentiellement spécialisées dans l'extraction de caractéristiques de bas niveau tels que les bords, il va normalement en prenant un filtre d'une taille spécifiée (une règle de pouce est 3x3 ou 5x5... etc.) et faites défiler l'image de haut en bas à droite. Pour chaque point de l'écranimage, la valeur est calculée sur la base du filtre, nous les combinons en une seule matrice. Lorsque construire le réseau, nous spécifions aléatoirement des valeurs pour les poids, qui sont en continu mise à jour à mesure que le réseau est formé (rétropropagation) le processus va.

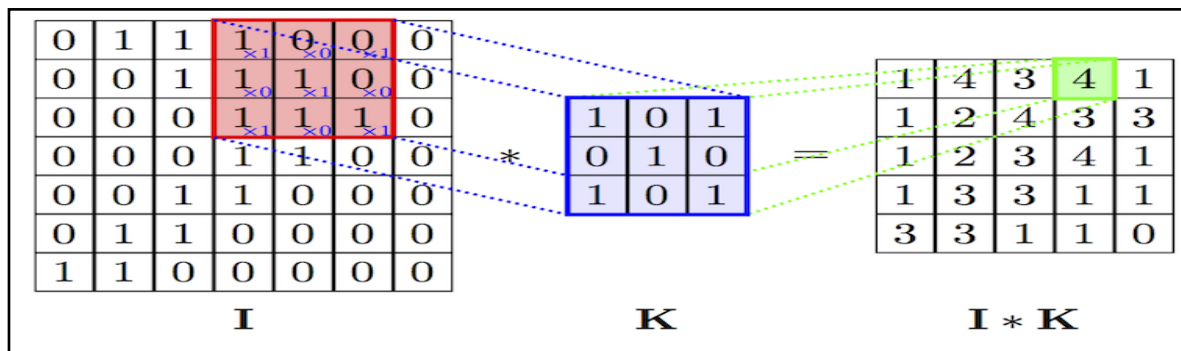


FIG-4.2:Exemple de principe du filtre convolutif [Web9].

4.2.2 Pooling Layer:

Est utilisée pour réduire le volume spatial de l'image d'entrée après la convolution. Sa fonction est de réduire progressivement la taille spatiale de la représentation pour réduire le nombre de paramètres et de calculs dans le réseau, et donc de contrôler également l'overfitting.

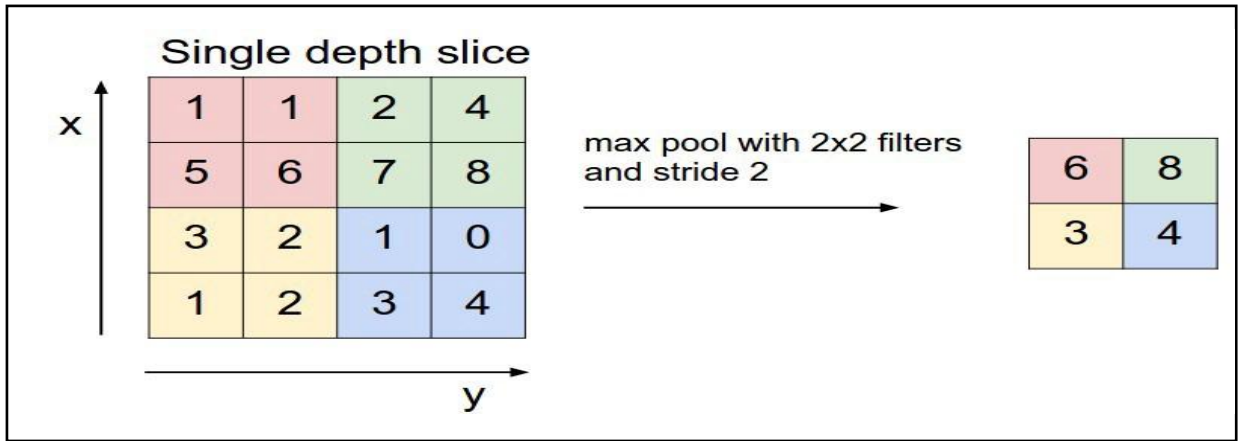


FIG-4.3:Exemple de principe du Pooling[Web20].

4.2.3 Fully-connected Layer (FC Layer):

Après plusieurs couches de convolution et de max-pooling le raisonnement de haut niveau dans le réseau neuronal se fait via des couches entièrement connectée. Il connecte les neurones d'une couche aux neurones d'une autre couche. Il est utilisé pour classer les images entre différentes catégories par formation [7].

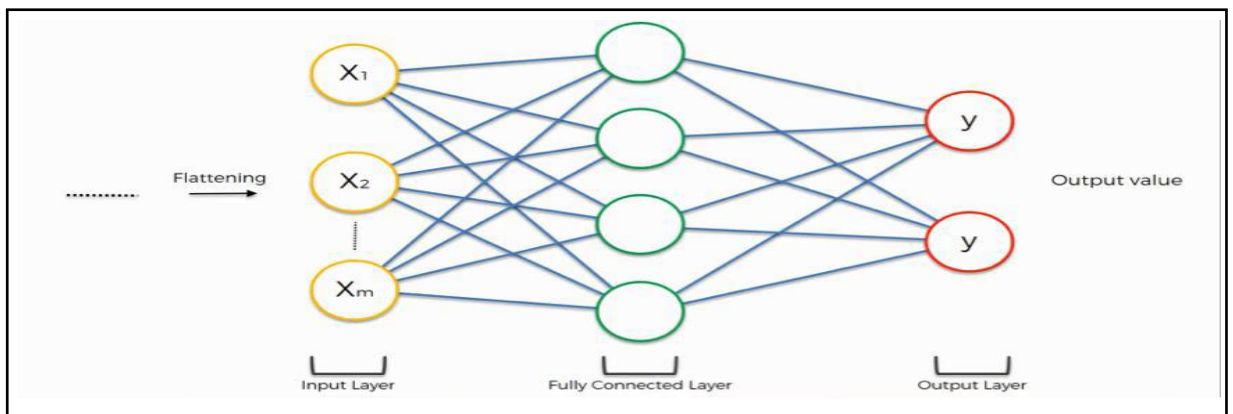


FIG-4.4:Principe de la couche entièrement connectée (FC) [8].

4.3 Fonction d'activation :

4.3.1 Fonction ReLU (Rectified Linear Unit):

ReLU est une fonction réelle non-linéaire définie par $\text{ReLU}(x) = \max(0, x)$. La couche de correction ReLU remplace donc toutes les valeurs négatives reçues en entrées par des zéros. Elle joue le rôle de fonction d'activation[9].

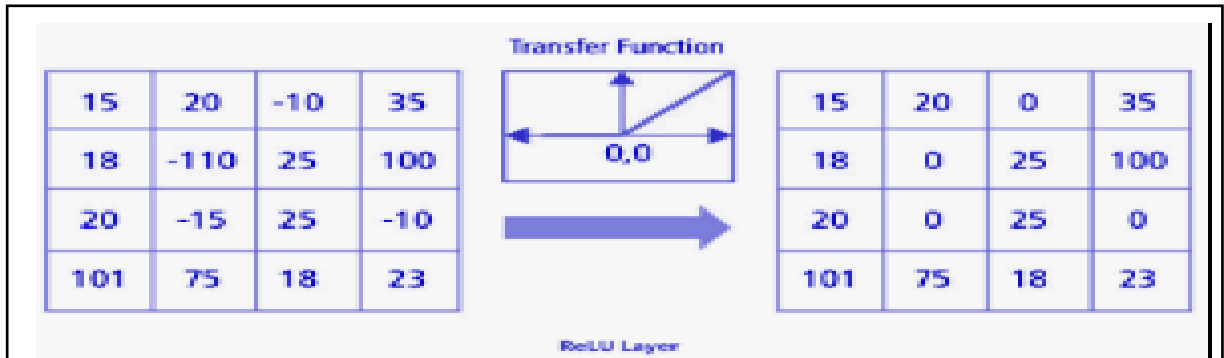


FIG-4.5 : Fonction ReLU [9].

4.3.2 Fonction Sigmoid :

Une fonction sigmoïde «FIG-4.6 » est définie comme suit :

$$f(x) = \frac{1}{1 + e^{(-x)}}$$

Pour $x \in \mathbb{R}$. Elle rend la sortie entre 0 et 1, comme le montre l'illustration suivante. La sortie non linéaire (en forme de S comme indiqué) améliore très bien le processus d'apprentissage, car elle ressemble beaucoup au principe suivant - faible influence : sortie faible et influence élevée : sortie élevée - et limite également la sortie dans une fourchette de 0 à 1.

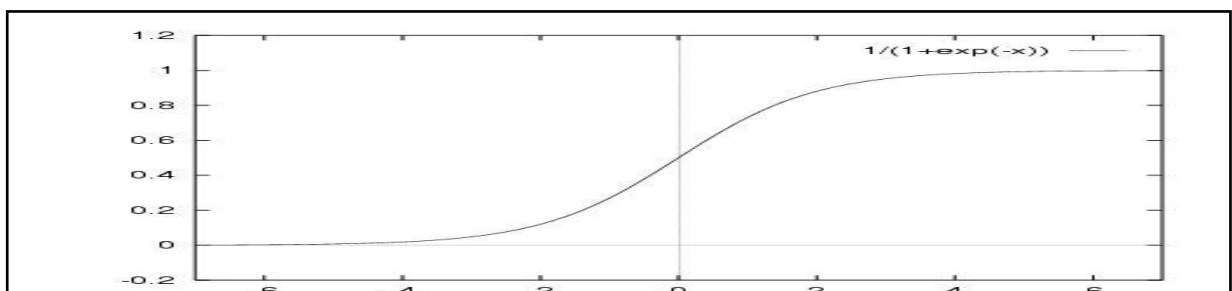


FIG-4.6 : Fonction sigmoïde[Web11].

4.3.3 Fonction Softmax :

La première chose qui vient à l'esprit en matière de classification binaire est l'algorithme de régression logistique. L'algorithme de régression logistique est un algorithme d'apprentissage automatique largement utilisé dans divers domaines. L'étape la plus importante consiste à mapper le produit du champ de nombres réels à travers le modèle linéaire à l'espace de nombres réels effectif $[0, 1]$ qui représente la distribution de probabilité, et la fonction sigmoïde a une telle fonction.

4.3.4 Dropout :

Dropout est une nouvelle méthode de régularisation qui a permis d'améliorer considérablement les résultats et la performance des réseaux de neurones. Cette méthode consiste à éteindre et rallumer les neurones aléatoirement, dans le cadre d'entraînements successifs. Une fois les séries d'entraînements terminées, on rallume tous les neurones et on utilise le réseau comme d'habitude. Cette technique a montré non seulement un gain dans la vitesse d'apprentissage, mais en déconnectant les neurones, on a aussi limité des effets marginaux, rendant le réseau plus robuste et capable de mieux généraliser les concepts appris[Web15].

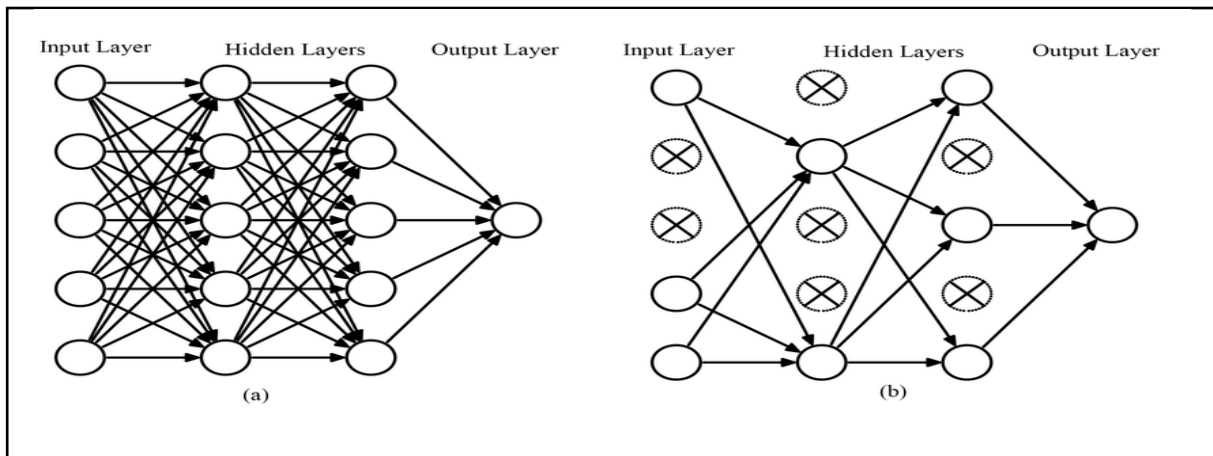


FIG-4.7 :Dropout [Web14].

4.4 Quelques Modèles de CNN :

4.4.1 VGG16 :

VGG16 est un modèle de réseau neuronal convolutif proposé par K. Simonyan et A. Zisserman de l'Université d'Oxford dans l'article « VeryDeepConvolutional Networks » pour la reconnaissance d'images à grande échelle ». Le modèle atteint 92,7 % de précision de test ImageNet, qui est un ensemble de données de plus de 14 millions d'images appartenant à 1000 classes. VGG16 est un réseau neuronal convolutif qui est formé sur plus d'un million d'images de la base de données ImageNet. Le réseau est profond de 16 couches et peut classer les images en 1000 catégories d'objets, telles que le clavier, la souris, le crayon et de nombreux animaux. Le réseau a une taille d'entrée d'image [Web12].

4.4.1.1 Architecture:

VGG16 est une architecture de réseau neuronal à convolution (CNN) qui a été utilisée pour remporter le concours ILSVR (Imagenet) en 2014. Elle est considérée comme l'une des meilleures architectures de modèle de vision jusqu'à ce jour. La chose la plus unique à propos de VGG16 est qu'au lieu d'avoir un grand nombre d'hyperparamètres, ils se sont concentrés sur des couches de convolution de filtre 3x3 avec une foulée 1 et ont toujours utilisé le même rembourrage et la même couche maxpool de filtre 2x2 de foulée 2. Il suit cet arrangement de couches de convolution et de pool maximum de manière cohérente dans toute l'architecture. Au final, il a 2 FC (couches entièrement connectées) suivies d'un softmax pour la sortie. Le 16 dans VGG16 fait référence à 16 couches qui ont des poids. Ce réseau est un réseau assez vaste et compte environ 138 millions de paramètres (environ)[Web12].

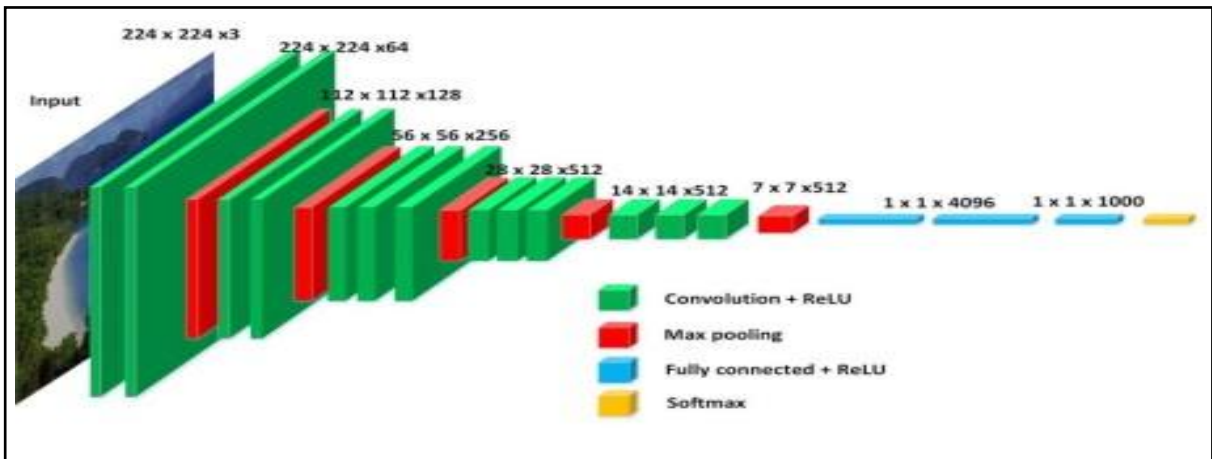


FIG-4.8:VGG16 Architecture [Web13].

4.4.2 RESNET 50 :

ResNet-50 est un réseau neuronal convolutif de 50 couches de profondeur. ResNet, abréviation de Residual Networks, est un réseau de neurones classique utilisé comme colonne vertébrale pour de nombreuses tâches de vision par ordinateur. La percée fondamentale avec ResNet a été qu'il nous a permis de former des réseaux de neurones extrêmement profonds avec plus de 150 couches. Il s'agit d'un réseau de neurones innovant qui a été introduit pour la première fois par Kaiming He, Xiangyu Zhang, Shaoqing Ren et Jian Sun dans leur article de recherche sur la vision par ordinateur de 2015 intitulé "Apprentissage résiduel profond pour la reconnaissance d'images". Les réseaux de neurones convolutifs présentent un inconvénient majeur : le « problème de gradient de fuite ». Pendant la rétropropagation, la valeur du gradient diminue de manière significative, donc pratiquement aucun changement ne vient aux poids. Pour surmonter cela, ResNet est utilisé. Il utilise "SKIP CONNECTION". [Web32]

4.4.2.1 Architecture:

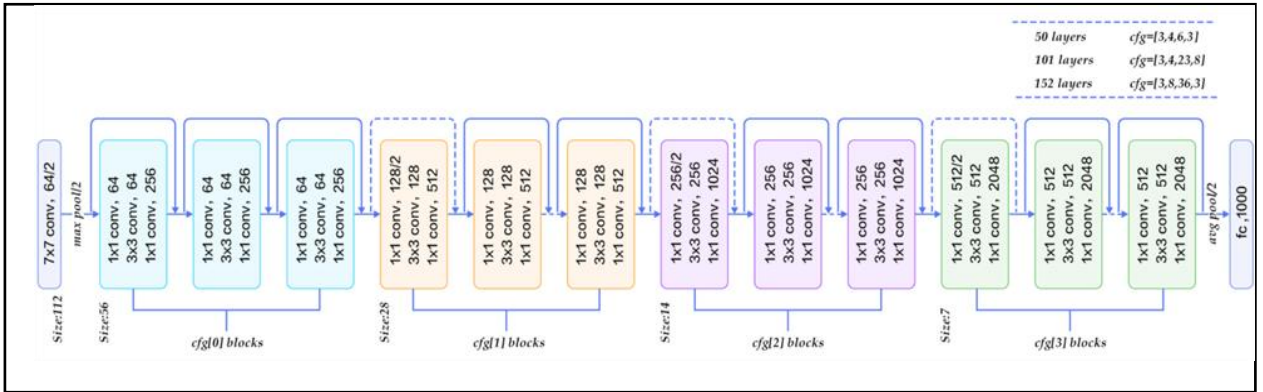


FIG-4.9:RESNET 50 Architecture [Web32].

4.5 Conclusion:

Dans ce chapitre, nous sommes concentrés sur l'un des algorithmes d'apprentissage en profondeur les plus puissants, le Deep Convolutional Neural Network (CNN), et avons expliqué ses couches et certains modèles, et cet algorithme vise à analyser des images visuelles.

5

Conception et Implémentation

5.1 Introduction :

Après avoir expliqué dans les chapitres précédents tout ce qui touche à la théorie du deeplearning et de la machine learning, dans ce chapitre nous commençons par la présentation des ressources, du langage et de l'environnement de développement utilisés, puis les étapes de réalisation du modèle et terminons par les tests effectués et les résultats obtenus.

5.2 Architecture de notre modèle :

Le modèle proposé pour le diagnostic automatisé des cas de COVID-19 est illustré à la FIG-5.1. Le modèle vise à classer une image radiographique pulmonaire donnée dans la catégorie normale ou COVID-19 qui comprend deux étapes vitales : le prétraitement (normalisation et augmentation) et classification à l'aide d'architectures CNN pré-entraînées. La description de chaque étape est détaillée dans les sections suivantes.

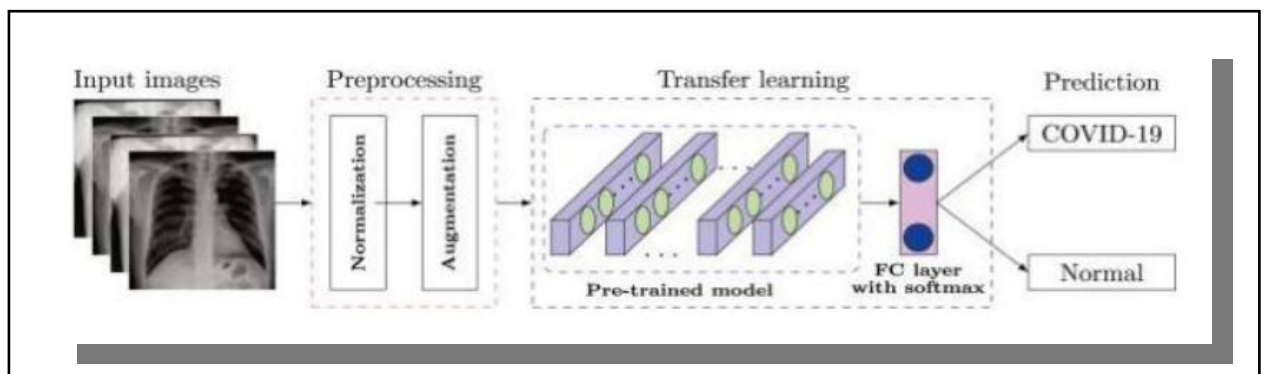


FIG-5.1: Pour une méthode de diagnostic automatisée proposée pour COVID-

19.

5.2.1 Prétraitement :

Les modèles DL ont besoin de données pour être entraînés. Plus nous avons de données, plus le modèle est performant car il sera capable de capturer plus de comportements dans le processus d'apprentissage :

5.2.1.1 Normalisation :

La normalisation des données est une étape essentielle et est généralement utilisée pour maintenir stabilité dans les architectures de CNN. Il fournit un apprentissage rapide pour elle. Par conséquent, dans cette étude, les valeurs de pixels des images d'entrée ont été normalisées entre la plage 0-1. Les images utilisées dans les ensembles de données étudiés sont des images en niveaux de gris et le redimensionnement a été réalisé en multipliant $1/255$ par les valeurs de pixel.

5.2.1.2 Augmentation :

Afin de générer un ensemble de données plus important, nous avons utilisé des techniques d'augmentation de données avec trois stratégies d'augmentation différentes : rotation aléatoire de -20% à 20%, bruit aléatoire et retournement horizontal.

5.3 Préparation de l'expérience :

5.3.1 Exécution :

Nous avons utilisé un GPU Google Colab (Tesla K80 12 Go GDDR5 VRAM), Python 3.7 et TensorFlow 2.2.0. Pour implémenter CNN, la bibliothèque d'apprentissage en profondeur TensorFlow 2.2.0 est utilisée, et les procédures de formation et de test sont effectuées sur la plate-forme Google Colab. Afin de construire le modèle, nous devons importer l'ensemble de bureaux suivant : Keras (Il est conçu pour permettre une expérimentation rapide des réseaux de neurones profonds) et Numpy (pour travailler avec des tableaux NumPy offre des fonctions mathématiques complètes), Matplotlib (complète pour créer des visualisations statiques, anime et interactives en Python).

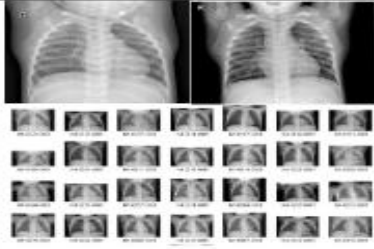
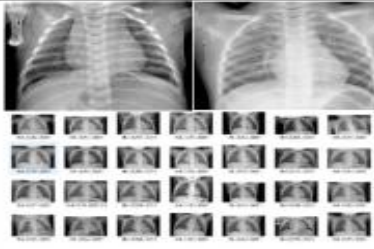
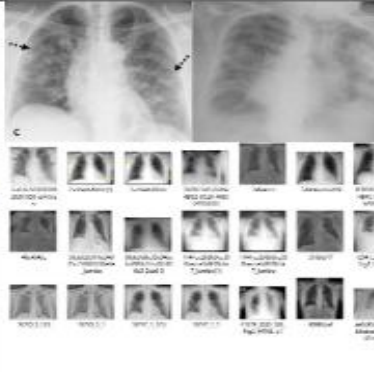

5.3.2 Base de données :

Alors que le COVID-19 attaque les cellules épithéliales qui tapissent nos voies respiratoires, nous utiliserons une radiographie pulmonaire pour analyser la santé des poumons d'un patient. Cet ensemble de données agrégées n'est pas destiné à revendiquer mais à rechercher la capacité de diagnostic de tout modèle d'apprentissage en profondeur. Dans ce travail, nous avons utilisé deux bases de données. L'ensemble de données comprenait un total de 6056 radiographies pulmonaires :

- Des images radiographiques de personnes infectées par le virus Covid-19 ont été obtenues auprès du Github du Dr. Mais nous n'avons besoin que des radiographies du virus Covid-19 et pour cela il existe un fichier (csv) contenant toutes les données. Nous allons travailler là-dessus pour extraire des images de personnes infectées par le virus Covid-19 et à partir de là le nombre d'images est de 930 où les images sont différentes en rayons x et il y a des coupes avant et arrière ainsi que transversales. Pour équilibrer ici avec la base de données non infectée, nous choisissons de travailler sur PA (Antérieur-postérieur)[Web22].
 - L'ensemble de données privé a été collecté à partir du référentiel Kaggle, qui contient des examens radiographiques pulmonaires pour le Covid-19 infecté, normal et de pneumonie. Cet ensemble de données est divisé en 5467 groupes de formation, de validation 965 normaux et de pneumonie. Nous avons travaillé sur cas normaux seulement [Web31].
 - Nous avons une base de données composée de patients et de non-patients. Nous avons obtenu les deux bases de données de nos sites Kaggle et Github. Grâce à ces deux sites, nous avons compilé une base de données et créé notre propre base de données.
- Ensuite, nous les avons organisés dans les dossiers Covid-19+ et Covid-19-, Après avoir obtenu une base de données, nous allons travailler à diviser la base de données train en base de données train.

[Web31]. https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia?fbclid=IwAR24z60oO4l3TdqWAUfEADz7mF-y4tP2XqE6MRWp1JaCB86ah8z_zqROv4o

- la base de données de test à l'aide du programme Spyde comme suit, où le fichier train contient des données pour les patients et les non-patients, ainsi que pour le fichier de test.
- Où nous l'avons divisé comme suit : 70 % de train et 30 % de test.
- Ensuite, nous l'avons téléchargé sur Drive afin de l'invoquer et de travailler dessus dans Colabe.

	Train	Tests
Healthy Person		
Covid-19 Infected Patient		

TAB. 5 .1 : distribution d'ensembles de données.

5.4 Extraction de fonctionnalités :

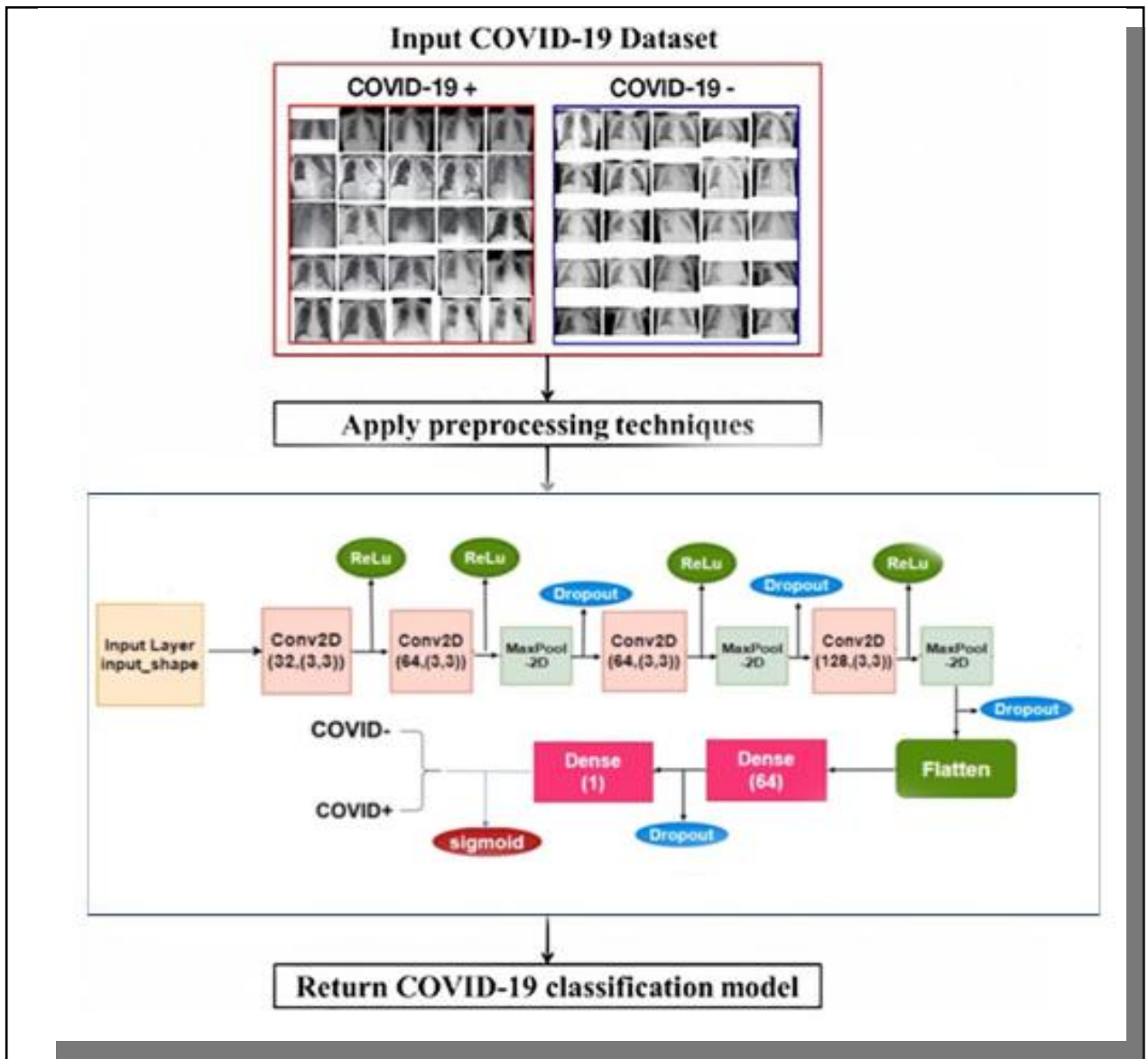


FIG-5.2 : Schéma fonctionnel du processus d'apprentissage du modèle de Classification COVID-19 basé sur CNN.

CNN est l'architecture la plus couramment utilisée dans laquelle l'image est insérée dans la couche d'entrée, qui se propage à travers plusieurs couches cachées et produit finalement la probabilité de prédiction à travers la couche de sortie. Ces couches cachées consistent en une série de couches convolutionnelles suivies de couches d'assemblage qui tentent de découvrir et d'apprendre des caractéristiques et des modèles très complexes appartenant à des images d'une classe particulière. C'est la raison du formidable succès des modèles CNN dans le domaine de la vision par

ordinateur et de ses diverses applications. Bien que CNN fonctionne très bien dans la tâche de classification des images, il existe certaines limites à cela. Le modèle CNN proposé considère des images de taille 224x224 dans la couche d'entrée. La couche d'entrée est suivie d'une couche convolutive, d'une couche de normalisation par lots et d'une couche Maxpooling. Cet ensemble de couches est répété quatre fois avec un nombre variable de filtres ayant une taille de (3 x 3). Tous les neurones de ces couches ont le ReLU (RectifiedLinear Unit) comme fonction d'activation. La couche de convolution finale se compose de 8 filtres de taille (3 x 3) suivis de couches entièrement connectées ayant une activation ReLU. Pour toutes ces couches, la foulée de taille (1x1) est utilisée dans les couches convolutives et la foulée de taille (2x2) est utilisée dans les couches Maxpooling. À partir de la dernière couche convolutive, un total de 800 les caractéristiques sont générées à partir de chaque échantillon d'image. Paramètres totaux de ce l'architecture sont 66 079, où les paramètres entraînaibles sont 65 727 et les 352 autres sont des paramètres non entraînaibles.

5.5 Résultat de l'expérience :

5.5.1 Code d'explication :

- Cette figure montre le code pour créer un modèle en parcourant plusieurs couches jusqu'à la fin

```
# Build the model
x = Conv2D(32, (3,3) , padding = 'same', activation='relu', name = 'layer_1') (img_input)
x = Conv2D(64, (3,3) , padding = 'same', activation='relu', name = 'layer_2') (x)
x = MaxPool2D((2,2), strides=(2,2), name = 'layer_3') (x)
x = Dropout(0.25)(x)

x = Conv2D(64, (3,3) , padding = 'same', activation='relu', name = 'layer_4') (x)
x = MaxPool2D((2,2), strides=(2,2), name = 'layer_5') (x)
x = Dropout(0.25)(x)

x = Conv2D(128, (3,3) , padding = 'same', activation='relu', name = 'layer_6') (x)
x = MaxPool2D((2,2), strides=(2,2), name = 'layer_7') (x)
x = Dropout(0.25)(x)

x = Flatten(name='layer_8')(x)
x = Dense(64, name = 'layer_9')(x)
x = Dropout(0.5)(x)

x = Dense(2, activation='sigmoid', name = 'predections')(x)
```

FIG-5.3 : Modèle Resnet pour la segmentation.

- La figure montre comment créer le formulaire et lui donner un nom :

```
# generate the model
model = Model(inputs = img_input, outputs =x, name = 'CNN_COVID_19' )
```

FIG-5.4 : Génération du mode.

- Imprimer la structure du réseau

```
# pint network structure
model.summary()
```

FIG-5.5:Structure du réseau d'impression.

- démarrer Train / Test

```
# Start trian/Test
batch_size = 32
hist = model.fit(traindata,
                 steps_per_epoch = traindata.samples//batch_size,
                 validation_data = testndata,
                 validation_steps = testndata.samples//batch_size,
                 epochs =100
                )
```

FIG-5.6 : Code de démarrage du train et du test.

```
plt.plot(hist.history['loss'], label = 'train')
plt.plot(hist.history['val_loss'], label = 'val')
plt.title('CNN_In_8_Steps : Loss & Validation Loss')
plt.legend()
plt.show()

plt.plot(hist.history['accuracy'], label = 'train')
plt.plot(hist.history['val_accuracy'], label = 'val')
plt.title('CNN_In_8_Steps : Accuracy & Validation Accuracy')
plt.legend()
plt.show()
```

FIG-5.7 : Code de courbe.

```

# Confusion Matrix & Pres & Recall & F1-Score

target_names = ['COVID+', 'COVID-']
label_names = [0,1]

Y_pred = model.predict_generator(testndata)
y_pred = np.argmax(Y_pred , axis = 1)

cm = confusion_matrix(testndata.classes, y_pred, labels = label_names)

print('Confusion Matrix')
print(confusion_matrix(testndata.classes, y_pred))

print('classification_Report')
print(classification_report(testndata.classes, y_pred, target_names=target_names))

disp = ConfusionMatrixDisplay(confusion_matrix= cm, display_labels=target_names)
disp = disp.plot(cmap=plt.cm.Blues, values_format = 'g')
plt.show()

```

FIG-5.8 : Code de Matrice.

5.5.2 Résultats et discussion :

Dans cette partie nous nous intéressons aux résultats d'exécution de l'approche développée. Le résultat de la structure du réseau d'impression illustré à la FIG-5.10. Les paramètres de chaque couche décrit dans cette partie a bien détaillé.

Layer (type)	Output Shape	Param #
img_input (InputLayer)	(None, 224, 224, 3)	0
layer_1 (Conv2D)	(None, 224, 224, 32)	896
layer_2 (Conv2D)	(None, 224, 224, 64)	18496
layer_3 (MaxPooling2D)	(None, 112, 112, 64)	0
dropout_1 (Dropout)	(None, 112, 112, 64)	0
layer_4 (Conv2D)	(None, 112, 112, 64)	36928
layer_5 (MaxPooling2D)	(None, 56, 56, 64)	0
dropout_1 (Dropout)	(None, 56, 56, 64)	0
layer_6 (Conv2D)	(None, 56, 56, 128)	73856
layer_7 (MaxPooling2D)	(None, 28, 28, 128)	0
dropout_2 (Dropout)	(None, 28, 28, 128)	0
layer_8 (Flatten)	(None, 100352)	0
layer_9 (Dense)	(None, 64)	6422592
dropout_3 (Dropout)	(None, 64)	0
predictions (Dense)	(None, 2)	130

Total params: 6,552,898
 Trainable params: 6,552,898
 Non-trainable params: 0

FIG-5.9 : Le résultat de la structure du réseau d'impression.

5.5.3 Fonction de perte et fonction de précision :

- Représente la fonction de perte d'entraînement sur l'époque

```

10/10 [=====] - 14s 1s/step - loss: 0.0331 - accuracy: 0.9900 - val_loss: 0.1272 - val_accuracy: 0.9479
Epoch 86/100
10/10 [=====] - 14s 1s/step - loss: 0.0226 - accuracy: 0.9900 - val_loss: 0.1979 - val_accuracy: 0.9375
Epoch 87/100
10/10 [=====] - 14s 1s/step - loss: 0.0632 - accuracy: 0.9866 - val_loss: 0.0874 - val_accuracy: 0.9688
Epoch 88/100
10/10 [=====] - 14s 1s/step - loss: 0.0333 - accuracy: 0.9937 - val_loss: 0.0823 - val_accuracy: 0.9583
Epoch 89/100
10/10 [=====] - 14s 1s/step - loss: 0.0197 - accuracy: 0.9933 - val_loss: 0.0922 - val_accuracy: 0.9688
Epoch 90/100
10/10 [=====] - 15s 2s/step - loss: 0.0233 - accuracy: 0.9866 - val_loss: 0.0888 - val_accuracy: 0.9896
Epoch 91/100
10/10 [=====] - 14s 1s/step - loss: 0.0372 - accuracy: 0.9833 - val_loss: 0.0940 - val_accuracy: 0.9583
Epoch 92/100
10/10 [=====] - 14s 1s/step - loss: 0.0273 - accuracy: 0.9900 - val_loss: 0.1184 - val_accuracy: 0.9792
Epoch 93/100
10/10 [=====] - 14s 1s/step - loss: 0.0431 - accuracy: 0.9933 - val_loss: 0.2711 - val_accuracy: 0.9167
Epoch 94/100
10/10 [=====] - 14s 1s/step - loss: 0.0460 - accuracy: 0.9967 - val_loss: 0.0813 - val_accuracy: 0.9583
Epoch 95/100
10/10 [=====] - 14s 1s/step - loss: 0.0250 - accuracy: 0.9900 - val_loss: 0.1603 - val_accuracy: 0.9583
Epoch 96/100
10/10 [=====] - 14s 1s/step - loss: 0.0243 - accuracy: 0.9933 - val_loss: 0.1239 - val_accuracy: 0.9688
Epoch 97/100
10/10 [=====] - 14s 1s/step - loss: 0.0117 - accuracy: 0.9967 - val_loss: 0.1717 - val_accuracy: 0.9583
Epoch 98/100
10/10 [=====] - 14s 1s/step - loss: 0.0580 - accuracy: 0.9866 - val_loss: 0.0532 - val_accuracy: 0.9896
Epoch 99/100
10/10 [=====] - 14s 2s/step - loss: 0.0251 - accuracy: 0.9933 - val_loss: 0.1153 - val_accuracy: 0.9271
Epoch 100/100
10/10 [=====] - 14s 1s/step - loss: 0.0366 - accuracy: 0.9866 - val_loss: 0.1581 - val_accuracy: 0.9792
    
```

FIG-5.10 :Résultats de l’entraînement du modèle avec les images.

Après l’analyse des résultats obtenus, on constate les remarques suivantes :

- Si nous faisons une comparaison de FIG-5.11 et FIG-5.12 nous constatons qu’entre ces 100 itérations nous constatons que le taux d’erreur baisse tant dis que notre précision augmente.
- La précision de l’apprentissage et validation augmente avec le nombre d’époque. Cela signifie que notre modèle a été bien entrainer et réponds d’ailleurs à la définition du réseau de neurones qui disait que plus le réseau de neurones est profond meilleur sont ses performance.

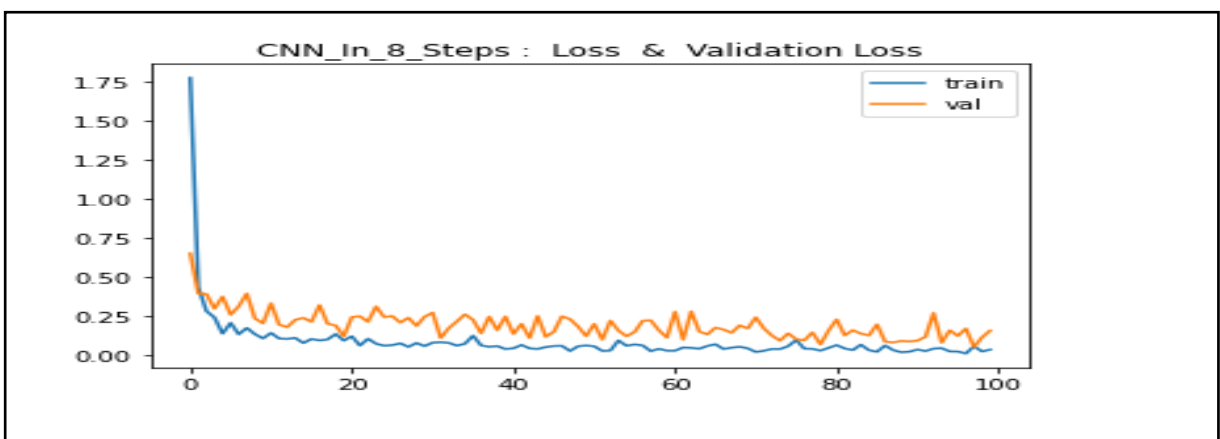


FIG-5.11 : courbe de loss et validationloss.

- Représente la fonction de précision d’entraînement sur l’époque.

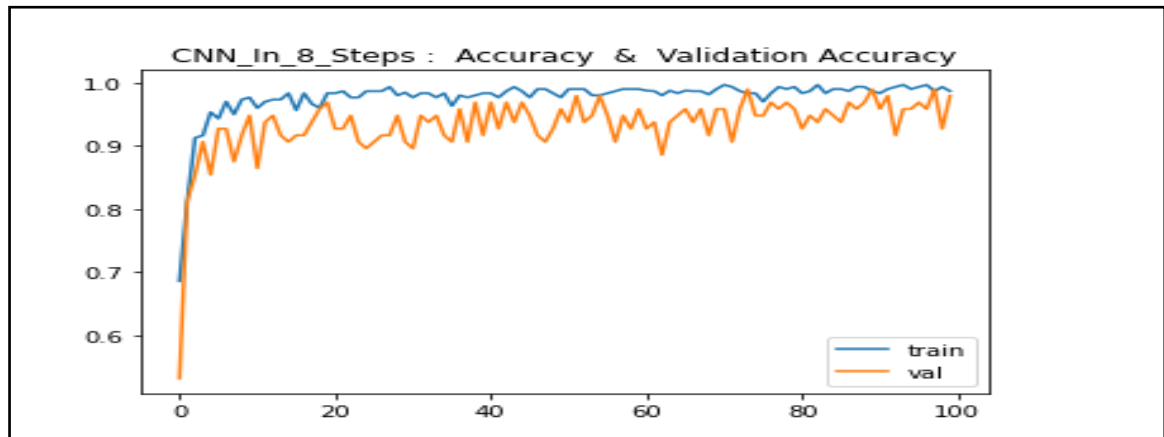


FIG-5.12 : courbe d'accuracy et validation accuracy.

5.5.4 La matrice de confusion (MC) :

La matrice de confusion est la mesure la plus simple et intuitive utilisée pour trouver l'exactitude du modèle. La MC est un tableau comportant deux dimensions (« prévu » et « réel ») et des groupes de « classes » dans les deux dimensions. Nos prévisions sont des lignes et les classifications réelles sont des colonnes. MC dans en soi n'est pas une mesure du rendement en tant que telle, mais presque toutes les mesures du rendement sont basé sur la matrice de confusion et les nombres à l'intérieur [23].

	P	N
P	True Positives(TP)	False Positives(FP)
N	False Négatives(FN)	True Négatives(TN)

TAB. 5 .2 :La matrice de confusion.

5.5.4.1 Termes de la matrice de confusion :

- ❖ Matrice de confusion Te False Négatifs (FN) : Les faux négatifs sont les cas où la classe réelle du point de données était 1 (Vrai) et la valeur prédite est 0 (Faux). Faux est parce que le modèle a prédit incorrectement et négativement parce que la classe prédite était négative (0) [22].
- ❖ Vrais positifs (TP) : Les vrais positifs sont les cas où la classe réelle du point de données était 1 (Vrai) et la prévision est également 1(Vrai).

❖ Faux positifs (FP) : Les faux positifs sont les cas où la classe réelle du point de données était 0 (Faux) et la valeur prédite est 1 (Vrai). Faux est parce que le modèle a prédit incorrectement et positivement parce que la classe prédite était positive (1).

❖ Vrais négatifs (TN) : Les vrais négatifs sont les cas où la classe réelle du point de données était 0 (Faux) et la prévision est également 0 (Faux).

5.5.4.2 Mesures d'évaluation du classificateur :

➤ Accuracy :

La précision des problèmes de classification est le nombre de prédictions correctes modèle sur toutes sortes de prédictions faites.

- Quand utiliser Accuracy ?

L'exactitude est une bonne mesure lorsque les classes de variables cibles dans les données sont presque équilibrées.

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

➤ Précision :

La précision est une mesure basée sur la MC qui nous dit quelle proportion de patients nous diagnostiqué comme ayant corona et effectivement eu corona. Les positifs prévus (Les gens prédit comme coronavirus sont TP et FP) et les personnes ayant réellement une corona sont TP.

- Quand ne pas utiliser la précision ?

L'exactitude ne doit JAMAIS être utilisée comme mesure lorsque les classes de variables cibles dans les données sont une majorité d'une class.

$$Precision = \frac{TP}{TP + FP}$$

➤ **Le Rappel :**

Le rappel est une mesure qui nous indique la proportion de patients qui ont déjà covid-19 qui sont diagnostiqués par l'algorithme comme ayant covid-19. Les vrais positifs (les personnes atteintes de corona sont TP et FN) et les personnes diagnostiquées avec covid-19 sont TP. (Remarque : FN est inclus car la personne a déjà une aura même si le modèle prédit le contraire.

$$Recall = \frac{TP}{TP + FN}$$

➤ **Spécificité :**

La spécificité est une mesure qui nous indique quelle proportion de patients n'ont PAS de corona, ont été prédits par le modèle comme non-coronavirus. Les négatifs réels (Les gens en fait PAS ayant corona sont FP et TN) et les personnes diagnostiquées par nous n'ayant pas de corona sont TN. (Remarque : Le PF est inclus parce que la personne n'avait PAS réellement de corona, même si le modèle prédit La spécificité est exactement le contraire de Rappé[24].

$$Spécificité = \frac{TN}{TN + FP}$$

➤ **Dice score:**

Le coefficient de similitude de Dice, aussi connu sous le nom d'indice Sorensen-Dice ou simplement coefficient de Dice, est un outil statistique qui mesure la similitude entre deux ensembles de données. Ce L'index est devenu sans doute l'outil le plus utilisé pour la validation de l'image algorithmes de segmentation créés avec l'IA.

$$Dice = \frac{2 \times TP}{(TP + FP) + (TP + FN)}$$

➤ **The F1 score :**

La partition F1 est un nombre compris entre 0 et 1 et est la moyenne harmonique de précision et rappel[6].

$$F1 = 2 * \frac{precision * recall}{precision + recall}$$

5.5.4.3 Résultat de la matrice de confusion :

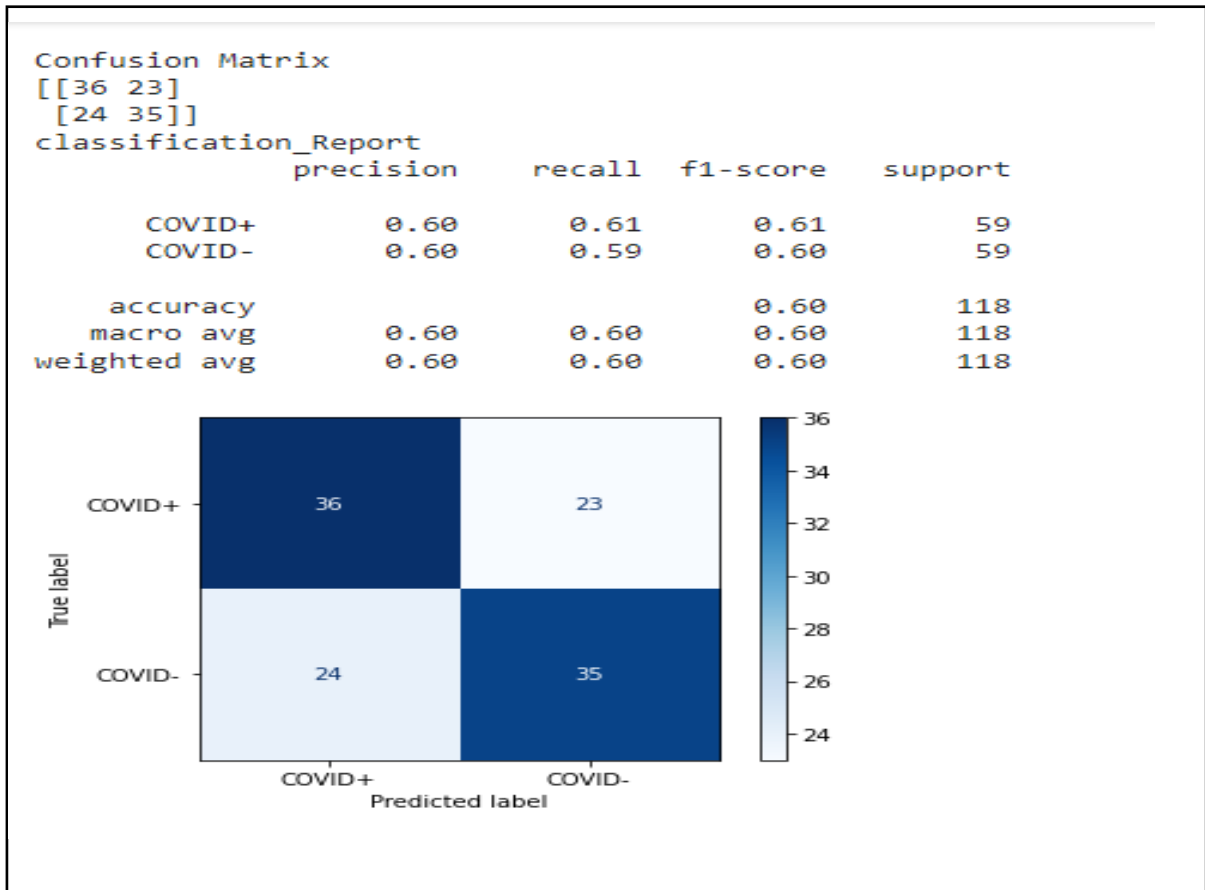


FIG-5.13: Résultat de la matrice de confusion.

5.5.5 Résultat général:

Nous avons utilisé les modèles pré-formés VGG16 et Resnet50 sur Imagenet et les avons utilisés de différentes manières, une fois sans changement, encore une fois avec juste une formation softmax et en gardant les couches précédentes figées, et encore avec la dernière formation de couche de convolution avec softmax et en gardant la précédente couches figées, la dernière fois que nous avons créé des transformateurs pour créer notre modèle et l'entraîner à partir de zéro, puis comparé les résultats extraits, et ils étaient les suivants :

	Forme	Parametre	Formation temps
VGG16	Softmax	4,098	44 m
	Last layer +Softmax	2,363906	86 m
Resnet50	Softmax	8,194	65 m
Nos modèles	Last layer +Softmax	6,552898	215 m

TAB. 5 .3 :Résultats extraits.

❖ Le tableau suivant résume les résultats de la précision de détection pour chacun d'eux :

	Formation	Accuracy
VGG16	Softmax	70.25%
	Last layer +Softmax	86.57%
Resnet50	Softmax	67.20%
	Last layer +Softmax	83.83%
Nos modèles	CNN-Covid-19	98.66%

TAB. 5 .4 :Accuracy.

Nous avons remarqué que le réglage fin VGG16 (modifié dans la dernière couche +soft-max) et le réglage fin Resnet50 (ajusté dans la dernière couche + soft-max) ont donné un résultat acceptable. Et par vertu de l'homme toujours à la recherche du meilleur, nous avons donc décidé d'essayer de créer un modèle et de former toutes ses couches et tous ses poids, et il nous a donné de meilleurs résultats comme le montre le tableau précédent.

5.5.6 Discussion :

L'étude des images radiographiques thoraciques pour la prédiction précise de l'infection au COVID-19 a suscité un grand intérêt depuis la publication de l'ensemble de données, après quoi plusieurs tentatives ont été faites pour développer un modèle de diagnostic fiable à l'aide de techniques DL. Il a été largement utilisé avec les réseaux basés sur CNN. Cependant, la plupart des méthodes précédentes ont été évaluées en utilisant une quantité limitée de données et, dans certains cas, les données sont également déséquilibrées. Les résultats obtenus ont montré que notre algorithme proposé est efficace, bien que notre solution soit efficace, elle se trompe parfois, nous

avons donc essayé plusieurs façons d'atteindre le dernier résultat. En comparant les modèles VGG16 et ResNet50 et le nôtre, notre modèle les a surpassés autant que possible car toutes leurs couches ont été entraînées et pondérées, et notre résultat était acceptable. Les modèles CNN ont donné de bons résultats (83,73% pour VGG16 et 73,18% pour ResNet50), cependant, nous voulions augmenter beaucoup plus la précision. C'est pourquoi nous avons créé un modèle, formé toutes ses couches et l'avons pondéré pour nous donner un bon résultat. En fin de compte, notre solution proposée a atteint notre objectif souhaité en détectant la maladie infectée avec une précision de 98,66%.

5.6 Conclusion :

Nous avons présenté dans ce chapitre une approche de classification basée sur les réseaux de neurones convolutionnels, nous avons présenté notre application en donnant quelques captures d'écran qui expliquent le déroulement et le fonctionnement de notre travail, ainsi que les résultats obtenus en termes de précision et d'erreur.

Conclusion générale & perspectives

L'Organisation mondiale de la santé a déclaré Covid-19 une pandémie mondiale le 11 mars 2020, le virus se propage entre les personnes par contact physique et les gouttelettes respiratoires de la toux ou des éternuements, et afin de limiter sa propagation, nous avons besoin d'une technologie plus rapide et relativement peu coûteuse pour détecter Covid19.

La technologie a pénétré notre quotidien avec beaucoup de succès, notamment dans l'intelligence artificielle, l'apprentissage automatique et l'apprentissage en profondeur. En raison des succès obtenus dans le domaine médical en utilisant l'intelligence artificielle dans divers domaines médicaux, nous avons travaillé à exploiter des images médicales aux caractéristiques différentes et à les relier au domaine de l'intelligence artificielle. Où l'apprentissage en profondeur est un moyen efficace d'offrir des solutions prometteuses basées sur un ensemble de données pour aider l'humanité à faire face au COVID-19.

Dans cette mémoire, nous proposons une méthode de diagnostic efficace pour identifier et caractériser les cas de COVID-19 par radiographie pulmonaire. Nous avons adopté la méthode CNN comme méthode de classification binaire pour diagnostiquer le Covid-19. Nous avons abordé les bases des réseaux de neurones en général et des réseaux de neurones convolutifs en particulier. Nous avons introduit un CNN en introduisant différents types de couches utilisées pour la classification des images en raison de la grande disponibilité de l'ensemble de données.

Notre méthode proposée a atteint une précision globale de 98.66%. Par rapport aux études précédentes, notre méthode présente certains avantages. On peut voir que notre modèle est d'une importance de référence pour le diagnostic de COVID-19 dans les travaux futurs. Nous prévoyons d'élargir l'ensemble de données et de présenter les images de tomographie en même temps et d'essayer d'améliorer les structures de réseau les plus efficaces pour améliorer encore les performances de

classification. Nous essaierons également d'appliquer cette méthode aux dispositifs médicaux ou de l'étendre à d'autres tâches médicales pour aider à dépister le COVID-19 et à diagnostiquer d'autres maladies.

Bibliographie

- [1] Boughaba Mohammed et Boukhris Brahim., L'apprentissage profond (Deep Learning) pour la classification et la recherche d'images par le contenu . Thèse de master: UniversityKasdiMerbah , Ouargla .2016-2017
- [2] S. Candemir, S. Jaeger, K. Palaniappan, J.P. Musco, R.K. Singh, Z. Xue, A. Karargyris, S. Antani, G. Thoma, C. J. McDonald, "Lung segmentation in chestradiographusinganatomicalatlaseswithnonrigid registration," IEEE Transactions on Medical Imaging, vol.33, no.2, pp.577-590, February 2014
- [3] G. Coppini, M. Miniati, M. Paterni, S. Monti, E. M. Ferdeghini, Computer-aided diagnosis of emphysema in COPD patients: Neuralnetwork-based analysis of lung shape in digital chest radiographs, Medical Engineering & Physics, Volume 29, Issue 1, January 2007, Pages 76-86.
- [4] F.M. Carrascal, J.M. Carreira, M. Souto, P.G. Tahoces, L. Gómez, J.J. Vidal, "Automatic calculation of total lung capacity from automatically traced lung boundaries in postero-anterior and lateral digital chest radiographs" Med Phys. 1998 Jul;25(7 Pt 1):1118-31.
- [5] Mr. Mokri Mohammed Zakaria.MMZ, 2017.Classification des images avec les réseaux de neurones convolutionnels ": mémoire de projet de fin d'étude. UniversitéAbouBakrBelkaidTlemcen.
- [6] (Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, Lou Y, GaoD,YangL,He D (2020) Preliminary estimation of the basic reproduction number of novel.)
- [7] A. HABBA et O. ISHAK. Année Universitaire 2018/2019. La classification des images satellitaires par l'apprentissage profonde (deeplearning). Mémoire de fin d'étude. Université Ahmed Draia - Adrar
- [8] Z.Sellami. 7 juillet 2019. Optimisation du CNN par l'algorithme Génétique pour la Reconnaissance de Visage. Mémoire de fin d'étude. Université Mohamed Khider de Biskra.
- [9] Shen, Zuowei, Haizhao Yang, and Shijun Zhang. "Optimal approximation rate of ReLU networks in terms of width and depth." Journal de MathématiquesPuresetAppliquées 157 (2022): 101-135.

- [10] Jie Cui, Fang Li et Zheng-Li Shi, « Origin and evolution of pathogenic coronaviruses », *Nature Reviews Microbiology*, vol. 17, no 3, mars 2019, p. 181–192 (ISSN 1740-1534, DOI 10.1038/s41579-018-0118-9).
- [11] Y. Chen, « The Novel Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Directly Decimates Human Spleens and Lymph Nodes », *University of Oxford - Immunology Network*, 2020.
- [12] P. Vincent, « Modèles à noyaux à structure locale », Thèse de Phd en informatique, Université de Montréal, 2003.
- [13] R. Acharya, R. Wasserman, J. Stevens, and C. Hinojosa. *Biomedical Imaging modalities: a tutorial. Computerized Medical Imaging and Graphics. VOL.19 N° 01* Année 1995.
- [14] I. Vasilev, D. Slater, G. Spacagna, P. Roelants, and V. Zocca, *Python Deep Learning : Exploring deep learning techniques and neural network architectures with Pytorch, Keras, and Tensor-Flow*. Packt Publishing Ltd, 2019.
- [15] Nichyporuk, Brennan. "Engineering deep learning systems for robust and accurate focal pathology segmentation and detection." (2022).
- [16] S. Chabane et N. Haddache, *Segmentation d'images mammographies*. Mémoire de Master en vue d'obtenir le diplôme de master professionnel en administration et sécurité des réseaux. Université Abderrahmane Mira Béjaïa. 2017.
- [17] MEKHALDI Nadia, *Analyse des images médicales par des techniques hybrides : Application à la détection des kistes réseaux*. Diplôme magister en informatique. université de science et de la technologie d'Oran Mohamed Boudiaf. 2014..
- [18] LEFEVRE, Sébastien. *Elaboration et validation d'un outil mathématique de segmentation et de reconstruction volumique sur imagerie IRM de tumeurs cérébrales*. Mémoire pour l'obtention du diplôme d'études approfondies, 1999..
- [19] Chua, Leon O. *CNN: A paradigm for complexity*. Vol. 31. World Scientific, 1998.
- [20] https://www-master.ufr-info-p6.jussieu.fr/2006/Ajouts/Master_esj_2006_2007/IMG/pdf/IntroDefinitions.pdf
- [21] (Xu, H., Chen, W., Zhao, N., Li, Z., Bu, J., Li, Z., Liu, Y., Zhao, Y., Pei, D., Feng, Y., et al. *Unsupervised anomaly detection via variational autoencoder for seasonal kpis in web applications*. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web*, pages 187–196).

- [22] (tekalp, A. Murat, Digital video processing, Prantice Hall.)
- [23] Xu, H., Chen, W., Zhao,N.,Li,Z.,Bu,J.,Li,Z.,Liu,Y.,Zhao,Y.,Pei, D., Feng, Y., et al. Unsupervised anomaly detection via variationalautoencoder for seasonal kpis in web applications. In Proceedings of the 2018 World WideWeb Conference on World Wide Web, pages 187-196).
- [24](International World WideWebConferences Steering Committee.)

Webographie

- [Web1] Virus respiratoires émergents, y compris le COVID-19 : méthodes de *détection*, de prévention, de réponse et de contrôle) .. Open WHO. Retrieved June 13, 2022, from <https://openwho.org/courses/introduction-au-ncov>
- [Web2] iee8023/covid-chestxray-dataset: We are building an open database of COVID-19 cases with chest X-ray or CT images. (n.d.).GitHub. Retrieved June 13, 2022, from <https://github.com/iee8023/covid-chestxray-dataset>
- [Web3] Qu'est-ce que l'apprentissage automatique ? (2020, July 15). IBM. Retrieved June 13, 2022, from <https://www.ibm.com/fr-fr/cloud/learn/machine-learning>
- [Web4] Retrieved June 13, 2022, from <https://www.scribd.com/presentation/467048988/Apprentissage-supervise>
- [Web5] Apprentissage non supervisé : on vous explique ! (n.d.). DataScientest.com. Retrieved June 13, 2022, from <https://datascientest.com/apprentissage-non-supervise>
- [Web6] Retrieved June 13, 2022, from <https://pageperso.univ-lr.fr/arnaud.revel/MesPolys/SemiSupervise>
- [Web7] Retrieved June 13, 2022, from <https://io.hsoub.com/tech/118455%D8%A7%D9%84%D8%B1%D8%A4%D9%8A%D8%A9%D8%A7%D9%84%D8%AD%D8%A7%D8%B3%D9%88%D8%A8%D9%8A%D8%A9-computer-vision>
- [Web8] Retrieved June 13, 2022, from <https://i.suar.me/LG0gn>
- [Web9] Retrieved June 13, 2022, from <https://i.stack.imgur.com/mkjZo.jpg>
- [Web10] Retrieved June 13, 2022, from <https://www.educative.io/api/edpresso/shot/4775123253460992/image/60948240%2000716800>
- [Web11] Retrieved June 13, 2022, from <https://humphryscomputing.com/Notes/Neural/Bitmaps/sigmoid.gif>
- [Web12] Retrieved June 13, 2022, from <https://towardsdatascience.com/step-by-step-vgg16-implementation-in-%20keras-for-beginners-a833c686ae6c>
- [Web13] Retrieved June 13, 2022, from <https://www.researchgate.net/publication/335184836/figur>

- [Web14] Retrieved June 13, 2022, from https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.researchgate.net%2Ffigure%2FDropout-Strategy-a-A-standard-neural-network-b-Appling-dropout-to-the-neural_fig3_340700034&psig=AOvVaw0XrN1JUP6CrRGWhtbYo991&ust=1652983443673000&source=images&cd=vfe&ved=0
- [Web15] Budhiraja, A. (2016, December 15). Dropout in (Deep) Machine learning | by Amar Budhiraja. Medium. Retrieved June 13, 2022, from <https://medium.com/@amarbudhiraja/https-medium-com-amarbudhiraja-learning-less-to-learn-better-dropout-in-deep-machine-learning-74334da4bfc5>
- [Web16] Retrieved June 13, 2022, from <https://academy.hsoub.com/questions/5777-%D9%85%D8%A7-%D9%87%D9%8A-%D9%84%D8%BA%D8%A9-python-%D9%88%D9%85%D8%A7-%D9%87%D9%8A-%D8%AA%D8%B7%D8%A8%D9%8A%D9%82%D8%A7%D8%AA%D9%87%D8%A7-%D9%88%D8%A8%D9%85%D8%A7%D8%B0%D8%A7-%D8%B3%D9%8A%D9%8F%D9%81%D9%8A%D8%AF%D>
- [Web17] Retrieved June 13, 2022, from <https://reliefweb.int/report/algeria/alg-rie-rapport-de-situation-sur-l-pid-mie-du-covid-19-451-du-22-juin-2021>
- [Web18] Retrieved June 13, 2022, from <https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/sars-cov-2-transmissio%20n.html%20>
- [Web19] Coronavirus. Retrieved June 13, 2022, from https://www.who.int/fr/health-topics/coronavirus/coronavirus#tab=tab_1
- [Web20] <https://arxiv.org/abs/2004.10641>. View at: Google Scholar
- [Web21] <https://arxiv.org/abs/2003.04655>. View at: Google Scholar
- [Web22] github. Retrieved June 13, 2022, from <https://github.com/ieee8023/covid-chestxray-dataset?fbclid=IwAR20JS-CyRv6OPMAXNS9ZYG0mwx0PJ5qVdGC3TfvhRoZ0MoKQrXjMGzD5II>
- [Web23] Retrieved June 13, 2022, from <http://www.sfr-radiologie.asso.fr>
- [Web24] What is Artificial Intelligence (AI)? Definition, Benefits and Use Cases. TechTarget. Retrieved June 13, 2022, from <https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence>
- [Web25] Retrieved June 13, 2022, from <https://www.google.com/url?sa=i&url=https%3A%2F%2Ffr.linkedin.com%2Fpulse%2Fr%25C3%25A9seau-neuronal-et-deep-learning-gabriel->

[viel&psig=AOvVaw0w6hCRNdAL31dDRwCYER8o&ust=1653818472324000&source=images&cd=vfe&ved=0CAwQjRxqFwoTCNDA_av4gfgCFQAAAAAdAAAAABAK](http://www.lagis.univ-lille1.fr/~bonnet/Outils_Simul/complements_image.pdf)

[Web26] Chapitre Généralités sur le traitement d'images Introduction. Retrieved June 13, 2022, from <http://thesis.univ-biskra.dz/2271/6/Chapitre%2003.pdf>

[Web27] Retrieved June 13, 2022, from http://www.lagis.univ-lille1.fr/~bonnet/Outils_Simul/complements_image.pdf

[Web28] Images *Imagerie Medicale* | Vecteurs, photos et PSD gratuits. (n.d.).Freepik. Retrieved June 13, 2022, from <https://fr.freepik.com/photos-vecteurs-libre/imagerie-medicale>

[Web29] Goodfellow, B. (n.d.). Introduction to Deep Learning. Stanford Graphics Lab. Retrieved June 13, 2022, from https://graphics.stanford.edu/courses/cs468-17-spring/LectureSlides/L10%20-%20intro_to_deep_learning.pdf

[Web30] Agarwal, R. (2019, September 17). The 5 Classification Evaluation metrics every Data Scientist must know. Towards Data Science. Retrieved June 13, 2022, from <https://towardsdatascience.com/the-5-classification-evaluation-metrics-you-must-know-aa97784ff226>

[Web31] Kaggle. Retrieved June 13, 2022, from https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia?fbclid=IwAR24z60oO4l3TdqWAUfEADz7mF-y4tP2XqE6MRWp1JaCB86ah8z_zqROv4o

[Web32] by AditiRastogi.Dev Genius. Retrieved June 24, 2022, from <https://blog.devgenius.io/resnet50-6b42934db431>