

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAITRISE EN MATHÉMATIQUES ET INFORMATIQUE
APPLIQUÉES

PAR
ABOUBACAR SALIF TAPSOBA

CONCEPTION D'UN SYSTÈME DE RECONNAISSANCE FACIALE
MASQUÉE

JUILLET 2023

Résumé

L'identification humaine est une tâche fondamentale dans le domaine de la vision par ordinateur. Elle est utilisée en surveillance vidéo, entre autres. Le volume de données acquises par les caméras de surveillance est énorme, ce qui constitue un défi. Il est donc nécessaire d'extraire des caractéristiques de base, qui peuvent être obtenues rapidement tout en minimisant l'espace mémoire.

La reconnaissance des visages a été au cœur de plusieurs nouvelles percées au cours des deux dernières décennies et a régulièrement proposé plusieurs applications pluridisciplinaires qui vont des logiciels commerciaux traditionnels aux applications essentielles d'application de la loi. Les récents développements révolutionnaires dans l'analyse du Big Data, le Cloud Computing, les réseaux sociaux et le Machine Learning ont considérablement transformé la vision conventionnelle de la manière de résoudre plusieurs problèmes complexes en vision par ordinateur. Dans ce projet, nous mènerons une étude approfondie des concepts du Cloud Computing, du Big Data, des réseaux de neurones et du Machine Learning dans une perspective contemporaine de la Reconnaissance des visages, et proposerons une nouvelle approche. Nous réaliserons des tests à une grande échelle et une comparaison avec d'autres méthodes.

Mots-clés :

Reconnaissance faciale, détection de visage, mégadonnées, infonuagique, apprentissage automatique.

REMERCIEMENTS

Avant tous, nous remercions ALLAH, pour nous avoir donné le pouvoir pour réaliser ce travail.

Un grand merci à mon encadreur M. Fathallah Nouboud, Professeur à l'Université du Québec à Trois-Rivières, qui a fourni des efforts énormes par ses conseils, ses orientations et ses informations.

Nous tenons à remercier les membres du jury et tous les enseignants du département Mathématiques et Informatiques Appliquées.

A mon père, ma mère, mon frère, ma sœur, mon âme sœur, mes enfants, mes amis bien-aimés et à toutes personnes ayant participé de près ou de loin à la réalisation de ce modeste travail.

Table des matières

Résumé	II
REMERCIEMENTS.....	III
INTRODUCTION GÉNÉRALE.....	10
CHAPITRE 1 : ÉTAT DE L'ART	12
CHAPITRE 2 : AU CŒUR DU BIG DATA	15
1. Introduction	15
2. Le paysage des données massives	16
2.1. L'activité des données	16
2.2. Évolution des Mégadonnées.....	16
3. Représentation des mégadonnées.....	17
3.1. Définition	17
3.2. Dimensions des Mégadonnées.....	18
4. Sources de données	19
5. Différents Types de données	19
5.1. Données Structurées	19
5.2. Données non structurées	20
5.3. Données semi-structurées	20
6. Infrastructure des mégadonnées	20
7. Cycle de vie des mégadonnées.....	21
8. Défis rencontrés lors de la gestion des mégadonnées	22
9. Techniques d'analyse des données	22
9.1. Analyse descriptive	22
9.2. Analyse prédictive.....	23
9.3. Analyse Prescriptive.....	23
10. Cas d'usages du Big Data	23
10.1. Soins de santé	23
10.2. Les télécoms	24
10.3. Les services financiers.....	24
11. Conclusion	24
CHAPITRE 3 : VISION PAR ORDINATEUR	25

1.	Introduction	25
2.	Aperçu de la vision par ordinateur	26
3.	Les cas d'usages de la vision par ordinateur	27
3.1.	Sécurité et sûreté.....	27
3.2.	Le domaine de l'automobile.....	27
3.3.	Le secteur de la vente	28
3.4.	Le domaine sportif	28
3.5.	Le domaine médical	28
4.	Les principales tâches de la vision par ordinateur.....	28
5.	Les défis de la vision par ordinateur	29
6.	Conclusion	29
CHAPITRE 4 : Le CLOUD COMPUTING		30
1.	Introduction	30
2.	La technologie Cloud Computing	30
2.1.	Définition	30
3.	Modèles de services cloud.....	31
3.1.	SaaS (Software as a Service)	31
3.2.	PaaS (Plateforme as a Service)	32
3.3.	IaaS (Infrastructure as a Service).....	32
4.	Modèles de déploiement.....	33
4.1.	Les Clouds publics	33
4.2.	Les Clouds privés.....	33
4.3.	Les Clouds communautaires	33
4.4.	Les Clouds hybrides.....	34
5.	Avantages et inconvénients du Cloud Computing.....	34
5.1.	Les avantages	34
5.2.	Les inconvénients.....	34
6.	Les caractéristiques du Cloud	35
6.1.	Libre-service à la demande.....	35
6.2.	Large accès au réseau	35
6.3.	Élasticité rapide.....	35
6.4.	Mise en commun des ressources	35
6.5.	Services mesurés	36

7. Cloud Computing et Big Data	36
8. Conclusion	37
CHAPITRE 5 : CONTEXTE D'ÉTUDE.....	38
1. Introduction	38
2. Reconnaissance faciale	38
2.1. Architecture de reconnaissance faciale.....	38
2.2. Les types de représentations	39
2.2.1. Les approches holistiques.....	39
2.2.2. Les approches locales	40
2.2.3. Les approches Hybrides	40
2.3. Les domaines d'applications	40
3. Les réseaux de neurones	41
3.1. Fonction d'activation.....	42
3.2. Algorithme d'apprentissage	43
3.3. Rétropropagation	43
3.4. Bases de données.....	44
3.5. Quelques notions.....	45
4. Présentation des réseaux de neurones convolutifs (CNN).....	45
4.1. Les images RVB	45
4.2. Les couches de convolutions	46
4.3. Couche de mise en commun.....	47
4.4. Les couches entièrement connectées	48
4.5. Présentation des réseaux de neurones LeNet-5	49
5. Amélioration du modèle.....	50
5.1. Augmentations des données.....	50
5.2. Dropout.....	51
5.3. Processus de normalisations par lots	51
5.4. Présentation de la taille des lots.....	51
5.5. Taux d'apprentissage	51
CHAPITRE 6 : MÉTHODOLOGIE DE RECHERCHE	52
1. Introduction	52
2. Environnement logiciel et matériel de développement.....	52
2.1. Configuration Matérielle	52

2.2.	Langage de programmation	53
2.3.	Présentation des bibliothèques	53
2.4.	Environnement de développement	55
3.	Données.....	56
3.1.	Collecte des données.....	56
3.2.	Organisation des données	56
3.3.	Détection de visages.....	58
3.4.	Alignement de visages	59
3.5.	Extractions des caractéristiques (prétraitement des données)	59
4.	Formation des modèles de reconnaissance faciale	60
4.1.	Première expérimentation	61
4.2.	Deuxième expérimentation.....	62
4.3.	Troisième expérimentation	62
4.4.	Bases de données.....	62
4.5.	Les métriques de performances.....	63
CHAPITRE 7 : RÉSULTATS ET DISCUSSIONS		65
1.	Présentations des résultats et discussions.....	65
2.	Résultats obtenus	71
CONCLUSION GÉNÉRALE.....		74
RÉFÉRENCES		76

Liste des figures

FIGURE 1 : IMAGE D'UN HOMME ET D'UN CHAT [38].	26
FIGURE 2 : LES MODÈLES DE SERVICE DU CLOUD COMPUTING [47]	31
FIGURE 3 : PROCESSUS DE RECONNAISSANCE FACIALE [54].	39
FIGURE 4 : ARCHITECTURE D'UN RÉSEAU DE NEURONES [61].	42
FIGURE 5 : UN NEURONE ARTIFICIEL [63].	43
FIGURE 6 : ÉVOLUTION DE LA COURBE D'ERREUR DURANT LA PHASE D'APPRENTISSAGE [61].	44
FIGURE 7 : ARCHITECTURE D'UN RÉSEAU DE NEURONES CONVOLUTIFS.	45
FIGURE 8 : UNE REPRÉSENTATION VISUELLE D'UNE COUCHE CONVOLUTIVE [67].	47
FIGURE 9 : OPÉRATION DE MISE EN COMMUN [69].	48
FIGURE 10 : OPÉRATION D'APLATISSEMENT [70].	49
FIGURE 11 : ARCHITECTURE LUNET-5 [73].	50
FIGURE 12 : COMMANDES D'INSTALLATION ET D'ACTIVATION.	55
FIGURE 13 : EXEMPLES D'IMAGES DE LA BASE DE DONNÉES FFHQ [5].	57
FIGURE 14 : EXEMPLES D'IMAGES DE VISAGES DÉTECTÉS.	58
FIGURE 15 : EXEMPLES D'IMAGES DE VISAGES RECADRÉS.	59
FIGURE 16 : EXEMPLES D'IMAGES APRÈS EXTRACTIONS DES CARACTÉRISTIQUES.	60
FIGURE 17 : ARCHITECTURE DU BIG DATA PROPOSÉ	63
FIGURE 18 : (1) GRAPHIQUE DE PRÉCISION D'ENTRAÎNEMENT ET DE VALIDATION ET (2) GRAPHIQUE DE PERTE D'ENTRAÎNEMENT ET DE VALIDATION POUR LA PREMIÈRE EXPÉRIMENTATION.	66
FIGURE 19 : (1) GRAPHIQUE DE PRÉCISION D'ENTRAÎNEMENT ET DE VALIDATION ET (2) GRAPHIQUE DE PERTE D'ENTRAÎNEMENT ET DE VALIDATION POUR LA DEUXIÈME EXPÉRIMENTATION.	66
FIGURE 20 : (1) GRAPHIQUE DE PRÉCISION D'ENTRAÎNEMENT ET DE VALIDATION ET (2) GRAPHIQUE DE PERTE D'ENTRAÎNEMENT ET DE VALIDATION POUR LA TROISIÈME EXPÉRIMENTATION.	67
FIGURE 21 : MATRICE DE CONFUSION DE LA PREMIÈRE EXPÉRIMENTATION.	68
FIGURE 22 ; MATRICE DE CONFUSION DE LA DEUXIÈME EXPÉRIMENTATION.	69
FIGURE 23 : MATRICE DE CONFUSION DE LA TROISIÈME EXPÉRIMENTATION.	70
FIGURE 24 : RÉSULTATS DE LA RECONNAISSANCE FACIALE DE LA 1ERE EXPÉRIMENTATION.	72
FIGURE 25 : RÉSULTATS DE LA RECONNAISSANCE FACIALE DE LA 2EME EXPÉRIMENTATION.	72
FIGURE 26 : RÉSULTATS DE LA RECONNAISSANCE FACIALE DE LA 3EME EXPÉRIMENTATION.	73

Liste des tableaux

TABLEAU 1 : CONFIGURATION MATÉRIELLE.....	53
TABLEAU 2 : TABLEAU DE COMPARAISON DES RÉSULTATS DES TROIS EXPÉRIMENTATIONS.....	71

INTRODUCTION GÉNÉRALE

Si vous êtes dans le monde de l'informatique ou des affaires, vous avez probablement entendu parler du phénomène Big Data. Le BD est l'une des dernières tendances technologiques qui impacte profondément la façon dont les organisations utilisent l'information. Il fait référence à la collecte, au traitement et à l'organisation des informations à partir de grands ensembles de données. Bien que personne ne sache vraiment ce que signifie le Big Data, il ne fait aucun doute que le « BD » n'est pas seulement un mot à la mode en matière d'informatique. Il s'agit plutôt d'un concept en évolution rapide dans les technologies de l'information et les données.

Le logiciel de reconnaissance faciale est devenu l'une des technologies les plus importantes et les plus controversées au cours de ces dernières années. La reconnaissance faciale est responsable de l'identification et de l'authentification. Elle compare les données à une base de données de visages connus. La technologie de reconnaissance faciale a eu une attention considérable en tant que domaine de recherche actif en raison de son utilisation potentielle dans les applications commerciales et des exigences accrues de sécurité [1].

La pandémie de Covid 19 et ses différentes variantes ont fait la une de la maladie au quotidien. Elle a engendré des changements significatifs au sein de la communauté universitaire. Des cours à distance et en ligne, le respect des mesures de sécurité pour les cours en présentiel... Autant de scénarios possibles pour fournir une formation efficace dans un contexte de crise sanitaire mondiale. Parlant de formation efficace, les universités mènent des activités quotidiennes en fournissant des services au personnel académique et aux étudiants. Ces gens sont tenus de porter des masques lors des activités académiques pour se protéger mais aussi pour contrôler la propagation du coronavirus [2]. Toutefois, cette mesure préventive soulève un défi pour les universités : comment assurer la surveillance efficace du respect du port de masques dans un environnement académique dynamique et diversifié?

A la lumière de cette question nous développons un système de reconnaissance de masque facial pour identifier les visages sans masques et encourager les gens à bien porter les masques. Notre stratégie pour atteindre notre objectif consiste à utiliser des réseaux de neurones convolutifs (CNN) [3], qui ont démontré leur efficacité dans l'extraction de caractéristiques, une étape très importante dans le processus de classification des images. Les réseaux de neurones

convolutifs utilisent des grands ensembles de données pour les applications de reconnaissance faciale masquée. Cependant, l'acquisition de quantités massives de données est souvent un défi crucial pour obtenir d'excellentes performances.

Une approche combinant la technique d'augmentation des données et le modèle LeNet-5 [4] est proposée pour la résolution de notre problématique. Notre travail mène des expériences de reconnaissances des visages masqués sur le jeu de données Flickr-Face HQ (FFHQ) [5].

Notre projet de recherche est structuré en sept (7) chapitres. Tout d'abord, le premier chapitre aborde l'état de l'art, en examinant les recherches existantes dans les domaines de la reconnaissance faciale et des réseaux de neurones convolutifs. Par la suite, dans le deuxième chapitre, nous nous penchons sur le concept du « Big Data », en décrivant en détail le paysage des mégadonnées, les différents types de données et les technologies utilisées pour leur traitement. Le troisième chapitre, intitulé « Vision par ordinateur », explore les différentes applications de cette discipline ainsi que les principales tâches qu'elle englobe. Ensuite, dans le quatrième chapitre, nous nous intéressons au « Cloud Computing », en étudiant les modèles de déploiement et de service, leurs avantages respectifs et les critères à prendre en compte lors du choix d'un modèle spécifique. Dans le cinquième chapitre, appelé « Contexte d'étude », nous approfondissons notre compréhension des technologies de reconnaissance faciale et des réseaux de neurones artificiels, en explorant également les diverses techniques susceptibles de garantir des résultats satisfaisants. Par la suite, le sixième chapitre, intitulé « Méthodologie de recherche », présente en détail la méthodologie que nous avons élaborée pour mener à bien notre projet, ainsi que les outils spécifiques que nous avons utilisés. Enfin, dans le dernier chapitre intitulé « Présentation des résultats et discussion », nous mettons en avant les résultats obtenus grâce à nos travaux de recherche et entamons une discussion approfondie autour de ces résultats.

CHAPITRE 1 : ÉTAT DE L'ART

La technologie de reconnaissance faciale est un domaine de recherche qui a attiré beaucoup d'attention aux cours de ces dernières années [1, 6]. Elle comprend généralement quatre étapes, à savoir la détection des visages, l'alignement, l'extraction des caractéristiques faciales et la classification des visages [7]. Les caractéristiques faciales utilisées dans les technologies de reconnaissance faciale incluent souvent le nez, la bouche et les yeux. Cependant, diverses situations et circonstances, comme la pandémie de COVID-19, ont rendu le port de masques courant, ce qui masque partiellement les visages et affecte donc la précision de la reconnaissance faciale [8, 9]. Cela renforce la nécessité de former des modèles d'apprentissage en profondeur capables d'identifier et de vérifier avec précision les systèmes de reconnaissance faciale masquée. De nombreuses recherches ont été entreprises pour relever le défi de la reconnaissance faciale masquée dans le domaine de l'apprentissage en profondeur.

N. U. Din et al [10] ont proposé un réseau basé sur un réseau antagoniste génératif (GAN) pour retirer automatiquement le masque du visage. Ils ont réussi à préserver la cohérence de la structure et de la forme du visage reconstruit en utilisant deux discriminateurs pour apprendre la structure générale de la zone masquée en profondeur. Leur approche surpasse qualitativement et quantitativement les approches représentatives précédentes dans l'état de l'art.

G. Deore et al [11] ont présenté l'identification des personnes masquées dans une caméra de vidéosurveillance. Ils ont utilisé l'algorithme de HOG (Histogramme des gradients) pour la détection de personne. Leur algorithme détermine qu'une personne porte un masque facial lorsqu'elle est détectée et que son visage ne l'est pas. Leurs travaux ont contribué à la construction d'un modèle d'identification du visage masqué, offrant ainsi une compréhension analogue de cette problématique.

M. S. Ejaz et al. [12] ont proposé l'approche d'augmentation des données des images de visages masqués et non masqués pour une meilleure précision à l'aide du réseau de neurones convolutifs en cascade multitâche (MTCNN), du modèle d'intégration Google FaceNet et du classificateur SVM (Séparateurs à vaste marge). Dans cette approche, les régions faciales sont détectées à l'aide du réseau MTCNN. Ensuite, l'extraction des caractéristiques faciales est effectuée à l'aide du modèle d'intégration Google FaceNet. Enfin, le classificateur SVM est utilisé pour

reconnaître les identités des individus. Les résultats obtenus par les auteurs démontrent une performance remarquable en termes de reconnaissance faciale masquée.

A. Anwar et al. [13] ont mené des recherches sur la reconnaissance faciale masquée dans le contexte de l'authentification sécurisée. Ils ont présenté le modèle MaskTheFace, un outil open-source qui génère un grand ensemble de données de visages masqués. MaskTheFace détecte les points de repère du visage pour identifier les principales caractéristiques du visage. Les auteurs ont évalué la précision de leur système réentraîné en utilisant l'ensemble de données de reconnaissance faciale masquée MFR2[14]. Les résultats obtenus ont démontré une précision comparable à celle des approches existantes.

P. Nagrath et al. [15] ont obtenu des résultats prometteurs avec une précision de 93% et un score F1 de 93% en utilisant le modèle SSDMNv2 pré-entraîné sur l'ensemble de données de visages masqués provenant de Kaggle [16]. L'objectif de leur recherche était de détecter en temps réel les masques faciaux. Pour atteindre cet objectif, ils ont utilisé le classificateur d'images MobileNetV2 pour classifier les images. Les performances élevées obtenues démontrent l'efficacité de leur approche dans la détection des masques faciaux en temps réel.

S. V. Militante et N. V. Dionisio [17] ont proposé une approche utilisant l'apprentissage profond et les réseaux de neurones convolutifs (CNN) pour les systèmes de reconnaissance faciale masquée en temps réel. Leur objectif principal était de déterminer si un visage était porteur d'un masque ou non. Les chercheurs ont réalisé une étude qui a abouti à des résultats optimaux pour la détection des masques faciaux. Après de nombreux essais, ils ont obtenu une précision de 96% pour leurs résultats.

S. Ge et al. [18] ont développé un modèle en utilisant un ensemble de données pour la détection de visages masqués et non-masqués. L'ensemble de données des visages masqués comportait 35 806 images de personnes portant des masques. Les auteurs ont utilisé un réseau neuronal convolutif pour concevoir leur modèle, qui comprend trois modules distincts : proposition, mise en œuvre et authentification. Grâce à leur travail, ils ont obtenu des résultats avec une précision de 76,1%.

M. Loey et al. [19] ont réalisé une étude portant sur un modèle hybride utilisant l'apprentissage par transfert pour détecter les masques faciaux dans des images réelles. Leur modèle exploite trois ensembles de données : l'ensemble de données

RMFD (Real-world Masked Faces Dataset) [20] pour les visages masqués du monde réel, l'ensemble de données SMFD (Simulated Masked Faces Dataset) [21] pour les visages masqués simulés, et l'ensemble de données LFW (Labeled Faces in the Wild) [22] pour les visages non masqués. L'approche proposée comprend l'utilisation du modèle ResNet-50 pour extraire les caractéristiques des visages, et du modèle Yolov2 pour la détection des masques faciaux. Les chercheurs ont obtenu d'excellents résultats avec une précision de 99,64% pour RMFD, 99,48% pour SMFD, et 100% pour LFW respectivement.

CHAPITRE 2 : AU CŒUR DU BIG DATA

1. Introduction

La croissance exponentielle des utilisateurs d'internet a engendré une saturation du flux de données au cours des dernières décennies. Les données générées exponentiellement à partir des applications informatiques, sont collectées puis traitées dans une panoplie de domaines d'applications tels que les services mobiles, les services financiers, les sciences naturelles et génie. Selon une enquête récente d'IBM 2,5 quintillions (2 500 000 000 000 000 000 000, soit $2,5 \times 10^{18}$) octets de données sont générés chaque jour [23].

La dernière décennie fut marquée par une variété de plateformes différentes des médias sociaux, chacune d'elle générant une vaste quantité de données. Ces données, plus ou moins extractibles, deviennent la nouvelle propriété intellectuelle. Les appareils intelligents, tout comme les réseaux sociaux, sont devenus incontournables dans le quotidien des utilisateurs. En effet les appareils intelligents et les réseaux sociaux évoluent parallèlement. Les appareils intelligents connectés à d'autres appareils quotidiens tels que nos réfrigérateurs et nos montres génèrent, d'autres part, une grande quantité d'informations.

Nous dépendons de plus en plus d'internet et du Web, et le secteur académique n'en fait pas l'exception à travers les recherches scientifiques et les activités éducatives. La pandémie de Covid-19 a déclenché une création de nouvelles données depuis ses débuts. Cette création se justifie par l'adaptation des environnements de travail ainsi que les divertissements à distance. Cependant la sphère mondiale des données connaîtra une croissance due aux interactions dans les années à venir [24].

Nul doute que les données sont la nouvelle source économique du fait que la taille de l'univers numérique augmente avec le temps. La quantité de données a connu un boom, et par conséquent le terme Big Data qui se réfère aux mégadonnées recèle un énorme potentiel pour améliorer nos vies.

2. Le paysage des données massives

2.1. L'activité des données

La quantité de données créées et stockées a considérablement augmenté. Depuis les années 80, la production de données s'est accrue avec l'informatique personnelle. Cela inclut les vidéos, les audios, les textes, les journaux et les photos, ainsi que les machines connectées aux objets avec lesquels les gens interagissent. Les machines connectées les unes aux autres ajoutent de manière significative à la quantité massive de données existantes. Le développement technologique a conduit à l'intégration de capteurs, de téléphones portables et de compteurs dans des objets quotidiens comme les véhicules, les vêtements et les sources d'énergie.

La croissance des données ne se limite pas à l'activité humaine et à la connexion des machines. Il existe des données publiées par de nombreuses organisations et entreprises. Comprendre les clients, les produits et les services d'une entreprise nécessite beaucoup de données et d'analyses. Ainsi les géants du Web tels que Google, Apple et Amazon accroissent l'utilisation de leurs services grâce à la consommation des technologies de l'information pour une plus grande puissance de traitement et de stockage de données

La démocratisation des appareils intelligents ainsi que l'omniprésence des réseaux sociaux constituent une nouvelle source de création de quantité de données. En effet, nos habitudes ont changé : au lieu d'attendre des ordinateurs qu'ils remplissent des fonctions spécifiques, nous attendons d'eux qu'ils s'adaptent aux nouvelles tendances.

2.2. Évolution des Mégadonnées

Toute nouvelle technologie est toujours due aux limites des technologies plus anciennes et à la nécessité de les surmonter. Cependant, la plupart des nouvelles technologies des données s'appuient sur des idées plus anciennes même si elles sont incrémentielles ou révolutionnaires. L'accroissement des capacités de stockage de données a permis d'améliorer les performances et de réduire les coûts. Conformément à la loi de Moore [25], le prix de diverses ressources informatiques a chuté de façon régulière au cours des 20 dernières années. Cela inclut l'espace de stockage et le nombre de nœuds.

L'augmentation exponentielle du volume de données, qui se mesure en téraoctets et en pétaoctets, s'est avérée être un défi pour le Système de Gestion des Bases de Données Relationnelles (SGBDR) dans la gestion d'un volume de données aussi massif. La gestion des données englobe des nouvelles technologies matérielles, réseau et logicielles telles que le Cloud Computing et la virtualisation.

La convergence de ces technologies émergentes et la réduction des coûts du stockage aux cycles de calcul ont transformé le paysage des données et rendu possibles de nouvelles opportunités.

La dernière technologie à émerger est le Big Data littéralement appelée les mégadonnées.

3. Representation des mégadonnées

3.1. Définition

« The ability of society to harness information in novel ways to produce useful insights or goods and services of significant value » and « things one can do at a large scale that cannot be done at a smaller one, to extract new insights or create new forms of value. » by Mayer-Schonberger and Cukier (2013) [26].

Le Big Data est littéralement défini comme des mégadonnées ou des données massives. C'est un ensemble d'entités de données hétérogènes en extensibilité permanente qui ne peuvent pas être pris en charge par les systèmes de gestion de données traditionnels. Il est aussi défini comme une architecture distribuée et évolutive pour stocker et traiter de gros volumes de données.

En effet, les données sont créées chaque jour à l'aide de différents outils. Le Big Data est un terme désignant des ensembles de données très volumineuses. Cet ensemble de données géré par les grandes sociétés Internet telles que Google, Facebook, Amazon et Yahoo a changé les processus, les objectifs et la prise de décision ces dernières années [27].

L'apparition du Big Data marque une nouvelle révolution industrielle similaire à celle de l'électricité. Il n'existe cependant pas de définitions universelles pour les mégadonnées car les données elles-mêmes sont un concept complexe qui nécessite une interprétation en fonction des communautés qui les utilisent.

Dans la communauté de l'analyse, les données massives sont conceptualisées par les trois dimensions suivantes : [28]

- Volumes de données en croissance exponentielle;
- Vitesse de transmission extrêmement élevée;
- Large variété de données.

3.2. Dimensions des Mégadonnées

Le Big Data se distingue par ses caractéristiques exceptionnelles aux dimensions variées. La première des dimensions est la taille des données. La taille des données augmente car il y a eu une diminution significative de coût de stockage et de CPU. La deuxième dimension du Big Data est la vitesse. Elle concerne la rapidité avec laquelle les données traitées peuvent être produites et analysées. Puis la variété, qui définit l'hétérogénéité des données qu'elles soient structurées, semi-structurées ou non structurées. D'où la troisième dimension.

3.2.1. Volume

Créer des solutions commerciales meilleures et plus adaptées est un objectif pour nombreuses entreprises. En effet, les solutions plus larges impliquent une augmentation significative des données à stocker. Le volume fait référence aux grandes quantités de données générées et enregistrées. Le volume Big Data mesure des téraoctets aux zettaoctets (1024 Go = 1 téraoctet ; 1024 To = 1 pétaoctet ; 1024 Po = 1 exaoctet ; 1024 EB = 1 zettaoctet ; 1024 ZB = 1 yottabyte) [23].

3.2.2. Vitesse

Les données doivent être stockées, traitées, analysées et livrées en temps réel. La rapidité avec laquelle les données sont reçues et enregistrées puis analysées est la vitesse. Les données sont collectées à un rythme élevé, ce qui rend difficile leur sauvegarde dans un court laps de temps. Les entreprises doivent appréhender la vitesse non seulement en termes de création de données, mais aussi sur le plan de leur traitement, de leur analyse et de leur restitution à l'utilisateur en respectant les exigences des applications en temps réel.

3.2.3. Variété

La variété fait référence au format de données pris en charge par le Big Data. Les données arrivent sous un format structuré, semi-structuré et non structuré. L'analyse des mégadonnées est possible grâce à l'essor de l'intelligence artificielle et de l'apprentissage automatique. Différentes sources de données : des avions aux trains et aux voitures, circulent constamment dans des systèmes distincts. Cela permet la classification, l'intégration et l'analyse des données de nouvelles façons.

3.2.4. Autres dimensions

D'autres dimensions définissent aussi le Big Data; à savoir la véracité et la valeur des données. La véracité fait référence à la fiabilité des données. Pour extraire la valeur des données, celles-ci doivent être nettoyées pour éliminer le bruit. C'est

pourquoi, des données précises et fiables recueillies à partir de mégadonnées sont véridiques. La valeur fait référence à la quantité de connaissances utiles qui peuvent être extraites des données. Les données ressemblent à une mine d'or potentielle, mais comme dans une mine d'or, il y'a peu d'or et beaucoup de résidus [29].

4. Sources de données

L'augmentation massive de la taille des données est due à la numérisation mondiale de divers appareils. Les données proviennent de multiples sources de différents domaines. Bon nombre de ces changements ont été causés par l'utilisation d'appareils électroniques, notamment les ordinateurs, les téléphones cellulaires et les achats en ligne. D'autres raisons de l'augmentation du volume de données sont les transactions par courriel, les paiements de factures et les publications sur les réseaux sociaux.

Quelques exemples de sources de données :

- Les dossiers de santé électroniques;
- Les flux des médias sociaux;
- Les systèmes internet des objets (IoT) envoyant des données en temps réel;
- Les courriels.

5. Différents Types de données

Généralement les données sont produites par les humains ou les machines. Les données générées par l'humain sont créées via des interactions entre les humains et les machines et, celles générées par les ordinateurs, les périphériques matériels ou les logiciels sont appelées données générées par la machine. Toutes les données sont représentées en trois catégories :

- Données structurées;
- Données non structurées;
- Données semi-structurées.

5.1. Données Structurées

Les données stockées dans un format de tableau avec des colonnes et des lignes sont appelées données structurées. Les données structurées proviennent de plusieurs types d'organisations. Elles sont plus faciles à comprendre et à analyser grâce à leurs propriétés organisationnelles spécifiques. Quelques exemples de données structurées incluent le nom de l'employé et le titre du poste; de nombreux autres exemples peuvent être trouvés dans les nombres, les dates et les chaînes. Ces données sont souvent stockées dans une base de données

structurées, implémentées à l'aide d'un schéma puis sont interrogées à l'aide d'un langage de requête structuré, ou SQL.

5.2. Données non structurées

Contrairement aux données structurées, les données non structurées sont généralement sans structures définies. Elles se présentent sous de nombreuses formes, mais elles ne suivent pas de directives spécifiques. Générées par l'homme ou par la machine, les données non structurées sont généralement de grande taille. Des exemples de données non structurées comprennent la vidéo, l'audio, les images, les courriels, les fichiers texte et les publications sur les médias sociaux. Ceci contribuant toujours à la croissance des données.

5.3. Données semi-structurées

Les données semi-structurées sont une combinaison de données structurées et non structurées. Cela signifie qu'elles intègrent certains éléments d'un modèle de données défini sans suivre un format spécifique comme celui d'une base de données ou XML. Les données recueillies à partir de capteurs ou de journaux de serveurs Web sont généralement semi-structurées.

6. Infrastructure des mégadonnées

Les entreprises ont longtemps utilisé les bases de données traditionnelles pour le traitement des données structurées. Cependant, avec la croissance exponentielle des données, ces bases de données traditionnelles sont souvent problématiques. Pour tirer le meilleur parti des données, les entreprises se sont penchées sur les nouvelles technologies et outils. Les outils et technologies de mégadonnées permettent de stocker, de traiter et d'analyser des données. Ceux-ci permettent l'utilisation d'applications sur de grands systèmes qui nécessitent beaucoup de données avec une latence, un coût et une évolutivité minimaux. La plupart des nouveaux systèmes sont open source et présentent une faible latence, un faible coût et une grande évolutivité. Les différentes technologies des mégadonnées sont : Hadoop, HDFS et MapReduce [30].

Apache Hadoop est une plateforme open source écrite en Java pour le traitement de grands ensembles de données. Il s'agit d'un cadre capable de traiter des données semi-structurées, structurées et non structurées à des vitesses élevées. Hadoop stocke une grande quantité de données dans un système de fichiers distribué et les traite en parallèle sur plusieurs ordinateurs. L'évolutivité de Hadoop fait référence à sa capacité de gérer des charges plus importantes tout en maintenant les performances. Chaque fichier Hadoop est écrit une fois et lu plusieurs fois. Les fichiers ne peuvent pas être modifiés. Un groupe d'ordinateurs connectés les uns aux

autres et vus comme un seul ordinateur est appelé un cluster. Les clusters sont utilisés pour stocker et analyser des quantités massives de données provenant de nombreuses sources de manière rentable. Certains sont utilisés pour créer, stocker et analyser des données avec Hadoop [31].

HDFS [30, 31 et 32] est un système de fichiers open source conçu pour gérer de grands ensembles de données avec un modèle d'accès en lecture ou en écriture en continu. Il est exécuté sur du matériel de base à faible coût. Les ensembles de données HDFS sont écrits une seule fois dans le système de fichiers et lus plusieurs fois. Ce modèle à écriture unique et à lecture multiple permet d'effectuer des analyses sur l'ensemble de données pour extraire des informations.

MapReduce [33] est le modèle de traitement par lots de Hadoop. Il utilise une approche divisée pour mieux régner qui est très fiable, tolérante aux pannes et évolutive. MapReduce traite tout type de données en parallèle et dans des environnements distribués; cela ne fonctionne qu'avec les charges de travail par lots. Le paradigme traditionnel du traitement par lots impliquait le déplacement des données entre la plate-forme de stockage et la plate-forme de traitement. En revanche, le traitement des données MapReduce a lieu là où les données sont stockées. L'approche traditionnelle prend plus de temps que l'approche MapReduce.

7. Cycle de vie des mégadonnées

Les mégadonnées apportent d'énormes avantages, des idées commerciales innovantes aux traitements non conventionnels des maladies, en passant par des défis à relever. Le défi se pose parce que la technologie d'aujourd'hui recueille de données. Les technologies du Big Data sont capables de les capturer et de les analyser efficacement. Le cycle de vie des mégadonnées est constitué de cinq phases. Celles-ci incluent la collecte de données, la classification des données, le nettoyage des données, la modélisation des données et la livraison des données.

La collecte de données implique le stockage de données provenant de différentes ressources. Les données brutes ainsi recueillies à partir de tous les processus analytiques sont introduites dans un système de stockage de données pour un traitement ultérieur. Ces données sont souvent redondantes, occupent de l'espace supplémentaire et augmentent ainsi les coûts de stockage.

La classification des données regroupe les données dans des catégories spécifiques. Cela peut être basé sur le fait qu'elles soient structurées, semi-structurées ou non structurées. Toutes les données sont intégrées et présentées de manière unifiées pour le nettoyage des données, la modélisation des données.

Le nettoyage des données garantit que les données sont exactes, cohérentes et complètes en supprimant les redondances et les incohérences. Il consiste à préparer les données pour la phase d'analyse.

Au cours de la phase de modélisation des données, une analyse des données est effectuée et le résultat est une quantité agrégée de données aux fins prévues. Des rapports sont ensuite générés à partir des résultats de la phase de modélisation des données, après la phase de livraison des données.

8. Défis rencontrés lors de la gestion des mégadonnées

Il existe un certain nombre de défis lorsqu'il s'agit de données volumineuses. Le véritable défi lors du traitement des données consiste à trouver un sens à des sources d'information disparates. Les données peuvent être structurées ou non structurées, avec des capteurs détectant à la fois les données vidéo et image. Les données des satellites, des transactions commerciales et des médias sociaux peuvent également être collectées de manière structurée. Cependant, pour trouver un sens aux données, il faut intégrer des données disparates provenant de différentes sources. Trouver du sens dans le traitement des données est plus difficile, car chaque traitement doit être effectué séparément.

9. Techniques d'analyse des données

Les méthodes d'analyse des données pour le Big Data sont principalement de trois (3) types:

9.1. Analyse descriptive

Les données sont trop volumineuses pour détecter des informations présentes mais masquées par le volume. C'est pourquoi les méthodes descriptives emploient des techniques de mise en évidence. La description consiste à analyser un groupe de données afin de fournir un aperçu de celui-ci. Ce processus implique la combinaison de plusieurs résultats d'analyse, tels que des mesures et des modèles. Parmi les techniques et algorithmes utilisés dans l'analyse descriptive, on cite :

- Analyse factorielle (ACP et ACM);
- Méthode des centres mobiles;
- Classification hiérarchique;
- Classification neuronale;
- Recherche d'association.

9.2. Analyse prédictive

Les prédictions utilisent les informations existantes pour estimer de nouvelles informations. L'analyse prédictive utilise des données historiques et en temps réel pour trouver des tendances qui aident à prédire les résultats futurs. Ces analyses sont plus fiables si les modèles utilisés sont fiables. Cette technique fait appels à de l'intelligence artificielle; les principales méthodes sont :

- Arbres de décision;
- Réseaux de neurones;
- Classification bayésienne;
- Support Vector Machine (SVM);
- K-plus proches voisins (KNN).

9.3. Analyse Prescriptive

Méthodes conçues pour prédire le meilleur plan d'action vers un objectif particulier. Cherchant à améliorer la prise de décision commerciale dans des situations spécifiques, l'analyse prescriptive utilise les données et l'intelligence artificielle pour fournir des conseils sur les choix commerciaux et les actions stratégiques. Elle est basée sur des méthodes d'analyse spécifiques et se concentre sur une plus grande variété de problèmes que l'analyse traditionnelle.

10. Cas d'usages du Big Data

Le Big Data a de nombreuses utilisations en dehors du monde académique. Ces applications incluent le transport, l'assurance, la banque, les télécoms, l'industrie et même les loisirs. Des exemples sont cités ci-dessous.

10.1. Soins de santé

De nombreux prestataires de soins de santé ont déjà mis en place des méthodes de collecte de données à grande échelle. D'autres sont ouverts à l'idée de le faire à l'avenir. Les organisations de santé utilisent fréquemment un traitement personnalisé, la prévision des patients admis à l'hôpital, ainsi que la gestion et l'optimisation de leurs pratiques en tant que cas d'utilisation populaires du Big Data. L'analyse des données peut être utilisée dans les organisations de santé pour améliorer la recherche sur les soins de santé.

10.2. Les télécoms

Les trois principales raisons pour lesquelles les gens utilisent le Big Data dans les télécommunications sont d'attirer de nouveaux clients, d'améliorer les performances du réseau et de conserver les clients actuels. À 87 %, l'industrie des télécommunications a déjà adopté le Big Data plus que toute autre industrie [34]. Les télécoms prévoient d'inclure de nouvelles utilisations de données pour leurs mégadonnées en analysant les appareils en fonction de leur emplacement et en incluant l'assurance des revenus. Ils veulent aussi optimiser les réseaux, les centres d'appels et les prix des équipements.

10.3. Les services financiers

De nombreuses institutions financières utilisent actuellement le Big Data. Les sociétés de services financiers utilisent le BD pour analyser les clients à des fins de personnalisation, d'évaluation des risques, de détection des fraudes et de menaces à la sécurité. Les autres utilisations des institutions financières destinées à être ajoutées de 2017 à 2018 comprenaient le trading algorithmique, l'analyse de la sécurité basée sur la localisation et l'analyse des données. La liste des priorités de 2018 a vu l'IA et l'apprentissage automatique dépasser l'analyse prédictive en tant que principaux investissements dans les services financiers. Ils sont arrivés en deuxième position avec 29% des voix. En 2017, ils étaient en tête de liste avec une plus grande part des voix. C'est ce que révèle le rapport d'enquête 2018 de Advanced Analytics & AI Summit [35].

11. Conclusion

Les données qui auparavant ne pouvaient pas être traitées avec des méthodes standards sont devenues le carburant de la révolution numérique.

Le Big Data nécessite des connaissances spécialisées et de grandes compétences en matière d'analyse et de manipulation des données. C'est un écosystème extrêmement complexe qui prend des années à maîtriser. Il demande de la compétence et de l'expertise dans la maîtrise et l'analyse des données.

Le BD affecte tous les secteurs d'activités humaines ; de la santé en passant par les télécommunications et les services financiers. Dans le domaine du transport, nous citons comme exemple le forum international du transport qui traite la problématique « Données massives dans les transports : applications, implications et limites ». Les participants du forum essaient de répondre à certaines questions relatives à l'utilisation du Big Data par les gouvernements dans le domaine du transport urbain [36].

CHAPITRE 3 : VISION PAR ORDINATEUR

1. Introduction

Prendre constamment des photos et des vidéos avec les appareils intelligents est une tendance actuelle. Nous prenons constamment des photos et des vidéos que nous partageons sur plusieurs plateformes des médias sociaux. Diverses tendances photographiques sont populaires depuis plusieurs années.

Internet contient désormais à la fois des informations textuelles et des images visuelles. Les services d'hébergement vidéo comme YouTube ont énormément gagné en popularité. Chaque minute, les utilisateurs téléchargent et regardent des millions d'heures de vidéo [37].

Depuis les attentats terroristes du 11 septembre 2001 des tours du World Trade Center à New York par des avions du groupe Al-Qaïda aux ÉTATS-UNIS [38], une corrélation significative s'est établie en ce qui concerne l'utilisation des ordinateurs. En réponse à cette tragédie, les programmes de sécurité et de surveillance ont été renforcés, ce qui a influencé la manière dont les ordinateurs sont utilisés. Les images vidéo avec analyse de mouvement offrent un large éventail d'avantages. En fait, la détermination des zones de mouvement dans une séquence d'images révèle souvent des événements clés sur lesquels un système de vision doit se concentrer.

Auparavant, les ordinateurs ne pouvaient comprendre les images qu'en remplissant une méta-description. Mais grâce aux récents progrès de la vision par ordinateur, les machines comprennent désormais les images par elles-mêmes.

2. Aperçu de la vision par ordinateur



FIGURE 1 : IMAGE D'UN HOMME ET D'UN CHAT [39].

Que contient l'image ci-dessous ? On voit bien qu'il s'agit d'une photo d'un homme et d'un chat. Nous comprenons instantanément la signification d'une image en tant qu'être humain. Les ordinateurs, en revanche, n'étaient pas capables de le faire jusqu'à tout récemment.

La vision par ordinateur est une discipline informatique qui apprend aux ordinateurs à interpréter le monde qui les entoure. Son but est de fournir aux ordinateurs des informations du monde réel en examinant leur environnement à travers des images.

Pour se rendre compte de l'utilité que peut avoir la vision par ordinateur, réfléchissez à la façon dont les ordinateurs peuvent vous être utiles lorsque vous ne voyez plus.

Nous recevons un des sens de notre environnement à travers nos yeux. Par conséquent, bon nombre de nos décisions quotidiennes sont influencées par ce que nous voyons. La vision par ordinateur consiste d'autres part à assister les personnes dans l'exécution des tâches.

Le fonctionnement de la vision par ordinateur est basé sur des outils matériels tels que les caméras, des ressources dans le cloud, les logiciels et l'intelligence

artificielle. Il existe plusieurs outils mais nous avons tendance à généraliser les algorithmes basés sur des appareils de prise de vue courants comme les smartphones.

Prendre des images sur Internet est une ressource inépuisable. Nous analysons des images déjà stockées sur le net, leur donnant une valeur significative même lorsqu'elles ne sont pas utilisées. Nombreuses tâches de vision par ordinateur utilisent ce fait à leur avantage en créant une énorme base de données d'images facile d'accès.

3. Les cas d'usages de la vision par ordinateur

Les applications de vision par ordinateur visent à améliorer les capacités des ordinateurs en imitant la perception et la cognition humaines. On recense aujourd'hui de nombreux domaines d'applications.

3.1. Sécurité et sureté

Dans le paysage de la sécurité d'aujourd'hui, la surveillance est nécessaire comme dans les magasins de ventes. Dans les rues, les parcs publics et les musées; les caméras de surveillance se trouvent partout. Elles filment et enregistrent des images. La reconnaissance visuelle permet d'identifier et d'authentifier les personnes dans les séquences qui ont été enregistrées.

Une inspection visuelle automatisée avec des drones est possible et entraine un gain de temps prodigieux dans la prévention des délits.

3.2. Le domaine de l'automobile

L'implication de l'informatique dans l'automobile aide les constructeurs automobiles à lutter contre la fatigue de la route [40]. La fatigue en circulation est l'une des causes des accidents de voitures entraînant des blessures graves et souvent des pertes humaines. Imaginez par exemple une voiture équipée d'une caméra et des systèmes de vision par ordinateur avec une vue de 360 degrés. Cette voiture identifiera les objets et réagira à des problèmes afin de réduire les taux d'accidents.

Les constructeurs comme TESLA [41], avec les systèmes de voitures sans pilote permettent à leurs véhicules de naviguer en toute sécurité dans la circulation tout en lisant les panneaux de circulations.

De plus en plus de véhicules utilisent des systèmes de vision par ordinateur mais le monde est toujours en attente de sa première voiture autonome.

3.3. Le secteur de la vente

Tout comme l'automobile et la sécurité, la vente est un domaine d'application des technologies de vision par ordinateur. Dans les magasins de ventes, les algorithmes de vision par ordinateurs sont utilisés pour l'amélioration de l'expérience clientèle tout en réduisant les vols et toutes autres pertes [42]. La vision par ordinateur détecte les défauts avec une méthodologie cohérente qui ne nécessite aucune intervention humaine. Cela permet de ne pas passer à côté de défauts critiques et augmente le risque d'inattention.

3.4. Le domaine sportif

La vision par ordinateur apporte une aide utile dans le domaine du sport de nos jours. Les sports comme le baseball, le football, et le basketball utilisent une technologie de suivi de balle avec précision. La coupe du monde de football 2022 en a été un exemple avec l'utilisation de la Goal line technologie [43].

3.5. Le domaine médical

La vision par ordinateur est utilisée pour diagnostiquer des problèmes de santé à l'aide de données recueillies par des caméras. Des interventions vitales peuvent être effectuées grâce aux algorithmes de vision par ordinateur. C'est le cas d'un chirurgien qui serait assisté par un logiciel de surveillance de la perte de sang en temps réel lors d'un accouchement par césarienne. En effet, le gain de temps pour le chirurgien et le gain de sécurité pour la femme améliorent les procédures médicales et chirurgicales [44].

L'imagerie médicale bénéficie des caractéristiques intrinsèques des machines de vision par ordinateur pour pouvoir être exécutée à grande échelle, avec un résultat constant. Plus de patients sont traités avec moins de risques d'erreurs.

4. Les principales tâches de la vision par ordinateur

Comprendre l'environnement humain, tout comme les humains, nécessite que les ordinateurs effectuent plusieurs tâches. Alors, quelles sont les tâches qui permettent aux ordinateurs de « voir »?

Nous énumérons la classification comme première tâche. Les modèles de classification sont formés pour regrouper des objets dans des catégories spécifiques en leur fournissant un ensemble de données d'objets définis. La classification implique l'attribution d'une étiquette ou d'une classe à une image entière. La détection est la combinaison de la classification et de la localisation. La localisation c'est le fait de connaître l'emplacement d'objets sur l'image. La détection c'est donc

de trouver des objets dans une image tout en déterminant leur position. La segmentation est la détermination des contours de différents éléments qui composent une image. Elle dessine des polygones qui circonscrivent l'image.

5. Les défis de la vision par ordinateur

Permettre à un véhicule de lire les panneaux de signalisation comme un être humain est difficile. La plus grande difficulté pour comprendre la vision artificielle est le manque de compréhension que nous avons du fonctionnement du système visuel humain. Malgré tous les progrès technologiques, la capacité de la vision humaine dépasse encore aujourd'hui celle de la vision par ordinateur.

De plus certaines tâches comme laver la vaisselle, lire un livre obligent l'humain à connaître suffisamment sur son environnement. De telles tâches sont différentes de la vision humaine simple, et nécessitent une compréhension de la présence et le comportement d'un objet dans une image.

6. Conclusion

En résumé nous avons donc vu que les systèmes de vision par ordinateur sont composés de trois tâches qui sont la classification, la détection et la segmentation. Plusieurs domaines tels que la santé utilisent les approches de vision par ordinateur pour diagnostiquer les problèmes de santé. Ce pendant le plus grand défi de ses systèmes est d'égaliser la vision humaine.

CHAPITRE 4 : Le CLOUD COMPUTING

1. Introduction

Internet n'a cessé de croître depuis sa création. Dans le paradigme du Big Data, des infrastructures matérielles fiables sont essentielles pour fournir un stockage fiable. La démocratisation de la technologie du BD profite également des services cloud, car les entreprises clientes n'ont plus à supporter le coût de la mise en place de l'infrastructure. Avec la promesse de fournir de meilleurs modèles commerciaux et services rapidement et à moindre coût, le Cloud Computing est devenu un moteur majeur d'innovation pour les entreprises de tous les secteurs. Les gens utilisent des logiciels et du matériel via des services en ligne connectés qui offrent des performances et une expansion supplémentaire [45].

La pandémie à COVID-19 nous a obligés à travailler à domicile via le Cloud par conséquent, nous en sommes venus à le considérer comme normal. Cela a conduit à une augmentation significative des discussions sur le sujet. Des enfants à l'école, aux adultes à la maison, de nombreuses personnes ont intégré le CC dans leur vie quotidienne; une journée typique passée en streaming, à télétravailler.

2. La technologie Cloud Computing

2.1. Définition

À ses débuts, toute nouvelle technologie fait souvent l'objet d'un battage médiatique important. En effet, les nouvelles technologies sont presque toujours en évolution.

La technologie du Cloud Computing en est encore à ses débuts, ce qui en fait un excellent candidat pour une étude plus approfondie. Ainsi il existe de nombreuses définitions courantes du CC et la plus connue est celle du National Institute of Standards and Technology (NIST).

« Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models. » [46]

Le Cloud Computing permet aux infrastructures logicielles et matérielles d'être adaptables et évolutives grâce à des services virtuels partagés via un réseau de centres de données. Cette technologie permet à ses utilisateurs d'accéder et de déployer des applications à partir d'Internet en fonction de leurs demandes et de la qualité de service souhaitée [47].

3. Modèles de services cloud

Dans le monde des services cloud, les trois principales catégories sont IaaS, SaaS et PaaS [48].

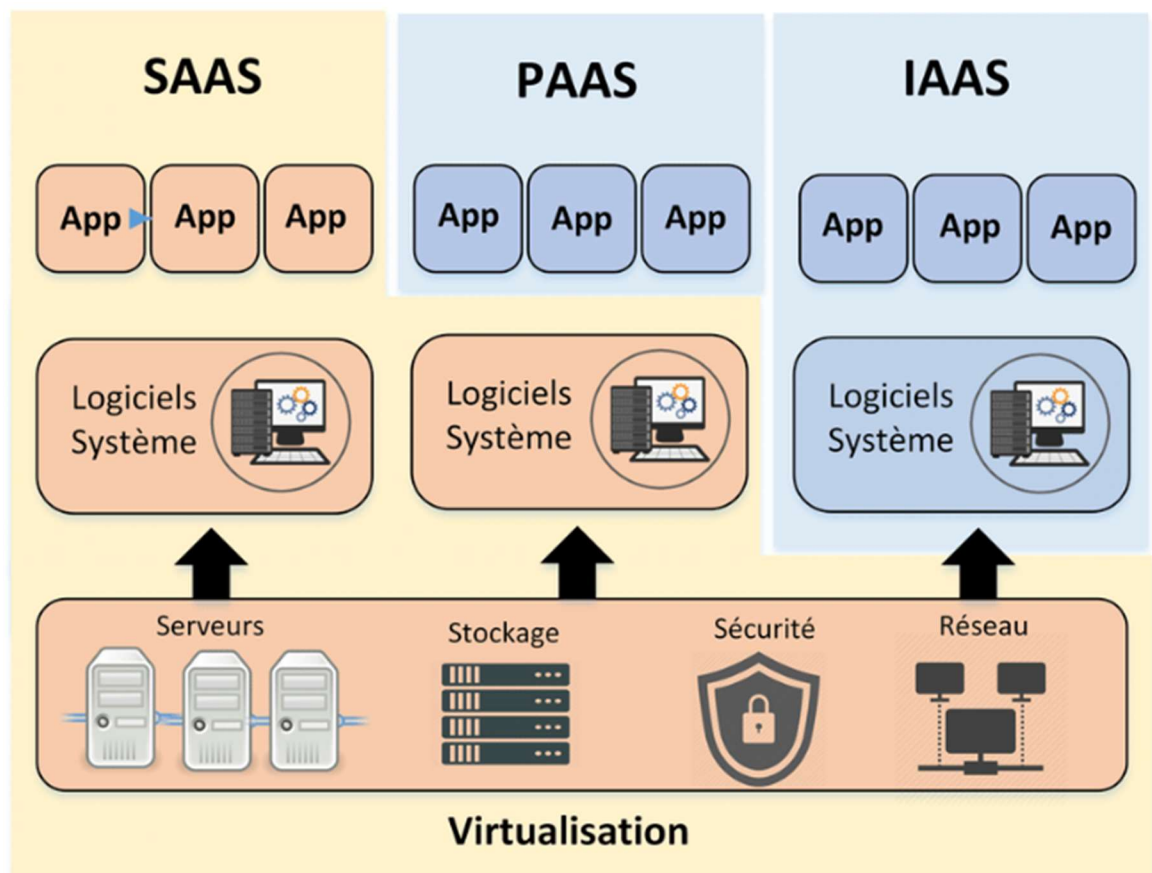


FIGURE 2 : LES MODÈLES DE SERVICE DU CLOUD COMPUTING [49]

3.1. SaaS (Software as a Service)

Cette architecture offre le service comme forme principale. Le logiciel est accessible depuis n'importe quel terminal disposant d'une connexion Internet. Les clients utilisent la plate-forme logicielle et matérielle du fournisseur de cloud SaaS, que le fournisseur gère lui-même sans s'occuper de la pile en dessous ni de

l'installation du logiciel en question. C'est par exemple le cas des services de messagerie électroniques (Outlook).

Cependant, certaines applications permettent aux clients de modifier leurs fonctionnalités en fonction de leurs besoins. L'application n'est ni contrôlée ni gérée par l'utilisateur final au lieu de cela, elle est gérée par le fournisseur de services.

3.2. PaaS (Plateforme as a Service)

Dans le modèle PaaS, les fournisseurs de services proposent des plateformes d'hébergement d'applications. Les fournisseurs de service sont disponibles sur une plateforme grâce aux langages, bibliothèques et services supportés par celle-ci. Cela permet aux développeurs d'utiliser les outils et les ressources disponibles pour créer et déployer leurs applications.

L'utilisateur PaaS n'a accès qu'à l'application qu'il déploie ; sa capacité à configurer la RAM et un CPU est limitée par le fournisseur PaaS. Il ne peut pas modifier l'infrastructure de son application, y compris les machines physiques ou virtuelles qu'elle utilise.

PaaS fixe ses tarifs d'utilisation aux applications en fonction de leur configuration technique, des services utilisés et du temps de déploiement de l'application.

Quelques exemples de PaaS sont : Heroku, Google AppEngine, AWS Elastic Beanstalk.

3.3. IaaS (Infrastructure as a Service)

Infrastructure as a Service (IaaS) est la combinaison des ressources informatiques de base, telles que le matériel, le réseautage et le stockage. Cette combinaison de ressources rend l'innovation cloud possible.

L'utilisateur d'IaaS est responsable du contrôle et de la gestion des ressources qu'il déploie. Par exemple, pour un serveur, il est responsable des mises à jour de sécurité du système d'exploitation, garantissant que le serveur est toujours en cours d'exécution, etc [50].

Exemples d'IaaS : Amazon EC2 (Web Services Elastic Compute Cloud) et AmazonS3 (Secure Storage Service), CloudSigma, RackSpace (Cloud Server et CloudFiles).

4. Modèles de déploiement

Chaque système nécessite des protections différentes en matière de tarification, de sécurité et de protections des données personnelles. De plus, les exigences changent en fonction du programme avec lequel il est utilisé. Il existe deux principaux types de services de CC qui se déclinent en versions publiques et privées. De plus, deux autres types de cloud dérivent des principaux à savoir le Cloud communautaire et le Cloud hybride [51].

4.1. Les Clouds publics

Les Clouds publics sont composés de matériel, de logiciels, de réseaux, de stockage et de diverses interfaces appartenant à des tiers et sont exploités par des personnes et des entreprises extérieures. Leurs services sont accessibles sur Internet par toute personne disposant d'un mode de paiement. L'objectif est d'héberger des applications sur un environnement avec un grand nombre d'utilisateurs. Les Clouds publics sont viables parce qu'ils gèrent généralement des charges de travail relativement simples et répétitives comme par exemple, le courrier électronique.

Nous pouvons citer certains fournisseurs de services publics de Cloud Computing : Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform, Scaleway, OVH Cloud, Alibaba Cloud, IBM Cloud, Oracle Cloud, DigitalOcean, Heroku, Tencent Cloud, Linode.

4.2. Les Clouds privés

Les Clouds privés sont similaires aux cloud publics du point de vue infrastructures mais sont créés et gérés par un tiers pour l'usage exclusif d'une entreprise. L'infrastructure cloud est déployée au sein de l'entreprise et n'est accessible qu'à ses employés. Les services utilisent une connexion réseau privée pour accéder aux ressources appartenant au service informatique de l'entreprise. L'infrastructure de cloud privé peut être établie par un fournisseur de cloud public ou une entreprise professionnelle dans le domaine.

4.3. Les Clouds communautaires

Les Clouds communautaires sont une version étendue des cloud privées. Leur déploiement de ressources est dédié aux communautés, groupes d'entreprises ou organisations ayant des intérêts communs. Cependant les ressources peuvent être contrôlées par un tiers ou une ou plusieurs organisations de la communauté.

4.4. Les Clouds hybrides

Les Clouds hybrides sont un mélange de Cloud privé combiné à l'infrastructure de Cloud public. Les modèles Cloud hybrides ne sont utilisés que pour étendre les fonctions d'un cloud privé d'entreprise à un cloud public.

5. Avantages et inconvénients du Cloud Computing

5.1. Les avantages

Le Cloud Computing offre de nombreux avantages significatifs pour les individus, les entreprises et les organisations. Voici quelques-uns des avantages du CC [52] :

- Réduction des coûts d'infrastructure matérielles et logicielles.
- Des ressources et services plus rapides à allouer et plus simple à utiliser;
- Accès aux ressources plus flexible;
- Utilisation plus efficace des ressources;
- Augmentation de la puissance de calcul;
- Grande capacité de stockage (quasi illimitée);
- Moins de problèmes d'entretien;
- Gestion simple et rapide des mises à jour des systèmes;
- Gestion de données personnelles;
- Disponibilité des infrastructures;
- Tout est considéré comme un service défini par entente de niveau de services (SLA);
- Réduction du temps de mise sur le marché.

5.2. Les inconvénients

Le Cloud Computing n'a pas que des avantages, il possède quelques obstacles et inconvénients qui sont abordés dans [52]. Parmi ces obstacles, il y a :

- Confidentialité des données;

- Chiffrement des données;
- Nécessité d'un accès réseau constant;
- Mauvais fonctionnement avec les connexions à basse vitesse;
- Faible niveau de la qualité de service dans le réseau;
- Risque d'engorgements lors des transferts de données;
- Problème d'interopérabilité;
- Problème de portabilité;
- Faible contrôlabilité;
- Manque de fonctionnalités d'audit.

6. Les caractéristiques du Cloud

Nous citons 5 caractéristiques essentielles du Cloud Computing.

6.1. Libre-service à la demande

Grâce à cette caractéristique, les utilisateurs peuvent contrôler en permanence la disponibilité des serveurs, le stockage réseau alloué et la capacité de calcul, assurant ainsi une surveillance continue.

6.2. Large accès au réseau

Les services informatiques utilisent généralement divers appareils et réseaux standard.

6.3. Élasticité rapide

Elle permet aux utilisateurs de dimensionner rapidement leurs ressources informatiques en fonction de leurs besoins changeants.

6.4. Mise en commun des ressources

Les ressources partagées telles que les serveurs, le stockage et les applications sont disponibles pour plusieurs utilisateurs sans pression. Une seule ressource peut fournir des services à plusieurs clients à la fois.

6.5. Services mesurés

Chaque application et ses utilisateurs sont suivis pour l'utilisation des ressources. Cela aide à la fois l'utilisateur et le fournisseur de la ressource à suivre leur utilisation. Ceci est utile aux deux parties en raison de la nécessité de surveiller la facturation et d'utiliser efficacement les ressources.

7. Cloud Computing et Big Data

Le Cloud permet aux utilisateurs d'accéder aux ressources informatiques nécessaires (comme le stockage et la puissance de calcul) sans avoir à investir dans l'achat de matériel ou de logiciels. Cette absence d'investissement initial permet aux utilisateurs de démarrer rapidement des projets sans se soucier de la mise en place d'une infrastructure physique.

L'une des principales caractéristiques du Cloud est sa capacité à évoluer sur plusieurs machines, ou son élasticité. L'élasticité du Cloud permet aux utilisateurs d'ajouter ou de réduire les ressources en temps réel en fonction des besoins changeants. Cela offre une flexibilité essentielle pour gérer des charges de travail variables, comme c'est souvent le cas dans le domaine du Big Data.

Étant donné que les ensembles de données du Big Data sont vastes et nécessitent une puissance de calcul importante, le Cloud offre la plate-forme idéale pour gérer, stocker et analyser ces données de manière efficace et évolutive.

L'expansion de l'informatique mobile a conduit à une augmentation significative du nombre de consommateurs, de professionnels et d'entreprises qui créent et accèdent à des données via des services basés sur le Cloud. Cela souligne comment le Cloud facilite l'accès et le partage de données à partir de divers appareils et endroits.

Google et Amazon.com utilisent le stockage et le traitement de données massives, appelé Big Data, pour améliorer leurs services. Par exemple Amazon.com utilise le Big Data pour personnaliser les recommandations de produits que vous voyez lorsque vous magasinez en ligne. Ces entreprises ont besoin de beaucoup de puissance informatique pour cela. Au lieu d'acheter et de gérer d'énormes quantités de serveurs elles-mêmes, elles utilisent le Cloud. Le Cloud leur permet de "louer" autant de puissance de calcul et de stockage que nécessaire, quand elles en ont besoin.

8. Conclusion

Pour clore ce chapitre, nous avons fourni une compréhension générale des idées clés et des points clés du Cloud Computing. Le CC a donné naissance à une nouvelle façon d'utiliser les technologies de l'information qui change radicalement la manière dont les entreprises investissent dans leur infrastructure. IL promet un stockage illimité à faible coût avec une puissance de calcul élevée. Mais, le Cloud Computing fait face à plusieurs obstacles qui doivent être résolus. La sécurité et la confidentialité des données personnelles sont des préoccupations majeures qui suscitent le scepticisme des utilisateurs [53].

CHAPITRE 5 : CONTEXTE D'ÉTUDE

1. Introduction

Dans ce chapitre, nous introduirons la reconnaissance faciale en mettant en évidence son importance dans différents domaines. Nous examinerons ensuite les fondements théoriques de cette technologie, en présentant différentes approches utilisées. Ensuite, nous explorerons l'architecture des réseaux de neurones artificiels, en nous concentrant particulièrement sur les réseaux de neurones convolutifs, avec une étude détaillée du modèle LeNet-5. Nous discuterons également des techniques qui peuvent être utilisées pour améliorer les performances des modèles de reconnaissance faciale.

2. Reconnaissance faciale

Ces dernières années, la reconnaissance faciale a fait l'objet d'une grande attention et est considérée comme l'une des applications les plus prometteuses dans le domaine de l'analyse d'images. La reconnaissance faciale est une méthode de reconnaissance d'un visage humain grâce à la technologie [1, 6, 54].

Elle est largement utilisée dans différents domaines tels que la sécurité, les guichets automatiques, l'identification criminelle, la vidéoconférence, la délivrance de permis de conduire et de passeports, ainsi que la surveillance des espaces publics. Par exemple, le Canada utilise un logiciel de reconnaissance faciale pour identifier les problèmes de sécurité publique comme les émeutes et surveiller les personnes dans les lieux publics comme les aéroports et les supermarchés.

Les systèmes de reconnaissance de visages (FR) nécessitent différentes fonctionnalités tout dépendamment du mode de fonctionnement du système et des données utilisées. La reconnaissance des visages satisfait les applications d'identification puisque le visage de l'utilisateur peut être capté à son insu et à distance [55].

2.1. Architecture de reconnaissance faciale

Le processus de reconnaissance faciale comporte généralement cinq étapes interdépendantes illustrées à la figure 3. La première étape consiste à acquérir l'image d'entrée, soit à partir d'une caméra ou d'une base de données. L'image est analysée et passe au modèle de détection de visage. Dans cette étape, le visage de la personne est détecté à partir de l'image. Une fois que le visage est détecté, il est

aligné afin de normaliser leur position et leur orientation. En outre, la troisième étape est l'extraction des caractéristiques faciales telles que les yeux, la forme du visage, etc. Ainsi dans cette étape, des représentations numériques sont générées. De même, la correspondance faciale est effectuée en utilisant des algorithmes de reconnaissance faciale. Les résultats de la correspondance détermineront l'identité du visage dans l'image d'où l'étape de la classification d'images.

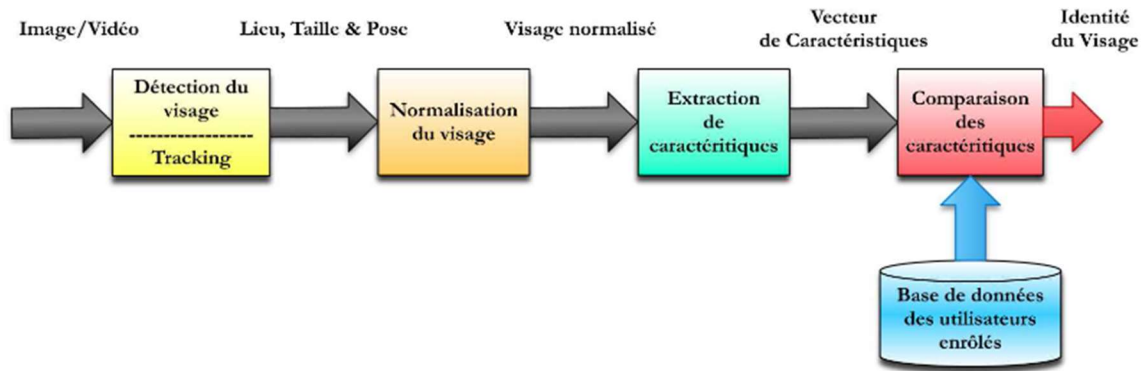


FIGURE 3 : PROCESSUS DE RECONNAISSANCE FACIALE [56].

2.2. Les types de représentations

Afin de vérifier l'identité d'un individu, un système de classification doit être entraîné. Ce processus consiste à décomposer le visage d'une personne en parties plus petites qui sont plus faciles à retenir [57].

Les systèmes FR peuvent être divisés en trois grandes approches afin de faciliter ce processus. : les approches holistiques qui sont une représentation entière du visage sur un ensemble de caractéristiques discriminantes d'un visage à l'autre, les approches locales qui résultent de l'extraction des caractéristiques particulières à certains points d'intérêt du visage et enfin les approches hybrides qui combinent les deux approches précédentes.

2.2.1. Les approches holistiques

Appelées aussi approches globales, elles se concentrent sur l'identification des caractéristiques du visage qui peuvent être différenciées. Cela permet de créer un classificateur qui peut étiqueter les visages comme appartenant à un groupe particulier.

Leur avantage est la conservation implicite des informations de texture et de la forme pour la reconnaissance faciale, en plus des aspects d'organisation structurelle globaux de la face. Néanmoins, l'inconvénient majeur réside dans la dimension très grande de l'espace de l'image ce qui reflétera négativement sur la classification [58].

À titre d'exemples d'approches holistiques nous citons PCA, LDA, eigenfaces, LBP et HOG dans [59].

2.2.2. Les approches locales

Les approches locales déterminent quelles caractéristiques faciales spécifiques sont les plus efficaces pour distinguer un individu. Ceci est fait afin de résoudre le problème de la représentation des visages. La difficulté principale liée à ces techniques consiste à trouver une manière adéquate de représenter les interrelations qui existent entre ces caractéristiques, qu'il s'agisse de positions relatives sur le visage, d'interconnexions géométriques, ou de regroupements par graphes

Parmi les approches les plus répandues, nous énumérons les méthodes HMM, DCP [60] et les approches purement géométriques.

2.2.3. Les approches Hybrides

Cette approche est un résultat d'une fusion entre deux autres approches : approche holistique ou globale et l'approche locale afin d'améliorer les performances des systèmes de reconnaissance. En effet les caractéristiques locales et globales sont complètement différentes. Chacune de ces méthodes a ses inconvénients, par ailleurs l'une peut être complémentaire de l'autre dans le but d'améliorer la classification.

2.3. Les domaines d'applications

La reconnaissance faciale offre une multitude d'application dans divers domaines. Tout d'abord, dans le secteur de la sécurité, elle est utilisée pour l'identification des individus dans les systèmes de vidéosurveillance, permettant une détection rapide des personnes recherchées ou suspectes. De plus, dans le domaine de la technologie mobile, la reconnaissance faciale est utilisée pour déverrouiller les smartphones et garantir un accès sécurisé aux données personnelles. Par ailleurs, dans le secteur médical, cette technologie est utilisée pour l'identification des patients, facilitant la gestion des dossiers médicaux et l'accès aux informations pertinentes. De même, dans le domaine de la gestion de l'identité, la reconnaissance faciale est utilisée dans les aéroports et les postes frontaliers pour vérifier l'identité des voyageurs et faciliter les procédures d'embarquement. Enfin, la reconnaissance faciale est également utilisée dans le domaine du marketing, où elle permet de

personnaliser les publicités en fonction des caractéristiques faciales des consommateurs.

3. Les réseaux de neurones

Un réseau de neurones est un système informatique qui est analogue au fonctionnement des neurones du système nerveux humain. Parmi les différents types de réseaux de neurones, le perceptron multicouche est le plus populaire. Il a été introduit en 1957 par Frank Rosenblatt [61].

Le réseau de neurones est généralement organisé en trois parties successives. Ces couches sont les suivantes :

- La couche d'entrée : elle reçoit et traite toutes nos données brutes fournies en entrée. Nos données sont sous forme vectorielle [62];
- La/ les couche(s) cachée(s) : est constituée d'unités qui transforment nos données d'entrée en informations utiles pour la couche de sortie. Un réseau peut contenir une ou plusieurs couches cachées [62];
- La couche de sortie: qui est la couche ou les informations traitées sont présentées. Elle donne les résultats du calcul interne [62].

Chaque couche du réseau comprend un ou plusieurs neurones. Ce sont des entités qui traitent et transmettent les informations travers le réseau.

L'architecture d'un réseau de neurones est présentée à la figure suivante. La couche d'entrée comprend 2 neurones, les couches cachées comprennent 10 neurones et la couche de sortie 1 neurone.

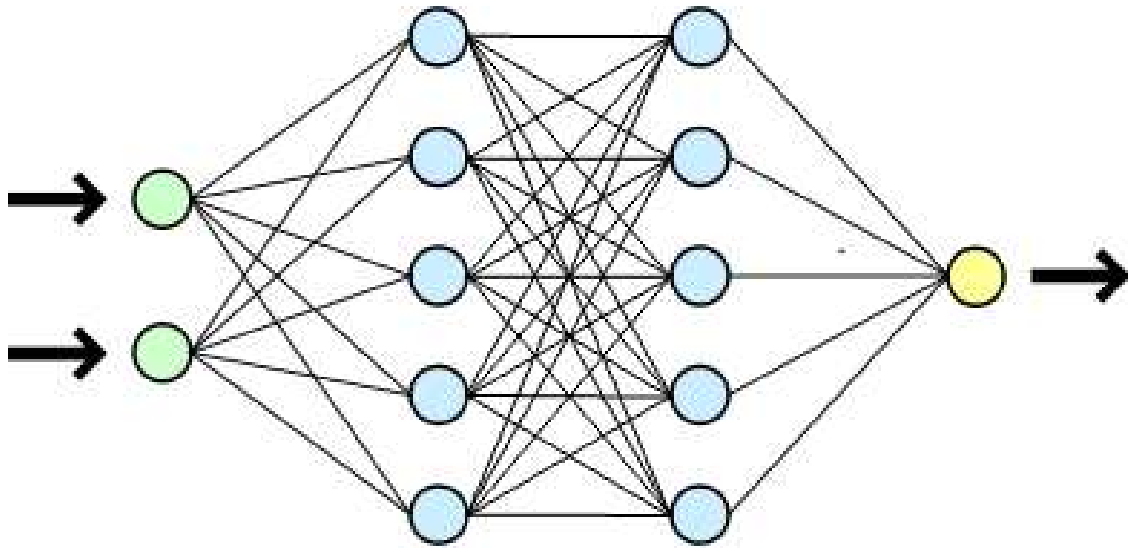


FIGURE 4 : ARCHITECTURE D'UN RÉSEAU DE NEURONES [63].

Les réseaux de neurones sont particulièrement adaptés pour l'apprentissage automatique à partir des données, qu'il s'agisse de l'apprentissage supervisé ou non supervisé.

- ✓ L'apprentissage supervisé est utilisé lorsque nos données fournies au réseau sont étiquetées. Un problème de classification se présente lorsque nos données de prédiction sont regroupées en classes distinctes. Si par contre, nos données sont des valeurs quantitatives, un problème de régression intervient.
- ✓ L'apprentissage non supervisé correspond au cas où nous ne disposons pas de bases d'apprentissage du réseau, c'est à dire nos données dont les sorties ne sont pas connues. Dans ce cas, l'apprentissage repose sur un critère interne qui évalue le comportement du réseau par rapport à des spécifications générales et non sur des observations.

3.1. Fonction d'activation

Chaque couche de neurones est reliée aux neurones des autres couches par des poids synaptiques. Ces poids sont des valeurs aléatoires qui modifient la quantité d'influence des sorties des neurones. Tous les poids sont multipliés par leurs valeurs d'entrée correspondante, puis additionnés pour former une somme pondérée. Le biais est ensuite ajouté à la somme pondérée, puis le résultat est transformé par une fonction appelé fonction d'activation.

La fonction d'activation est une fonction mathématique qui s'applique uniquement à la sortie des neurones des couches cachées [64].

Quelques exemples de fonctions d'activation couramment utilisées dans les réseaux de neurones sont la fonction sigmoïde, la fonction tangente hyperbolique, la fonction Relu (Rectification Linéaire), etc.

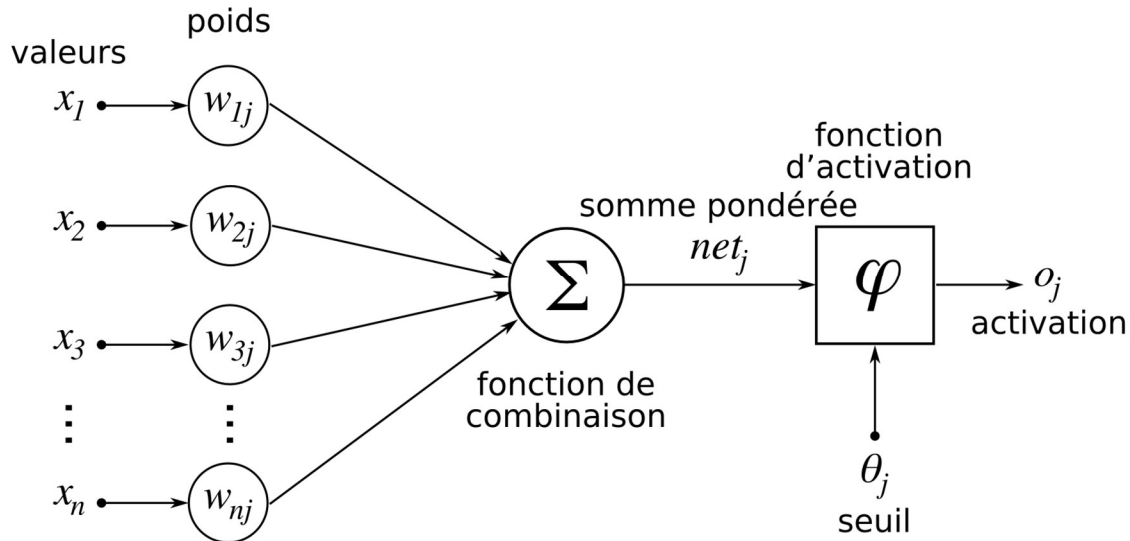


FIGURE 5 : UN NEURONE ARTIFICIEL [65].

3.2. Algorithme d'apprentissage

Lors de l'apprentissage supervisé, des données formées d'observations et d'étiquettes sont fournies au réseau de neurone. Ces données sont transmises à travers les différentes couches du réseau pour générer une prédiction en sortie.

L'algorithme d'apprentissage analyse les données d'entrée et compare les prédictions du réseau avec les étiquettes fournies. Il calcule aussi la différence entre les prédictions et les sorties réelles, également appelé erreur. La fonction de perte détermine cette erreur.

3.3. Rétropropagation

Le processus de rétropropagation commence en calculant l'erreur de prédiction par rapport aux poids synaptiques des neurones. Une fois l'erreur de prédiction calculée, l'algorithme d'optimisation intervient pour la quantifier. Elle utilise ainsi la technique de descente du gradient.

Parmi les algorithmes d'optimisation on retrouve la descente de gradient stochastique, l'algorithme d'optimisation par Adam ou encore l'algorithme d'optimisation par RMSprop.

3.4. Bases de données

Dans le domaine de l'apprentissage automatique, une base de données est généralement divisée en trois ensembles : l'ensemble d'entraînement, l'ensemble de validation et l'ensemble de test.

L'ensemble d'entraînement et l'ensemble de validation sont utilisés lors de la phase d'entraînement de notre réseau. L'ensemble d'entraînement est utilisé pour entraîner le réseau et l'ensemble de validation est utilisé pour évaluer les performances du réseau pendant l'entraînement. L'ensemble de validation est utilisé pour ajuster les hyperparamètres du modèle tels que le taux d'apprentissage, le nombre d'itération afin d'obtenir les meilleures performances.

Quant à l'ensemble de test, il est utilisé pour évaluer les performances du modèle sur les nouvelles données.

Pour atteindre une capacité de généralisation satisfaisante, un modèle doit obtenir de bons résultats à la fois sur l'ensemble d'entraînement et sur l'ensemble de validation. Une fois que le modèle a atteint cet objectif, il est recommandé à l'utilisateur de terminer l'apprentissage (voir figure).

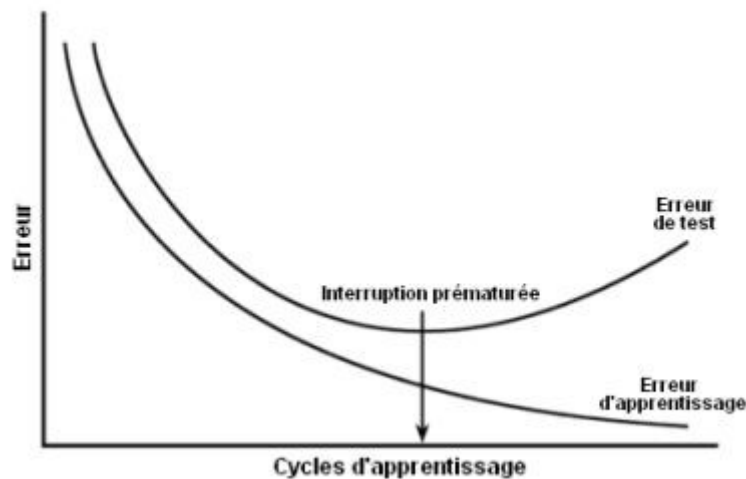


FIGURE 6 : ÉVOLUTION DE LA COURBE D'ERREUR DURANT LA PHASE D'APPRENTISSAGE [63].

3.5. Quelques notions

Le terme époque est définie comme la période de temps nécessaire pour passer complètement les ensembles de données à travers les réseaux de neurones.

4. Présentation des réseaux de neurones convolutifs (CNN)

Le réseau de neurones convolutifs, également appelé CNN, est une architecture neuronale profonde conçue pour travailler avec des images et des vidéos. Les CNN sont récemment devenus populaires en raison de la grande quantité de données sur Internet ainsi que de la puissance des processeurs informatiques et de la puissance de calcul.

Un CNN est composé de trois types de couches: convolution, mise en commun et couches entièrement connectées [66]. Les couches de convolution et de mise en commun effectuent l'extraction des caractéristiques tandis que la couche entièrement connectée mappe les caractéristiques extraites dans la sortie finale. Une couche est un bloc de construction.

Les différentes couches seront expliquées dans les sections suivantes.

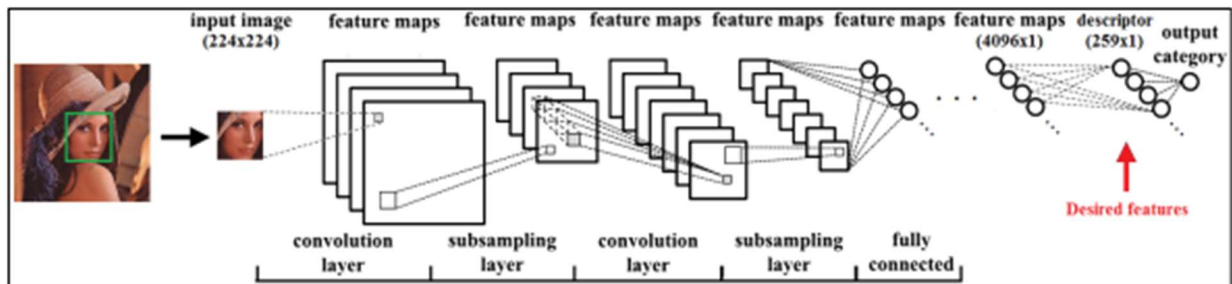


FIGURE 7 : ARCHITECTURE D'UN RÉSEAU DE NEURONES CONVULUTIFS [67].

4.1. Les images RVB

Les images sont généralement décomposées en trois canaux primitifs : Rouge, Vert et Bleu; c'est à dire qu'elles sont représentées sous trois cellules de couleurs de taille identique. Chaque cellule de la matrice contient la valeur du pixel. Ces pixels constituent le niveau des couleurs.

Ces valeurs sont comprises entre 0 et 255 où chacun de ces nombres représente l'intensité des pixels. Les lignes représentent la largeur de l'image, les colonnes correspondent à la hauteur de l'image et le nombre de canaux correspond à la profondeur de l'image.

4.2. Les couches de convolutions

L'architecture d'un réseau de neurones convolutifs repose sur une ou plusieurs couches convolutives, qui sont connectées aux couches précédentes de façon sélective. La couche de convolution est une combinaison d'une opération de convolution et d'une fonction d'activation.

Les convolutions sont des opérations linéaires de somme multipliée par élément entre une section de l'image d'entrée et la matrice de poids ou le filtre. Lors de l'opération de convolution, la matrice de poids parcourt l'image de sorte que tous les pixels soient couverts au moins une fois pour donner une sortie convoluée [68]. Cette sortie convoluée est la carte de caractéristiques.

La matrice de poids ou le filtre extraient les informations particulières de la matrice d'image d'entrée. Chaque filtre a une fonction spécifique. C'est à l'utilisateur de définir la taille du filtre et le nombre de filtre et, de les définir comme paramètres.

Le filtre se déplace sur toute l'image d'entrée en faisant des pas(stride). Un stride est la distance entre deux positions successives du filtre. Ainsi si le stride se déplace d'un pixel à la fois, nous l'appelons un stride de 1. Le stride 1 est le choix commun lors d'une convolution.

La taille de l'image diminue progressivement à mesure que la valeur du stride augmente. Pour pallier à ce problème, la méthode de rembourrage zéro est utilisée à l'image d'entrée. Le rembourrage est utilisé pour conserver les dimensions spatiales de l'image d'entrée après des opérations de convolution sur une carte d'entités.

Dans un CNN, les premières couches extraient les caractéristiques les plus génériques, tandis que, à mesure que le réseau s'approfondit les dernières couches obtiennent les caractéristiques cachées. En sortie de chaque couche de convolution, une fonction d'activation est finalement introduite [69]. Cette dernière est une fonction de non-linéarité permettant au réseau de neurones de s'adapter facilement à une variété de données.

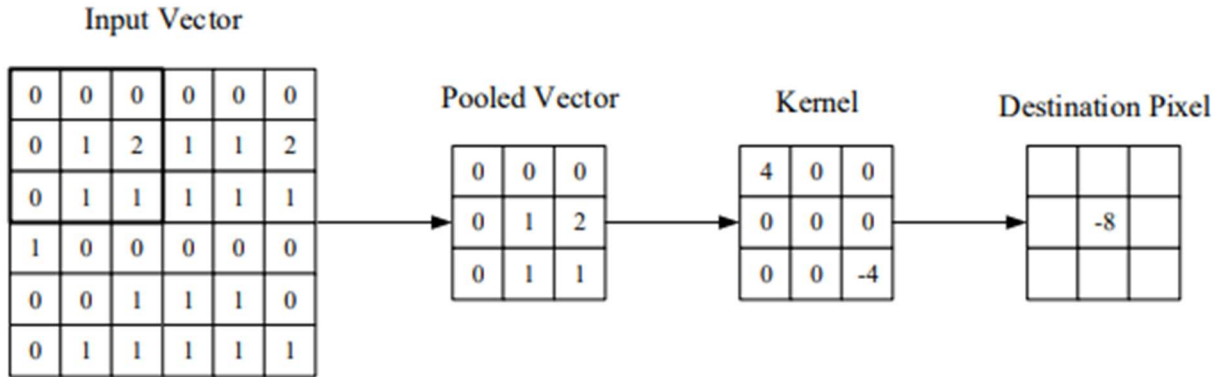


FIGURE 8 : UNE REPRÉSENTATION VISUELLE D'UNE COUCHE CONVOLUTIVE [70].

4.3. Couche de mise en commun

La couche de mise en commun est souvent placée entre deux couches convolutives. Cette couche reçoit des entrées de plusieurs caractéristiques pour simplifier les données. Un sous-échantillonnage est effectué sur ces données; ce qui réduit la quantité de puissance de calcul requise pour traiter les données.

Les couches de regroupement se connectent aux neurones des sorties de la couche précédente en utilisant le même nombre de neurones que les couches de convolution. Leur fonctionnement est d'autant plus simple qui consiste à renvoyer le résultat d'une fonction d'agrégation soit maximale ou moyenne comme le montre la figure 9.

La mise en commun peut être divisée en deux types : la mise en commun maximale et la mise en commun moyenne. Un atout de la mise en commun est l'amélioration du réseau tout en éliminant le problème de surapprentissage. Cependant la perte d'information suite à la réduction est l'inconvénient majeur des couches de mise en commun. Cette problématique est résolue en utilisant les petits filtres 2 x 2 dans l'exemple illustré dans la figure.

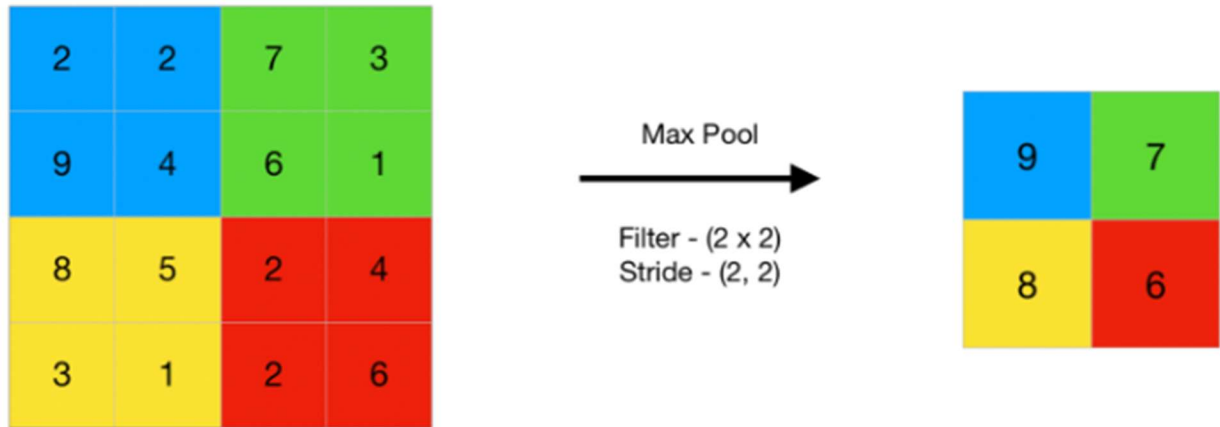


FIGURE 9 : OPÉRATION DE MISE EN COMMUN [71].

4.4. Les couches entièrement connectées

Après avoir effectué une série d'opérations de regroupement et de convolution, nous créons des couches entièrement connectées. C'est ce que nous appelons opération d'aplatissement. Elle consiste à condenser les aspects de notre image en un seul vecteur (Figure 10).

Nous classons l'image de sortie à l'aide d'un perceptron multicouche qui comprend des couches entièrement connectées. Nous pouvons utiliser avec la dernière couche entièrement connectée une fonction « Softmax » ou « Relu » pour la génération des prédictions des classes.

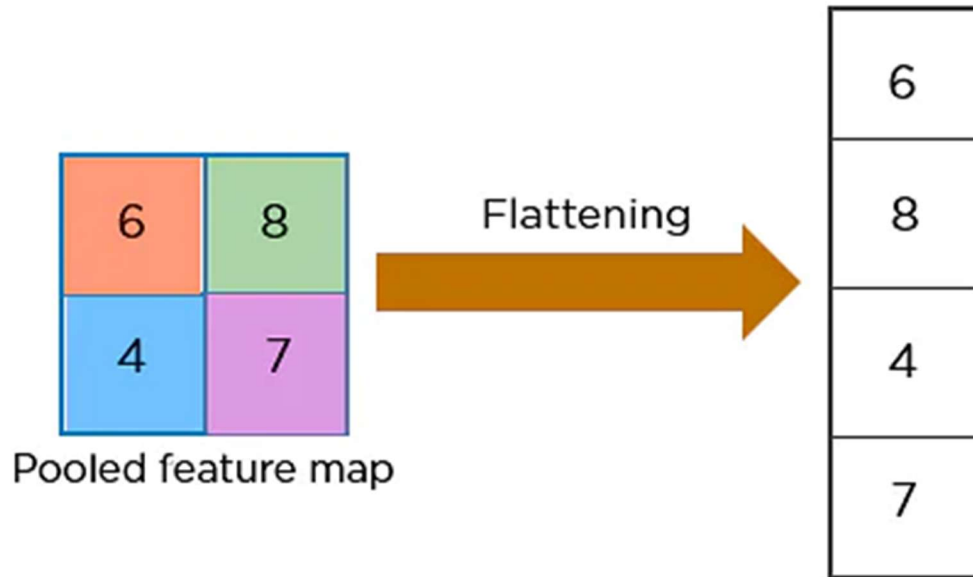


FIGURE 10 : OPÉRATION D'APLATISSEMENT [72].

4.5. Présentation des réseaux de neurones LeNet-5

LeNet-5 est le plus simple et le plus modeste des réseaux de neurones de convolutions. Présenté par Yann LeCun [4], il s'agit d'un réseau de convolutions à Cinq (5) couches pour la classification; ce qui en fait un choix idéal en tant que premier CNN.

Ce modèle prend en entrée des images de tailles (32x32x1). Il est constitué de trois ensembles de couches de convolution (voir figure 11) accompagnés d'une couche de mise en commun moyen. Après les couches de convolution et de mise en commun moyen, nous avons deux couches entièrement connectées. Un filtre de taille 5x5 de même qu'un pas (stride) de taille 1 et un rembourrage de taille 1 sont utilisés pour chaque couche de convolution.

La fonction d'activation Relu [73] est utilisée pour chaque couche de convolution et la technique du zéro rembourrage, de manière à conserver la taille de nos images après chaque couche de convolution.

L'architecture LeNet-5 est divisée en deux parties : le processus de convolution et le processus de réseaux de neurones.

Notre processus convolutif est constitué de trois blocs de convolutions constituant chacun une couche de convolution et ont successivement des profondeurs de 6, 16 et 16.

L'opération d'aplatissement est utilisée à la suite du dernier bloc de couches de convolution. Ensuite nous avons le classificateur qui comprend 3 couches. La première couche comporte 120 neurones. La suivante est une couche entièrement connectée avec 84 neurones qui aboutissent à la sortie à 84 valeurs et la fonction d'activation utilisée ici est Relu.

La dernière couche est la couche de sortie avec 10 neurones et la fonction Softmax qui classe les images dans leur classe respective [74].

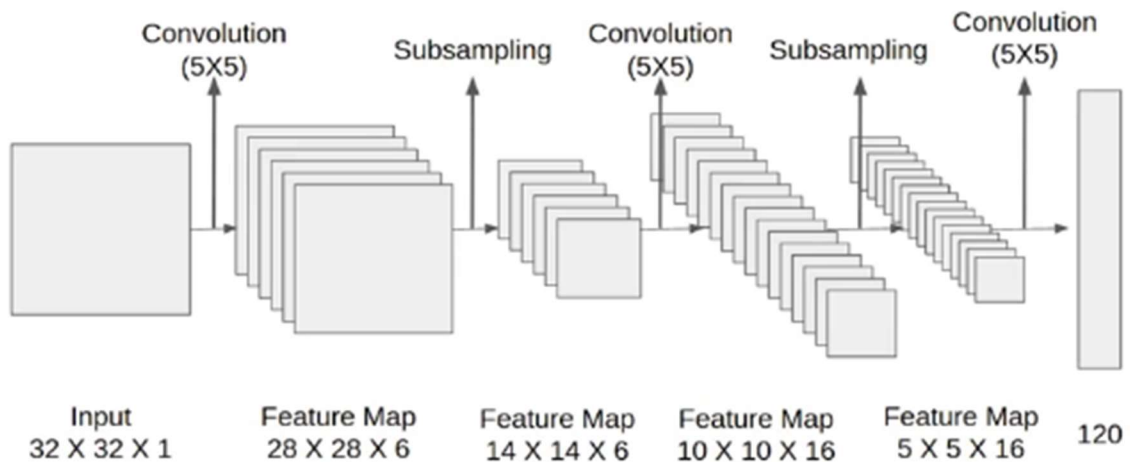


FIGURE 11 : ARCHITECTURE LENET-5 [75].

5. Amélioration du modèle

5.1. Augmentations des données

Avoir un grand ensemble de données est crucial pour la performance d'un modèle de réseau de neurones convolutifs. Cependant, nous pouvons améliorer les performances du modèle en augmentant les données dont nous disposons déjà. Cela aide également à généraliser sur différents types d'images. L'augmentation des données est une technique qui utilise les échantillons d'apprentissage existants pour générer davantage de données d'apprentissage [76]. Ceci est possible en définissant une variété de transformations aux données originales.

Dans l'augmentation, nos transformations englobent un ensemble d'opérations dans le domaine de la manipulation d'images, telles que le zoom, la rotation, le redimensionnement et diverses autres transformations.

5.2. Dropout

Dropout est une approche de régularisation utilisée dans les réseaux de neurones qui empêche le problème de sur-ajustement [77]. Cette technique rend le processus d'entraînement "bruyant" en les neurones de couche cachée à gérer les entrées de manière aléatoire, avec une probabilité variable. Par exemple, lorsqu'un "Dropout" avec une probabilité de 0.25 est appliqué à une couche cachée, cela signifie que pendant la phase d'apprentissage, 25% des neurones de cette couche sont désactivés de manière aléatoire.

5.3. Processus de normalisations par lots

Éviter le sur-ajustement est l'un des défis les plus fréquents en matière d'apprentissage automatique. Une solution à un tel défi est la régularisation. Les techniques de régularisation contribuent à l'amélioration d'un modèle en favorisant une convergence plus rapide. L'une des techniques de régularisations est le processus de normalisation par lots.

La normalisation par lots [78] est un processus qui améliore notre modèle en ajoutant des couches supplémentaires dans notre approche de réseau de neurones profond. La nouvelle couche effectue les opérations de standardisation et de normalisation sur l'entrée d'une couche provenant d'une couche précédente.

5.4. Présentation de la taille des lots

La taille du lot (Batch Size) définit le nombre d'échantillons qui traversent généralement le réseau neuronal en une seule fois [79]. Elle est également appelée mini-lot. La taille d'un mini-lot varie de 1 jusqu'au nombre total d'exemples de nos données. Pour illustrer, dans le cas d'un jeu de données comprenant 50 images, la taille du mini-lot peut être choisie parmi des valeurs telles que 2, 6, 10, 24 ou toute autre valeur comprise entre 1 et 50. Plus la taille du lot est élevée, plus nous aurons besoin de plus d'espace mémoire.

5.5. Taux d'apprentissage

Le taux d'apprentissage (Learning Rate) est un hyperparamètre qui régule la manière dont notre modèle est modifié en réponse à l'erreur estimée lors de la mise à jour des poids du réseau neuronal [80]. Le choix du taux d'apprentissage est un processus délicat, car une valeur trop petite peut entraîner une formation lente et qui pourrait être bloqué dans certains cas, tandis qu'une valeur plus grande entraîne un apprentissage rapide et non optimal des poids ou un processus d'entraînement instable. Il représente métaphoriquement la vitesse à laquelle un modèle d'apprentissage automatique apprend.

CHAPITRE 6 : MÉTHODOLOGIE DE RECHERCHE

1. Introduction

Ce chapitre décrit les trois parties principales : la première est la description des outils que nous avons utilisés pour concevoir notre modèle; la deuxième aborde les démarches de collecte et d'organisation des données, et la dernière est la formation de notre réseau de neurine et de son évaluation. Les trois étapes sont expliquées en détail ci-dessous.

2. Environnement logiciel et matériel de développement

Il existe différentes techniques pour effectuer du Machine Learning (Apprentissage Automatique), mais elles nécessitent toutes des ressources matérielles relativement importantes, en fonction des données à traiter et du contexte de recherche.

Dans notre cas, nos données fournissent suffisamment de caractéristiques pour leur analyse, ce qui implique la nécessité d'un ordinateur suffisamment puissant pour la réalisation des expérimentations.

2.1. Configuration Matérielle

Fabricant	HP
Processeur	CPU Intel® Core TM i7-1065G7 CPU @ 1.30GHz 1.50GHz
Mémoire physique (RAM) installée	8.00 Go
Nom du système d'exploitation	Windows 11 Famille
Version du système d'exploitation	21H1

Architecture	64 bits
Connexion	À distance en utilisant le protocole RDP (Remote Desktop Protocol)

TABLEAU 1 : CONFIGURATION MATÉRIELLE

2.2. Langage de programmation

Python [81] est un langage de programmation largement utilisé dans le domaine de la vision par ordinateur et de la science des données. Depuis sa création en 1989, le langage Python open source extensible est dynamique, compact, libre d'utilisation et modulaire. Il encourage une approche orientée objet de la programmation sans la forcer. C'est un langage très avancé pour les systèmes de reconnaissance faciale.

Son plus grand avantage est le large éventail de bibliothèques qu'il contient. Un autre de ses avantages est sa forte communauté qui pousse constamment à son développement. Pour toutes ces qualités, notre choix s'est tourné vers Python, ayant déjà fait ses preuves dans le domaine de la vision par ordinateur.

2.3. Présentation des bibliothèques

2.3.1. Numpy (Numerical Python)

Numpy [82] est la bibliothèque s'occupant des opérations scientifiques liées aux tableaux unidimensionnels et multidimensionnels. Il exécute un certain nombre de routines en utilisant un tableau multidimensionnel d'objets. Numpy est une bibliothèque rapide et polyvalente.

2.3.2. OpenCV (Open Computer Vision)

OpenCV [83] est l'une des bibliothèques open source les plus célèbres et les plus utilisées pour les tâches de vision par ordinateur telles que l'analyse d'images, la détection de visages, la reconnaissance faciale, etc. Son autre utilisation majeure est l'exécution de tâches d'apprentissage automatique.

OpenCV a été développé en 2002 par Intel et écrit en C++, ce pendant la bibliothèque a fourni des liaisons en Java et en Python. Il est facile à lire et à utiliser.

OpenCV fournit une large gamme de programmes et de composants grâce à ses diverses fonctionnalités, permettant de concevoir des interfaces graphiques de base.

2.3.3.Pandas

Pandas [84] est une bibliothèque open source largement utilisée par les scientifiques des données du monde entier pour effectuer des analyses et des manipulations de données. Elle repose sur le langage de programmation Python. Pandas est rapide, flexible et très puissant.

2.3.4.SciPy

La bibliothèque open source SciPy [85] est utilisée pour résoudre des problèmes d'ingénierie, mathématiques et scientifiques. Construit sur l'extension python Numpy, SciPy contient une variété de sous-packages pour différentes applications. SciPy est rapide et a une puissance de calcul élevée.

2.3.5.Scikit Learn

Scikit Learn [86] est l'une des meilleures bibliothèques pour l'apprentissage automatique en Python. Pour tout projet de science des données, l'utilisation de la bibliothèque Scikit Learn est indispensable. C'est un outil simple et très rapide pour l'analyse prédictive des données et la modélisation statistique.

Cette bibliothèque est construite en utilisant python au-dessus de NumPy, SciPy et matplotlib.

2.3.6.TensorFlow

TensorFlow [87] est une bibliothèque mathématique symbolique qui utilise le flux de données et la programmation différentiable pour effectuer de nombreuses tâches liées à la formation et à l'inférence sur les réseaux de neurones profonds. C'est une plate-forme open source qui permet aux développeurs de créer des applications d'apprentissage automatique à l'aide de divers outils, bibliothèques et ressources communautaires.

TensorFlow est la bibliothèque d'apprentissage en profondeur la plus célèbre de Google. Google utilise l'apprentissage automatique dans tous ses produits, y compris son moteur de recherche, le sous-titrage des images et les recommandations.

2.3.7. Matplotlib

Matplotlib [88] a été introduit par John Hunter en 2002. C'est une bibliothèque de visualisation étonnante en Python pour les tracés 2D de tableaux de données. Elle est l'une des bibliothèques de visualisation de données les plus utilisées au monde.

Matplotlib est simple et ne nécessite aucune compétence particulière. Elle est construite sur des tableaux Numpy et conçue pour fonctionner avec la pile SciPy plus large.

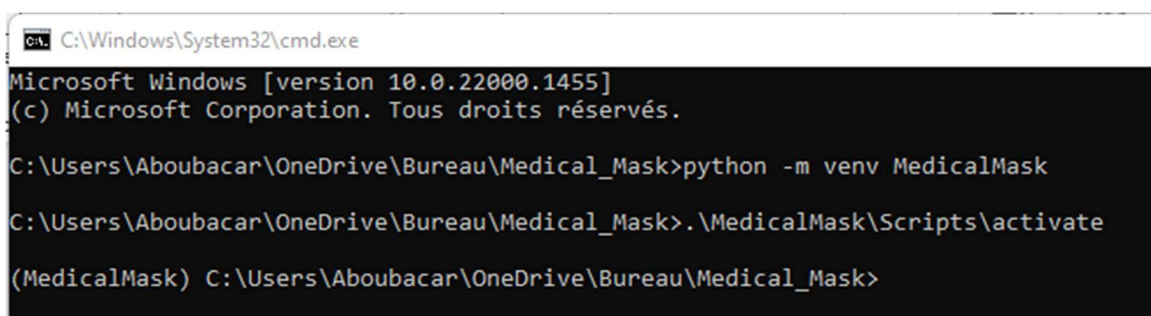
2.3.8. Jupyter

Jupyter [89] est l'environnement de développement intégré (IDE) choisi pour notre travail de recherche. C'est un IDE qui combine de manière transparente les bases du langage Python avec des fonctionnalités et des commodités supplémentaires disponibles dans un IDE. De plus, nous avons utilisé les frameworks CAFFE [90] permettant de réaliser des applications interactives depuis Python.

2.4. Environnement de développement

Étant donné que nous implémentons une application pour notre projet, nous avons créé et installé un environnement virtuel dans Python avec toutes les bibliothèques requises.

La figure illustre les commandes d'installation et d'activation de notre environnement de développement.



```
C:\Windows\System32\cmd.exe
Microsoft Windows [version 10.0.22000.1455]
(c) Microsoft Corporation. Tous droits réservés.

C:\Users\Aboubacar\OneDrive\Bureau\Medical_Mask>python -m venv MedicalMask
C:\Users\Aboubacar\OneDrive\Bureau\Medical_Mask>.\MedicalMask\Scripts\activate
(MedicalMask) C:\Users\Aboubacar\OneDrive\Bureau\Medical_Mask>
```

FIGURE 12 : COMMANDES D'INSTALLATION ET D'ACTIVATION.

3. Données

3.1. Collecte des données.

La première étape consiste à collecter l'ensemble de données afin que nous puissions préparer les images de visage pour former nos modèles. Pour cela, nous avons extrait le chemin du jeu de données Flickr-Face HQ (FFHQ) [5]. FFHQ est divisée en quatre (4) différentes catégories de visages de distribution uniforme. La première catégorie est « Mask » qui signifie des personnes avec le visage masqué, tandis que la deuxième catégorie « Mask_Chin » représente des personnes avec le menton uniquement couvert. Pour les individus dont le nez n'est pas couvert c'est la troisième catégorie « Mask_Mouth_Chin » et, la catégorie « Mask_Nose_Mouth » dont les personnes avec la bouche non couverte est la quatrième.

Étant donné que notre ensemble de données se subdivise en quatre (4) sous-ensembles, nous appliquons quatre étiquettes pour représenter les quatre classes.

3.2. Organisation des données

FFHQ se compose de 3000 images de faces de dimensions (300 x 300), dont 2700 images proviennent de l'ensemble d'entraînement et 300 images de l'ensemble de validation.

Nos données ont été organisées en quatre dossiers respectifs qui correspondent à nos 4 classes.



FIGURE 13 : EXEMPLES D'IMAGES DE LA BASE DE DONNÉES FFHQ [5].

3.3. Détection de visages

Une fois notre jeu de données collecté, labellisé et organisé, nous détectons les visages. La détection de visage est le processus d'identification du visage à partir de l'image capturée ou de l'image spécifiée à partir de la base de données. Pour une image donnée, nous identifions l'emplacement du visage.

Nous utilisons le modèle FP16 (Floating-Point), du module de réseau neuronal profond pré-entraîné de OpenCV appelé le DNN (Deep Neural Network) pour la détection du visage par implémentation CAFFE.

CAFFE (Convolutional Architecture for Fast Feature Embedding) est développé par Yangqing [90]. Il prend en charge différents types d'architectures de réseaux de neurones profonds spécialisés pour le traitement d'images, comme dans notre projet.

Nous disposons de deux fichiers nécessaires à l'exécution de CAFFE afin de charger les poids des couches du réseau FP16. Le premier fichier avec l'extension « .caffemodel » qui fournit les poids des couches intermédiaires lors de l'exécution du réseau et le second fichier, « .prototxt » qui définit l'architecture de l'algorithme FP16 au format texte lisible et modifiable.

Nous appliquons la fonction "cv2.dnn.blobFromImage" pour enregistrer l'entrée dans le modèle. Nous refactorisons l'algorithme en réduisant la taille et le poids avec les valeurs moyennes RVB (104,117,123) des données ImageNet [91].

Nous évaluons par la suite le score de précision qui détermine l'image contenant le visage détecté. Dans notre travail, ce score de confiance est supérieur à 0.8 (>0.8). Avec face_detection_model.forward(), nous déterminons les scores de confiance.

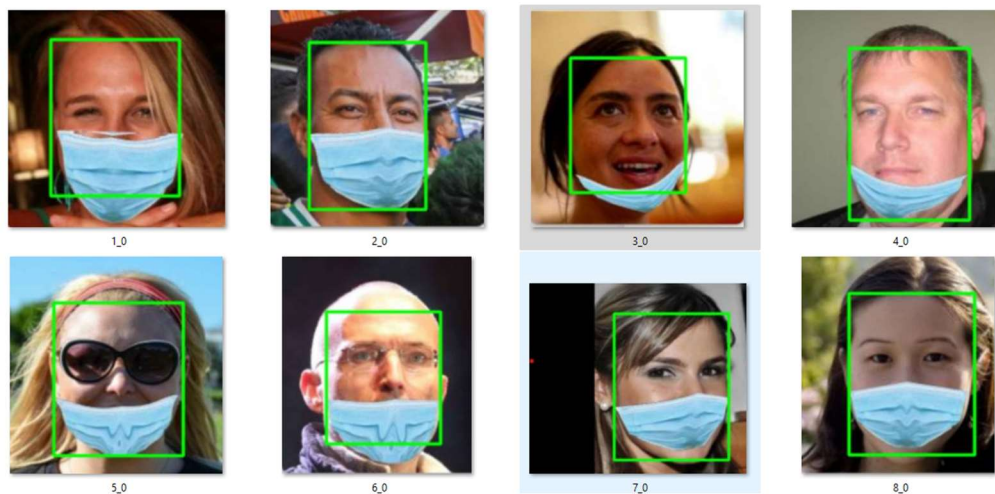


FIGURE 14 : EXEMPLES D'IMAGES DE VISAGES DÉTECTÉS.

3.4. Alignement de visages

L'alignement du visage est un processus de recadrage de la partie du visage à partir d'images et l'image recadrée représente les traits du visage.

Nous recadrons le visage une fois que nous le détectons dans le but d'obtenir la région d'intérêt de l'image qui contient le visage. Ce qui nous permet d'extraire les caractéristiques du visage afin qu'il soit facile pour notre modèle de s'entraîner.



FIGURE 15 : EXEMPLES D'IMAGES DE VISAGES RECADRÉS.

3.5. Extractions des caractéristiques (prétraitement des données)

Après avoir extrait les visages en sortie du processus, nous redimensionnons les images afin d'optimiser la base de données pour le traitement de notre modèle. Nous avons choisi la taille (100X100) dans notre projet.

Comme la fonction "cv2.dnn.blobFromImage" est utilisée pour l'entrée de notre modèle, plusieurs opérations sont faites avant d'obtenir cette entrée. Nous réduisons la dimension de l'image avec « np.squeeze(blob) ». L'image est ensuite normalisée puis retirée des valeurs négatives avec « np.maximum(blob_flip,0)/blob_flip.max() »

L'augmentation des données est la technique que nous avons abordée à la section pour la réalisation de ces opérations.

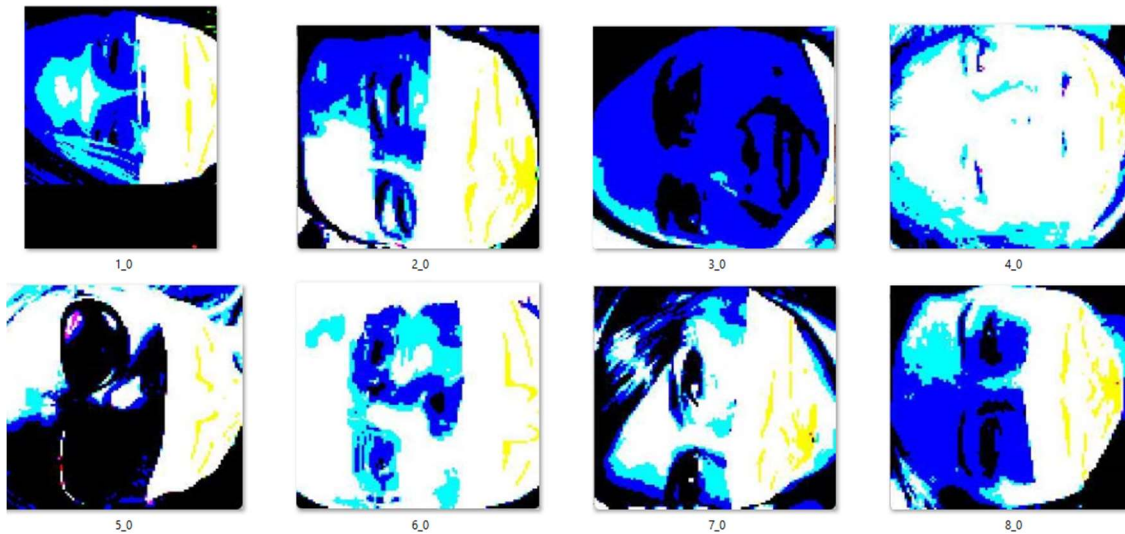


FIGURE 16 : EXEMPLES D'IMAGES APRÈS EXTRACTIONS DES CARACTÉRISTIQUES.

Après toutes les tâches de prétraitement, nous sauvegardons nos images prétraitées ainsi que nos étiquettes dans des tableaux Numpy. Cette bibliothèque prend en charge la manipulation des tableaux multidimensionnels. Elle fonctionne également avec des fonctions mathématiques relatives à la manipulation de matrices. Après cette étape, nous utilisons ces données comme entrées de notre réseau de neurones.

4. Formation des modèles de reconnaissance faciale

Nous entraînons trois types de modèles de CNN dans le but d'obtenir un modèle à la fois stable et performant. Nous chargeons nos données prétraitées pour la formation de nos modèles. Afin d'améliorer la précision des modèles, nous appliquons l'opération de pré-traitement du format « OneHotEncoding ». Notre ensemble de données comporte quatre (4) classes. : « Mask » ; « Mask_Chin »; « Mask_Mouth_Chin » et « Mask_Nose_Mouth ». L'opération « OneHotEncoding » est une technique utilisée pour représenter chaque classe par un vecteur binaire distinct.

Nous importons « train_test_split » de « sklearn.model_selection » pour diviser notre jeu de données. Ainsi, nous passons 0,1 en tant que paramètre dans le « train_test_split » qui divisera les données de sorte que 10 % des données seront dans la partie validation et les 90 % restantes seront dans la partie d'entraînement.

Nous fixons les valeurs des paramètres tels que le nombre de lots de traitement, la taille des images traitées et le nombre d'itérations puis, nous définissons le taux d'apprentissage, qui est un hyperparamètre qui contrôle la mesure dans laquelle le modèle répond à l'erreur d'estimation chaque fois que les poids du modèle neuronal de notre réseau sont mis à jour.

Une fois que nous avons décrit nos hyperparamètres, nous implémentons notre propre architecture, ce qui nous a permis de faire des modifications et des expérimentations.

Un réseau neuronal convolutif se compose de plusieurs couches de convolution et de mise en commun. Notre processus convolutif est constitué de trois (3) blocs de convolutions constitués chacun d'une couche de convolution et ont successivement la profondeur 3 des cellules 16, 32, 64. La taille cible de nos images est 100x100x3.

Nous utilisons une couche de sortie pour créer notre sortie finale. La fonction Dense () est utilisée ainsi avec le neurone 4, car nous avons que quatre (4) classes.

Nous avons utilisé la fonction d'activation Relu (Rectification Linéaire) au niveau des couches cachées pour apprendre à classifier nos visages aux 4 différentes classes et, pour la dernière couche entièrement connectée nous optons pour la fonction d'activation « Softmax » pour obtenir une probabilité pour nos 4 différentes classes.

Nous compilons notre modèle avec l'optimisateur « Adam » et la fonction de perte « categorical_crossentropy ». Il s'agit d'une fonction de perte d'entropie croisée pour les ensembles de données catégorielles.

Puis nous utilisons les métriques fondamentales telles que le score F1, le rappel, la précision, la matrice de confusion et la perte d'entropie croisée pour afficher les performances de notre modèle. Nous affichons les graphes de précision et de perte. Nous reviendrons plus en détail sur ces métriques dans le chapitre suivant dans la section dédiée aux résultats et discussions.

4.1. Première expérimentation

Pour notre première expérimentation, nous avons ajouté une couche de Batch normalisation dans chaque bloc de convolution et avons ajouté le paramètre « Dropout » qui a une valeur de 0.5 entre la deuxième couche de réseau de neurones artificiels et la dernière couche. La taille de lots est égale à 15 et le taux d'apprentissage est de 0.001. Notre modèle est entraîné sur 20 itérations.

4.2. Deuxième expérimentation

Notre deuxième expérience présente quelques variations par rapport à la première. Nous avons utilisé la même taille de lot et nous l'avons entraîné pour le même nombre d'itérations « epochs ». Nous avons juste une couche de Batch normalisation dans le premier bloc de convolution avec le paramètre « Dropout » de même valeur 0.50.

4.3. Troisième expérimentation

Notre modèle pour cette expérimentation a été entraîné sur LeNet-5 sans aucune couche de normalisation et de paramètres « Dropout ». Le taux d'apprentissage a été réduit à 0.0001 nous avons divisé notre jeu de données à une proportion 25 % et 75 %; donc 25 % des données seront dans la partie test et les 75 % restantes seront dans la partie d'entraînement.

4.4. Bases de données

Nous stockons nos données de reconnaissance faciale en utilisant l'architecture Big Data. Notre application utilise Python pour détecter un grand nombre d'images d'entrée à partir du jeu de données. Lors de la mise en œuvre d'un système de reconnaissance faciale, l'objectif est d'identifier les individus qui constituent un danger, cependant une de ses problématiques est la détection de ces individus en temps réel. Parmi les trois (3) architectures Big data, notre choix s'est porté sur l'architecture Lambda permettant la fusion de traitement par bloc de données (batch) et des nouvelles données entrées (temps réels).

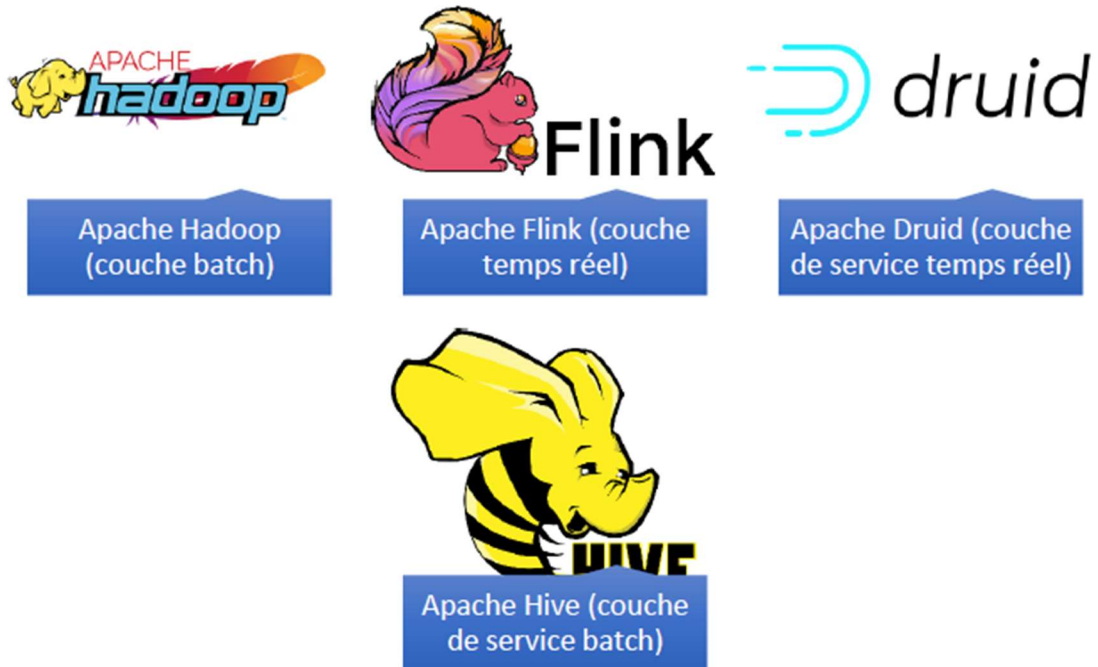


FIGURE 17 : ARCHITECTURE DU BIG DATA PROPOSÉ

4.5. Les métriques de performances

Pour l'évaluation de nos différents modèles, nous avons utilisé les mesures de performance suivantes :

4.5.1. Précision, rappel & F1

La précision [92] représente le nombre de prédictions correctement classées sur le nombre total de prédictions :

$$\text{précision} = \frac{\text{vraies positives}}{(\text{vraies positives} + \text{faux positives})}$$

Le rappel [92] détermine l'exactitude de la classe positive d'une prédiction en échange de la prise en compte des faux négatifs. Il est déterminé par la formule ci-dessous :

$$\text{rappel} = \frac{\text{vraies positives}}{(\text{vraies positives} + \text{faux négatives})}$$

La mesure F1 [92] est l'équilibre de la précision et du rappel. La mesure F1 est petite si l'une des métriques est petite, F1 est donnée par la formule :

$$F1 = 2 \frac{\textit{précision} * \textit{rappel}}{(\textit{précision} + \textit{rappel})}$$

4.5.2. Matrice de confusion

Les matrices de confusion [92] sont utilisées pour montrer les performances sur les problèmes de classification. Elles comparent votre prédiction aux résultats réels pour déterminer l'exactitude.

La matrice de confusion a 4 valeurs : Vrai Positif, Vrai Négatif, Faux Positif et Faux Négatif.

VP (Vrai Positif): représente le nombre de Vrais Positifs

VN (Vrais négatifs) : représente le nombre de Vrais Négatifs

FP (Faux Positif) : indique le nombre de Faux Positifs

FN (Faux Négatif) : définit le nombre de Faux Négatifs

CHAPITRE 7 : RÉSULTATS ET DISCUSSIONS

1. Présentations des résultats et discussions.

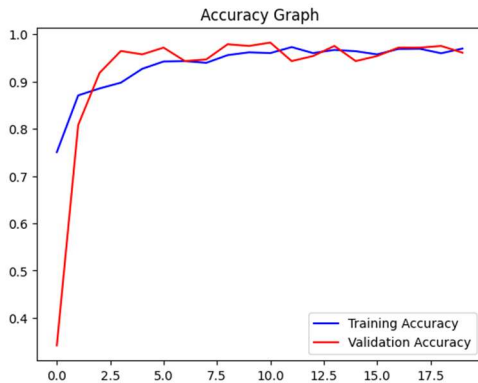
Dans ce chapitre, il sera question de présenter les trois (3) expérimentations menées au cours de notre travail. Les figures 18, 19 et 20 illustrent le graphique de précision d'entraînement et de validation ainsi que le graphique de perte d'entraînement et de validation de nos modèles.

Nos expérimentations utilisent une taille de lot constante de 15 pour nos échantillons d'apprentissage et de validation, et elles sont entraînées sur 20 itérations. Après le processus de formation, nous avons généré deux graphiques : (1) le graphique de précision d'entraînement et de validation, et (2) le graphique de perte d'entraînement et de validation.

Nous observons que l'ajout de la couche de normalisation à chaque bloc de convolution lors de la première expérimentation, ainsi que l'utilisation du paramètre « Dropout » et la réduction du taux d'apprentissage dans la troisième expérience, contribuent à renforcer la stabilité du réseau neuronal. Ces modifications améliorent les performances du réseau à la fois sur les données d'apprentissage et sur les données de validation.

Dans nos graphiques, les lignes bleues représentent la précision et la perte d'entraînement, tandis que les lignes rouges représentent la précision et la perte de validation. L'axe des abscisses correspond au nombre d'époques d'entraînement, tandis que l'axe des ordonnées représente les valeurs de précision et de perte.

(1)



(2)

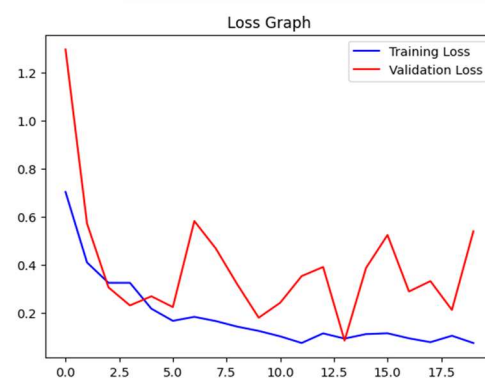
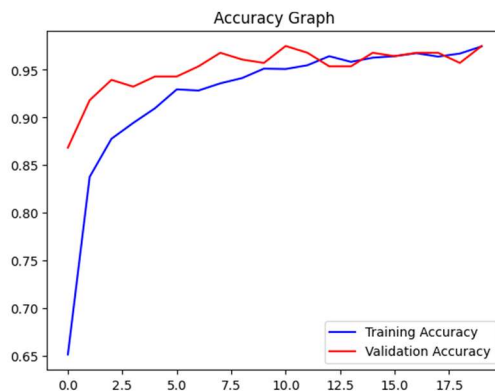


FIGURE 18 : (1) GRAPHIQUE DE PRÉCISION D'ENTRAÎNEMENT ET DE VALIDATION ET (2) GRAPHIQUE DE PERTE D'ENTRAÎNEMENT ET DE VALIDATION POUR LA PREMIÈRE EXPÉRImentation.

D'après la Figure 18, nous observons que la perte d'entraînement diminue à mesure que le nombre d'époques augmente, tandis que la perte de validation diminue initialement, puis oscille avec l'augmentation des époques. Parallèlement, la précision de l'entraînement et la précision de la validation augmentent au fur et à mesure des époques, atteignant une performance de 97,15 % après 20 époques. À partir de ces graphiques, nous en concluons que les meilleures performances sont obtenues lorsque le nombre d'époques est fixé à 20 et que le taux d'apprentissage est de 0,001.

(1)



(2)

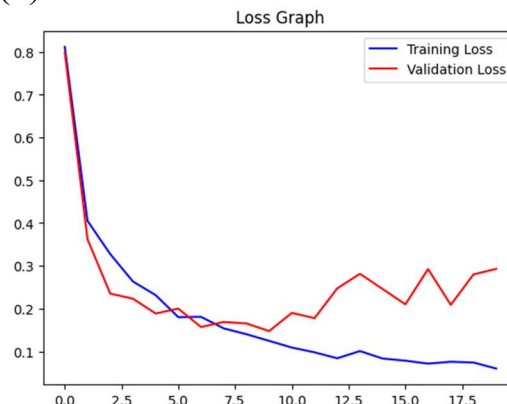


FIGURE 19 : (1) GRAPHIQUE DE PRÉCISION D'ENTRAÎNEMENT ET DE VALIDATION ET (2) GRAPHIQUE DE PERTE D'ENTRAÎNEMENT ET DE VALIDATION POUR LA DEUXIÈME EXPÉRImentation.

La figure 19 présente les graphiques de précision et de perte de la deuxième expérimentation. Initialement, la précision augmente rapidement au cours des deux premières époques, puis se stabilise près de 1. Cela indique que notre modèle ne nécessite pas un grand nombre d'époques pour être entraîné efficacement. La précision atteint finalement 98,80 % sur l'ensemble d'entraînement. La perte d'entraînement diminue progressivement et se stabilise près de zéro, tandis que la perte de validation augmente lentement.

Au début de l'entraînement, les pertes sont élevées, mais elles diminuent légèrement au fil du temps et se stabilisent. Après l'entraînement, un test est effectué, et le graphique de précision se stabilise rapidement près de 1, ce qui indique une performance élevée du modèle.

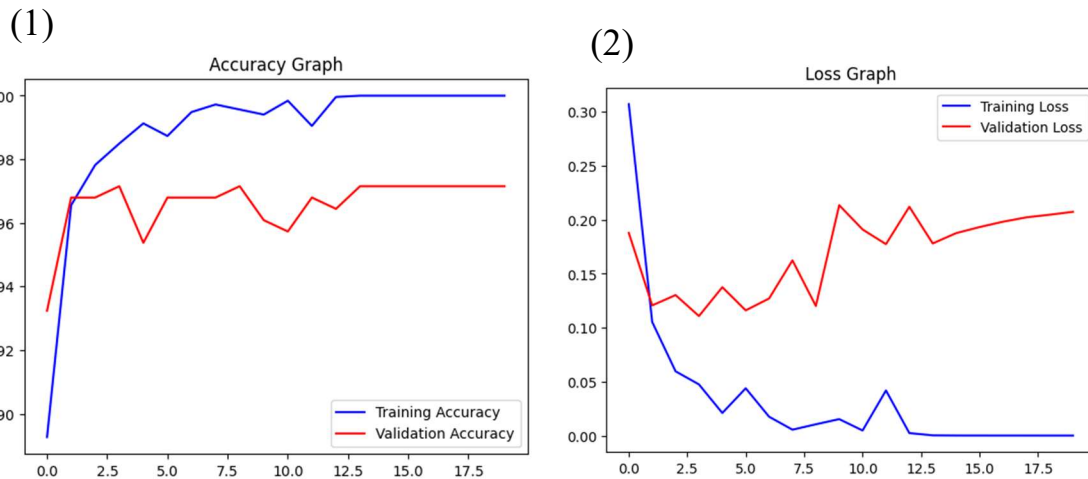


FIGURE 20 : (1) GRAPHIQUE DE PRÉCISION D'ENTRAÎNEMENT ET DE VALIDATION ET (2) GRAPHIQUE DE PERTE D'ENTRAÎNEMENT ET DE VALIDATION POUR LA TROISIÈME EXPÉRIMENTATION.

En analysant le graphique de la figure 20, nous constatons que le modèle présente un sur-ajustement. Cela se manifeste par une augmentation de la précision d'entraînement et une diminution de la perte d'entraînement. Au cours du processus d'entraînement, la précision d'entraînement est passée de 65 % à 98 %. Cependant, notre modèle montre également une amélioration de la perte de validation, ce qui indique qu'il ne généralise pas suffisamment bien sur notre ensemble de validation.

Une matrice de confusion est utilisée pour évaluer la performance de nos modèles de classification avec notre ensemble de données de test. La matrice de confusion est un tableau qui compare les valeurs réelles aux valeurs prédites. Les figures 21 à 23 montrent que notre modèle est capable de classer correctement les

classes. La matrice montre des valeurs élevées de TP et de TN par rapport aux faibles valeurs de FP et de FN, nous pouvons donc dire que notre modèle est capable de prédire correctement les échantillons corrects avec une bonne précision.

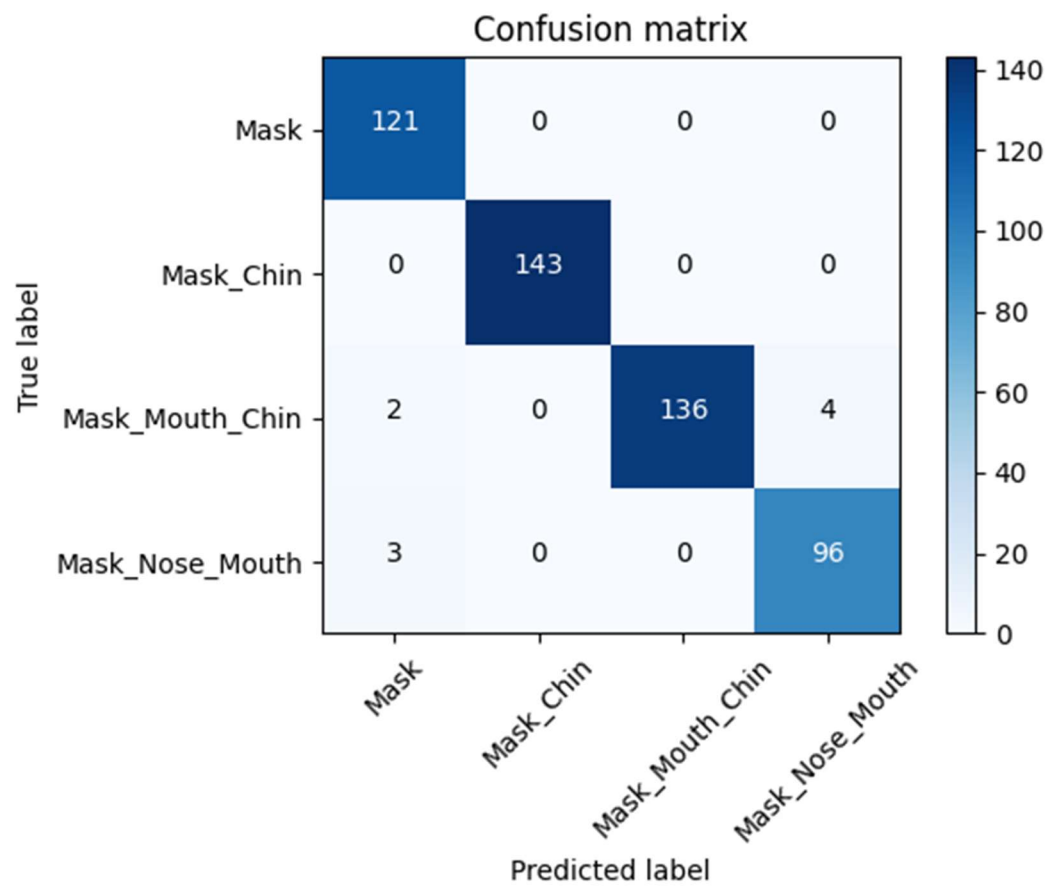


FIGURE 21 : MATRICE DE CONFUSION DE LA PREMIÈRE EXPÉRIMENTATION.

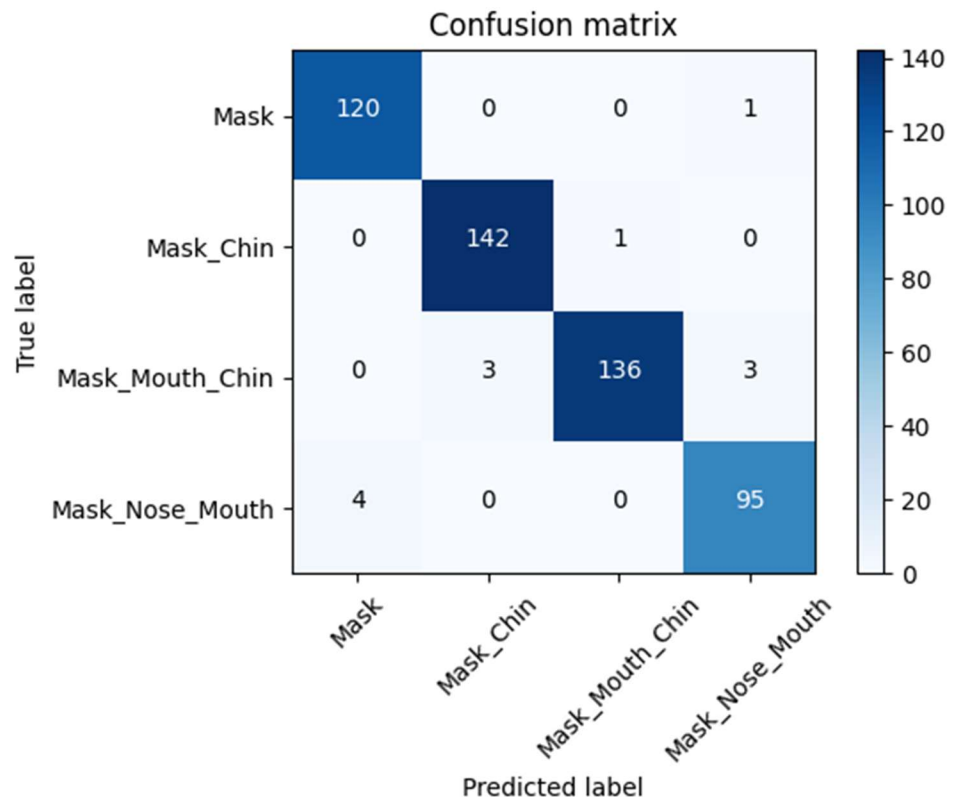


FIGURE 22 ; MATRICE DE CONFUSION DE LA DEUXIÈME EXPÉRIMENTATION.

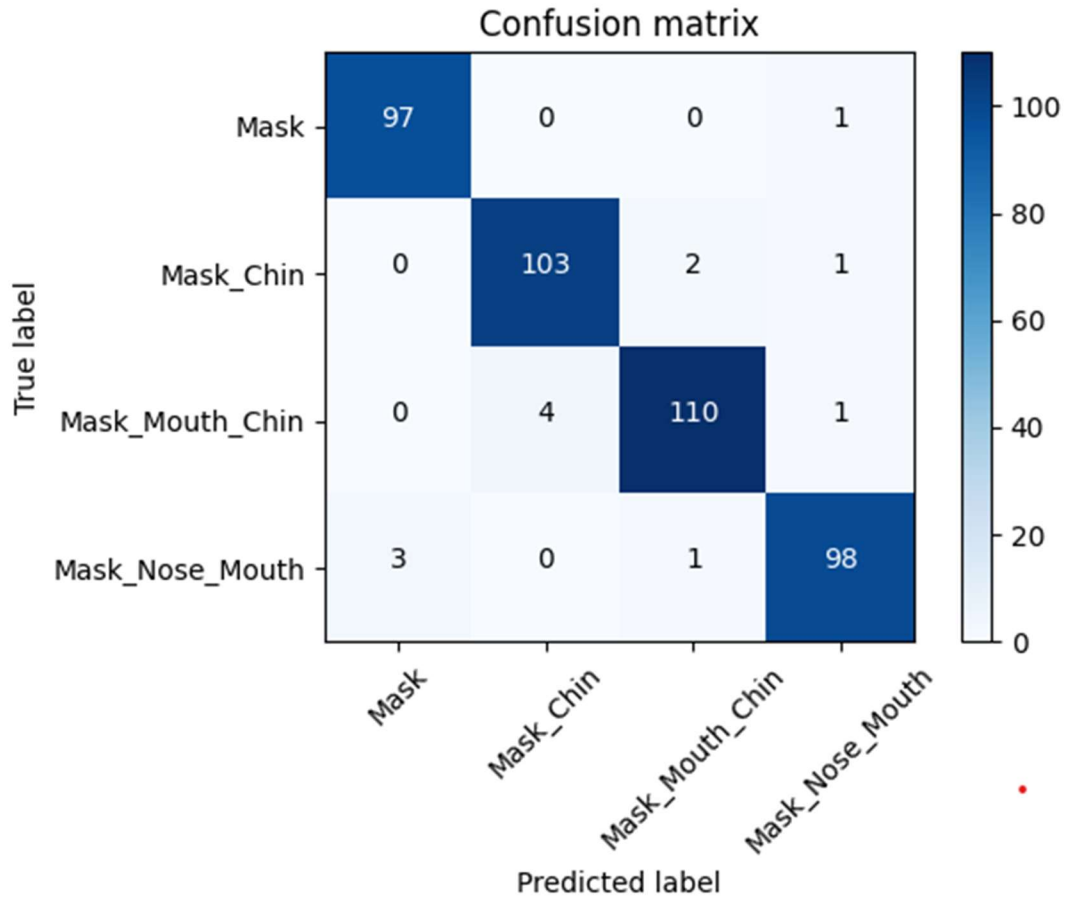


FIGURE 23 : MATRICE DE CONFUSION DE LA TROISIÈME EXPÉRIMENTATION.

Méthode de validation	Classification CNN			
	Class	Précision	Rappel	Score-F1
1ere experimentation	0	0.87	0.99	0.93
	1	0.99	0.94	0.97
	2	0.95	0.97	0.96
	3	0.99	0.89	0.94
2eme experimentation	0	0.97	0.99	0.98
	1	0.99	0.98	0.99
	2	0.95	0.99	0.97
	3	0.98	0.93	0.95
3eme experimentation	0	1	0.99	1
	1	0.99	0.99	0.99
	2	0.99	0.99	0.99
	3	0.98	0.99	0.99

TABEAU 2 : TABLEAU DE COMPARAISON DES RÉSULTATS DES TROIS EXPÉRIMENTATIONS.

Notre modèle LeNet-5 a obtenu le rappel le plus élevé pour la classification multi-classe, avec une valeur de 0,99 pour la classe 0 dans nos trois expérimentations, ainsi que pour les classes 1, 2 et 3 dans notre troisième expérimentation. En ce qui concerne la précision de la classification multi-classe, notre modèle a également atteint la valeur la plus élevée de 0,99 pour la classe 1 dans la première et la deuxième expérimentation, ainsi que la valeur de 1 pour la classe 0 dans la troisième expérimentation.

En ce qui concerne la méthode de partitionnement des données utilisée dans nos expérimentations, le score F1 le plus élevé obtenu est de 0,97 pour les deux premières expérimentations, pour la classe 0.

2. Résultats obtenus

L'étude de notre modèle nous a permis d'apporter des améliorations significatives, nous conduisant finalement à un résultat satisfaisant compte tenu des contraintes auxquelles nous étions confrontés.

Les contraintes dans notre recherche et comme toujours dans les problèmes de classification sont relatives aux données traitées. La préparation et le prétraitement des données sont cruciaux pour la formation aux données. Aussi du fait que d'autres images étaient contenues dans certaines classes, accentuent aussi cette problématique de classification.

Nous avons créé un système capable de reconnaître les images de visage masquées et non masquées.

Ci-dessous quelques exemples de reconnaissance de visage en utilisant les trois modèles :



FIGURE 24 : RÉSULTATS DE LA RECONNAISSANCE FACIALE DE LA 1ERE EXPÉRIMENTATION.

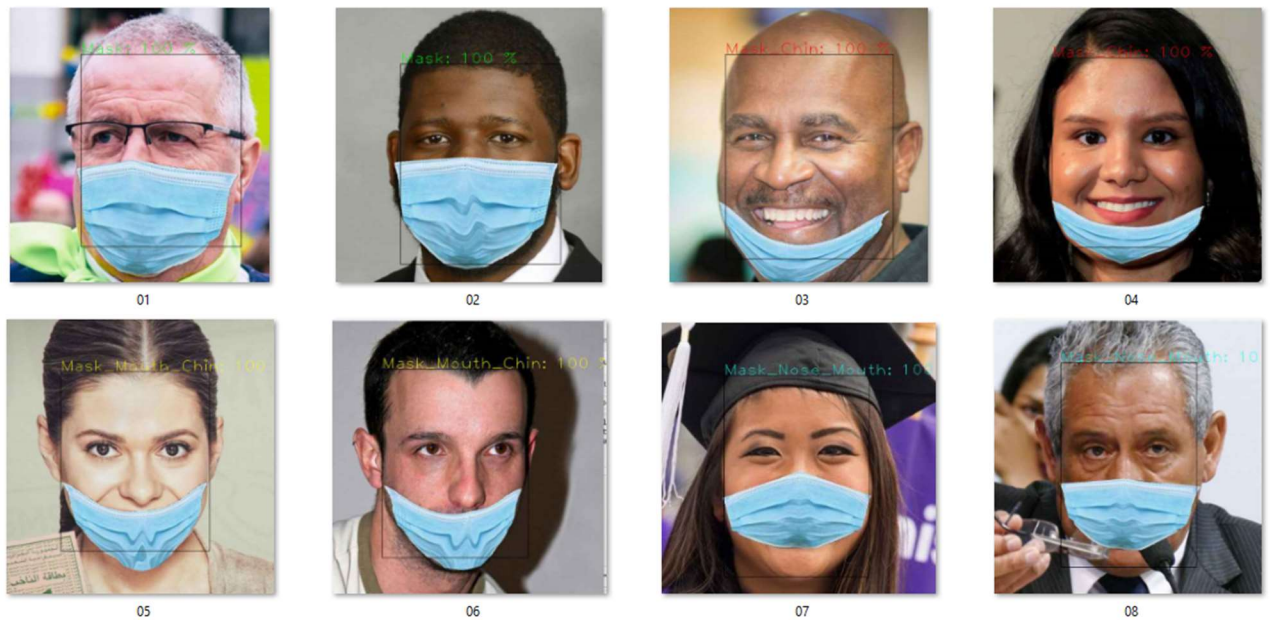


FIGURE 25 : RÉSULTATS DE LA RECONNAISSANCE FACIALE DE LA 2EME EXPÉRIMENTATION.



FIGURE 26 : RÉSULTATS DE LA RECONNAISSANCE FACIALE DE LA 3EME EXPÉRIMENTATION.

CONCLUSION GÉNÉRALE

Le présent rapport nous synthétise le travail effectué dans le cadre du projet de fin d'études à l'Université du Québec à Trois-Rivières. Ce projet s'est concentré sur l'étude de la reconnaissance faciale et la détection des masques faciaux en utilisant des techniques d'apprentissage en profondeur, en particulier les réseaux de neurones convolutifs.

Nous avons d'abord entrepris une étude approfondie des concepts du Big Data, de la vision par ordinateur et du Cloud Computing. Ensuite, nous nous sommes familiarisés avec le contexte général de notre approche pour la reconnaissance faciale en utilisant les réseaux de neurones convolutifs.

Le but de ce rapport est d'identifier les visages sans masques et de sensibiliser les individus à l'importance de bien porter des masques. L'analyse principale se concentre sur la classification des visages humains masqués et l'influence de la proportion du visage humain sur la précision.

Nous avons proposé quatre classes, chacune représentant une catégorie de classification. Pour automatiser cette classification des visages, nous avons réalisé trois expérimentations en utilisant le modèle LeNet-5 [4]. Dans ces expérimentations, nous avons appliqué des techniques telles que l'augmentation de données, l'ajout de couches de normalisation par lots (Batch Normalization) et de « Dropout », et nous avons testé différents taux d'apprentissage.

Les performances des modèles ont été évaluées sur l'ensemble de données Flickr-Face HQ (FFHQ) [5], qui comprend 3000 images. Nos résultats de test ont démontré un taux de précision élevé dans l'identification des personnes dont le visage est masqué, celles dont seulement le menton est couvert, et celles dont le nez n'est pas couvert. Notre modèle a atteint une précision d'entraînement de 98,80 % grâce à l'utilisation de la normalisation par lots et du "Dropout".

Notre projet de reconnaissance faciale masquée présente un potentiel d'extension considérable dans le futur. Cependant, bien que notre modèle ait obtenu de bons résultats de validation, il est important de reconnaître que ces résultats peuvent ne pas être généralisables à tous les types d'ensembles de données d'entraînement. Par exemple les masques n'ont pas tous la même couleur bleue.

Dans le cadre de nos travaux futurs, nous envisageons d'utiliser le même jeu de données d'entraînement pour tester divers autres modèles, tels que MobileNet [93]. Cette approche nous permettra d'explorer différentes options et de sélectionner un modèle plus stable et mieux adapté à la reconnaissance faciale masquée.

En outre, Nous pouvons ajouter plus de valeurs qui ne se limitent pas seulement à la reconnaissance faciale, mais incluent également la reconnaissance des genres. Bien sûr, cela nécessitera des changements de nos jeux de données, où nous devons inclure des annotations pertinentes pour ces nouvelles tâches.

Par ailleurs, nous nous aventurons à explorer l'utilisation d'autres types de réseaux neuronaux tels que les réseaux antagonistes génératifs [10] pour la réalisation de la reconnaissance faciale masquée. Cette approche nous permettra de repousser les limites et d'expérimenter de nouvelles méthodes pour améliorer les performances de notre système.

RÉFÉRENCES

[1] P.S. Prasad, R. Pathak, V.K. Gunjan et H.V.R. Rao, «Deep Learning Based Representation for Face Recognition, » Springer: Berlin, Germany, 2019; pp. 419–424.

[2] V. C.-C. Cheng, S.-C. Wong, V. W.-M. Chuang et al., «The role of community-wide wearing of face mask for control of coronavirus disease 2019 (covid-19) epidemic due to sars-cov-2, » Journal of Infection, vol. 81, no. 1, pp. 107–114, 2020. [En ligne]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7177146/pdf/main.pdf>.

[3] « Convolutional neural network, » Visité en Septembre 2020, [En ligne]. Available: https://en.wikipedia.org/wiki/Convolutional_neural_network.

[4] « LeNet, » Visité en Septembre 2020, [En ligne]. Available : <https://en.wikipedia.org/wiki/LeNet#:~:text=In%20general%2C%20LeNet%20refers%20to,in%20large%2Dscale%20image%20processing>.

[5] « Flickr-Face HQ Dataset, » visité en Octobre 2020, [En ligne]. Available: <https://github.com/NVlabs/ffhq-dataset>.

[6] Y. Taigman, M. Yang, M. Ranzato et L. Wolf, « Deepface: Closing the gap to human-level performance in face verification, » In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 1701–1708. [En ligne]. Available: https://openaccess.thecvf.com/content_cvpr_2014/papers/Taigman_DeepFace_Closing_the_2014_CVPR_paper.pdf.

[7] G. Hu, Y. Yang, D. Yi, J. Kittler, W. Christmas, S.Z. Li et T. Hospedales, « When face recognition meets with deep learning: An evaluation of convolutional neural networks for face recognition, » In Proceedings of the IEEE International Conference on Computer Vision Workshops, Santiago, Chile, 11–12 December 2015; pp. 142–150. [En ligne]. Available : https://www.cvfoundation.org/openaccess/content_iccv_2015_workshops/w11/papers/Hu_When_Face_Recognition_ICCV_2015_paper.pdf.

[8] W. Hariri, « Efficient masked face recognition method during the covid-19 pandemic, » 2021. [En ligne]. Available: <https://arxiv.org/abs/2105.03026>.

[9] H. Deng, Z. Feng, Z. Qian, X. Lv, H. Li et G. Li « MFCosface: A masked-face recognition algorithm based on large margin cosine loss, » Appl. Sci. 2021, 11, 7310. [En ligne]. Available: <https://www.mdpi.com/2076-3417/11/16/7310>.

- [10] N. U. Din, K. Javed, S. Bae et J. Yi, « A novel GAN-based network for unmasking of masked face, » IEEE Access 2020, 8, 44276-44287. [En ligne]. Available: <https://ieeexplore.ieee.org/ielx7/6287639/8948470/09019697.pdf>.
- [11] G. Deore, R. Bodhula, V. Udpikar et V. More, « Study of masked face detection approach in video analytics, » 2016 Conference on Advances in Signal Processing (CASP), Pune, 2016, pp. 196-200, doi: 10.1109/CASP.2016.7746164.
- [12] M. S. Ejaz et M. R. Islam, « Masked Face Recognition Using Convolutional Neural Network, » 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh, 2019, pp. 1-6, doi:10.1109/STI47673.2019.9068044.
- [13] A. Anwar et A. Raychowdhury, « Masked face recognition for secure authentication, » arXiv preprint arXiv:2008.11104 (2020).
- [14] Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong, H. Wu, P. Yi, K. Jiang, N. Wang, Y. Pei et al., « Masked face recognition dataset and application, » arXiv preprint arXiv:2003.09093, 2020.
- [15] P. Nagrath, R. Jain, A. Madan, R. Arora, P. Kataria et J. Hemanth, « SSDMNv2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2, » Sustainable cities and society, March 2021. [En ligne]. Available: <https://pdf.sciencedirectassets.com/>.
- [16] « Face masked dataset, » Visité en Septembre 2020, [En ligne]. Available: <https://www.kaggle.com/datasets/manishkc06/face-mask-dataset>.
- [17] S. V. Militant et N. V. Dionisio, « Real-Time Facemask Recognition with Alarm System using Deep Learning, » IEEE Control and System Graduate Research Colloquium (ICSGRC), 2020. [En ligne]. Available: <https://ieeexplore.ieee.org/ielx7/9223048/9232452/09232610.pdf>.
- [18] S. Ge, J. Li, Q. Ye et Z. Luo, « Detecting Masked Faces in the Wild with LLE-CNNs, » IEEE Conference on Computer Vision and Pattern Recognition, 2017. [En Ligne]. Available: https://openaccess.thecvf.com/content_cvpr_2017/papers/Ge_Detecting_Masked_Faces_CVPR_2017_paper.pdf.
- [19] M. Loey, G. Mangogaran, T. M.H.N et K. N.E.M., « A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic, » National Library of Medicine, 1 January 2021. [En ligne]. Available: <https://pubmed.ncbi.nlm.nih.gov/32834324/>.
- [20] « Real-World Masked Face Dataset » visité en Décembre 2020, [En ligne]. Available: <https://github.com/X-zhangyang/Real-World-Masked-Face-Dataset>.
- [21] « Simulated Masked Face Dataset » visité en Décembre 2020, [En ligne]. Available:] <https://github.com/prajnasb/observations>.

- [22] G. B. Huang, M. Mattar, T. Berg, et E. Learned-Miller, «Labeled faces in the wild: A database for studying face recognition in unconstrained environments, » Technical report, 2007.
- [23] B. Balusamy, R. N. Abirami, S. Kadry et A.H. Gandomi, « Big Data Concepts, Technology, and Architecture, » Published by JohnWiley & Sons, Inc., New Jersey, USA, 2021.
- [24] J-S. Bergé, S. Grumbach et V. Zeno-Zencovich, « The ‘Datasphere’, Data Flows beyond Control, and the Challenges for Law and Governance, » Forthcoming in the European Journal of Comparative Law and Governance n. 5 (2018) [En ligne]. Available: <https://deliverypdf.ssrn.com/>
- [25] P. Lemberger, M. Batty, MÉ Morel et J-L. Raffaëlli, « Big Data and machine learning: Manuel du data scientis, » 5 rue Laromiguière, 75005 Paris (2015).
- [26] V. Mayer-Schonberger et K. Cukier (2013), « Big data: A revolution that will transform how we live, work, and think, » UK: Hachette.
- [27] B. Braian de Souza et A. Cicero Aparecido, « L'importance des données pour Big compréhension des habitudes d'achat des consommateurs, l'augmentation de l'avantage concurrentiel dans le détail, » Magazine scientifique multidisciplinaire du Centre du savoir. Numéro 08. Année 02, vol 03. p 18-41, Novembre 2017. ISSN: 2448-0959.
- [28] D-S. Singh, et G. Singh, «Big data – A Review, » International Research Journal of Engineering and Technology (IRJET), Apr -2017, Volume: 04 Issue: 04, 822-824. [En Ligne] Available: <https://d1wqtxts1xzle7.cloudfront.net/53502016/IRJET-V4I4167-libre.pdf>.
- [29] J. Hurwitz, A. Nugent, F. Halper et M. Kaufman, « Big data for dummies, » Published by JohnWiley & Sons, Inc., New Jersey, USA, 2013. <https://jan.newmarch.name/IoT/BigData/Big%20Data%20For%20Dummies.pdf>
- [30] J. Singh et V. Singla, « Big Data: Tools and Technologies in Big Data, » International Journal of Computer Applications (0975 – 8887), f Computer Applications (0975 – 8887) Volume 112 – No 15, February 2015, 6-10.
- [31] T. Bikku, NS. Rao et AR. Akepogu, « Hadoop based feature selection and decision making models on Big Data, » Indian Journal of Science and Technology. 2016 Mar; 9(10). DOI: 10.17485/ijst/2016/v9i10/88905.
- [32] L. Greeshma et G. Pradeepini, « Big Data Analytics with Apache Hadoop MapReduce FrameworkI, » Indian Journal of Science and Technology, Vol 9(26), DOI: 10.17485/ijst/2016/v9i26/93418, July 2016.
- [33] A. Elsayed, O. Ismail, et M. E. El-Sharkaw, « MapReduce: State-of-the-Art and Research Directions, » IJCEE 2014 Vol.6 (1): 34-39 ISSN: 1793-8163 DOI: 10.7763/IJCEE.2014.V6.789.

- [34] « Big Data analytics market, » visité en Janvier 2021, [En Ligne]. Available: https://www3.microstrategy.com/getmedia/cd052225-be60-49fd-ab1c-4984ebc3cde9/Dresner-Report-Big_Data_Analytic_Market_Study-WisdomofCrowdsSeries-2017.pdf
- [35] Advanced Analytics & AI Summit: BFSI 2019. Survey Report: Big Data analytics for financial services (2018).
<https://www.intelligentautomation.network/events/advancedanalyticsfinance/downloads/big-data-analytics-for-financial-services-survey-report-2018>.
- [36] D. Mavaluru, R. Shriram et V. Sugumaran, « Big Data Analytics in Information Retrieval: Promise and Potential », Proceedings of 0 8 th IRF International Conference, 05 th July – 2014.
- [37] « 30+ Youtube Statistics and facts [CurrentYear, » visité en Janvier 2021, [En Ligne]. Available: <https://www.comparitech.com/tv-streaming/youtube-statistics/>.
- [38] « Attentats du 11 Septembre, » visité en Janvier 2021, [En Ligne]. Available: https://fr.wikipedia.org/wiki/Attentats_du_11_septembre_2001.
- [39] « homme chat » visité en Février 2021, [En Ligne]. Available: <https://www.istockphoto.com/fr/photos/homme-chat>.
- [40] « Media post: Can car technology help prevent fatigued driving? » visité en Février 2021,[En Ligne]. Available: <https://bestsellingcarsblog.com/2021/09/media-post-can-car-technology-help-prevent-fatigued-driving/>
- [41] S. Ingle et M. Phute, « Tesla Autopilot : Semi-Autonomous Driving, an Uptick for Future Autonomy, » International Research Journal of Engineering and Technology (IRJET), Volume: 03 Issue: 09 | Sep-2016 , 369-372.
- [42] « Computer vision in retail: 7 use cases and one way to succeed, » visité en Mars 2021, [En Ligne]. Available: <https://mindtitan.com/resources/industry-use-cases/computer-vision-in-retail/>
- [43] P. Spagnolo et al, « Non-Invasive Soccer Goal Line Technology: A Real Case Study, »in National Research Council of Italy , IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2013, 1011-1018. DOI : 10.1109/CVPRW.2013.147.
- [44] SV. Doctorvaladan et al, « Accuracy of Blood Loss Measurement during Cesarean Delivery, » Published by AJP Reports, Avril 2017, 93-100. Doi: [10.1055/s-0037-1601382](https://doi.org/10.1055/s-0037-1601382).
- [45] « Qu'est ce que le Cloud Computing » visité en Aout 2023, [En ligne]. Available:] <https://www.salesforce.com/fr/learning-centre/tech/cloudcomputing/>.

- [46] NB. Ruparelia, « Cloud Computing, » Published The MIT Press,2016.
<https://mitpressbookstore.mit.edu/book/9780262546478>.
- [47] Vaquero, L, Rodero-Merino, L, Caceres, J et Lindner M (2009), « A break in the clouds: towards a cloud definition, » ACM SIGCOMM computer communications review. Volume 39, Issue 1, January 2009, 50-55. Doi: 10.1145/1496091.1496100.
- [48] I. Ashraf, « An Overview of Service Models of Cloud Computing, » International Journal of Multidisciplinary and Current Research. Vol.2 (July/Aug 2014), 779-783.
<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=1596e1fda5456adfbba56e530e4f69be08696c64>.
- [49] « Cloud Computing et ses différents modèles et topologies. » visité en Juin 2021, [En Ligne]. Available: <https://fr.linkedin.com/pulse/cloud-computing-et-ses-diff%C3%A9rents-mod%C3%A8les-topologies-saad-fauzi>.
- [50] A. Rashid et A, Chaturvedi, « Cloud Computing Characteristics and Services: A Brief Review, » International Journal of Computer Sciences and Engineering. Vol.-7, Issue-2, Feb 2019É, 421-426. [En Ligne]. Available: https://www.ijcseonline.org/pub_paper/70-IJCSE-05826.pdf.
- [51] A. Nayak et A. Yassir, « Cloud Computing as an Emerging Paradigm, » International Journal of Computer Science and Network Security, VOL.12 No.1, January 2012. 61-65.
- [52] M. Armbrust, A.D. Joseph, R.H. Katz et A. Patterson, « Above the Cloud: A Berkeley View of Cloud Computing, » Technical Report, EECS Department, University of California, Berkeley, 2009. [En Ligne]. Available: <https://www2.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf>.
- [53] S. P. Mirashe et N.V. Kalyankar, « Cloud Computing, » Journal of Compyting, Volume 2, Issue 3, March 2010, ISSN 2151-9617 Available: https://sites.google.com/site/Journal_of_Compyting/
- [54] « Face recognition systems, » visité en Juin 2021, [En ligne]. Available: https://fr.wikipedia.org/wiki/Syst%C3%A8me_de_reconnaissance_faciale
- [55] « Reconnaissance faciale automatisée dans les secteurs public et privé, » visité en Aout 2021, [En ligne]. Available: https://www.priv.gc.ca/media/1766/fr_201303_f.pdf
- [56] S. KO. BASEMA, « Identification des personnes par reconnaissance de visage pour la sécurité d'une institution bancaire, » Mémoire pour l'obtention du diplôme

de la licence en informatique de gestion, Institut supérieur pédagogique de Bukavu-RDC, 2010.

[57] W. Zhao, R. Chellappa, P. J. Phillips et A. Rosenfeld, « Face recognition: A literature survey, » ACM computing surveys (CSUR), vol. 35(4), pp. 399-458, 2003.

https://inc.ucsd.edu/mplab/users/marni/Igert/Zhao_2003.pdf

[58] K. Bouchra, « Mise au point d'une application de reconnaissance faciale, » Mémoire pour l'obtention du diplôme de master en informatique, Université Abou Bakr Belkaid-Tlemcen, novembre 2013.

[59] S. G. Bhele et V. H. Mankar, « A Review Paper on Face Recognition Techniques, » International Journal of Advanced Research in Computer Engineering & Technology. Volume 1, Issue 8, October 2012, 339-346.

[60] J.-T. Chien et C.-P. Liao, « Maximum Confidence Hidden Markov Modeling for Face Recognition, » IEEE transactions on pattern analysis and machine intelligence, vol. 30(4), pp. 606-616, 2008.

[61] F. Rosenblatt, «The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, » Psychological Review, 1958.

[62] P. Paquet, « L'utilisation des réseaux de neurones artificiels en finance ».

[63] « Les réseaux de neurones convolutifs, » visité en Aout 2021, [En ligne]. Available : <https://histrory.hashnode.dev/les-reseaux-de-neurones-convolutifs>

[64] P. Hensman et D. Masko, «The Impact of Imbalanced Training Data for Convolutional Neural Networks, » Degree project, in computer science, first level Stockholm, Sweden 2015.

[65] « Qu'est-ce que le Deep Learning et comment ça marche, » visité en Septembre 2021, [En ligne]. Available : <https://www.saagie.com/fr/blog/l-histoire-du-deep-learning/>

[66] Y. LeCun, K. Kavukcuoglu et C. Farabet, « Convolutional networks and applications in vision, » In: Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on (2010), pp. 253–256.

[67] O. M. Parkhi, A. Vedaldi, et A. Zisserman, « Deep Face Recognition, » Proceedings of the British Machine Vision, vol. 1(3), p. 6, 2015.

[68] « A comprehensive Guide to Convolutional Neural Network, » Visité en Septembre 2021, [En ligne]. Available: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>.

[69] A. M'BEGNAN NAGNAN, « Développement d'outils web de détection d'annotations manuscrites dans les imprimés anciens, » Mémoire pour l'obtention

du diplôme de maîtrise en mathématiques et informatique appliquées, Université du Québec à Trois-Rivières, avril 2021.

[70] K. O'Shea and R. Nash, «An Introduction to Convolutional Neural Networks, »arXiv: 1511.08458v2[CS.NET] 2 décembre 2015.

[71] « Max Pooling Layer, » Visité en Mai 2023, [En ligne]. Available: <https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/>

[72] « Flattenng Opération, » Visité en Mai 2023, [En ligne]. Available: <https://www.simplilearn.com/tutorials/deep-learning-tutorial/convolutional-neural-network>

[73] V. Nair et G.E. Hinton, « Rectified linear units improve restricted boltzmann machines[C], » International Conference on International Conference on Machine Learning. Omnipress, 2010:807-814.

[74] A. Krizhevsky, I. Sutskever, et G. E. Hinton, « Imagenet classification with deep convolutional neural networks, » 2021. [En ligne]. Available: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

[75] « The Architecture of Lenet-5, » Visité en Janvier 2022, [En ligne]. Available : <https://www.analyticsvidhya.com/blog/2021/03/the-architecture-of-lenet-5/>

[76] D.C. Cireşan, U. Meier, J. Masci, L.M. Gambardella et J. Schmidhuber, «Flexible, High-performance convolutional neural networks for Image Classification,» Arxiv preprint arXiv:1102.0183, 2011.

[77] B. Prenkaj, P. Velardi, et al, « A Survey of machine learning approaches for student dropout prediction in online courses, » ACM Computing Surveys Volume 53, Issue 3.

[78] S. Santurkar, D. Tsipras et al « How does batch normalization help optimization? » 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada. [En ligne]. Available : <https://proceedings.neurips.cc/paper/2018/file/905056c1ac1dad141560467e0a99e1cf-Paper.pdf>.

[79] I. Goodfellow, Y. Bengio et A. Courville, « Deep Learning, » MIT Press, Cambridge, MA, 2016.

[80] K. P. Murphy, « Machine Learning: A Probabilistic Perspective, » Cambridge: MIT Press. p. 247. ISBN 978-0-262-01802-9., 2012.

- [81] « Python, » Visité en Janvier 2022, [En ligne]. Available: <https://www.python.org>.
- [82] S. Gholizadeh., « Top Popular Python Libraries in Research, » Journal of Robotics and Automation Research, Rondebosch, Cape Town, 7701, South Africa, May 2022, 3(2), 142-145.
- [83] « OpenCV, » Visité en Février 2022, [En ligne]. Available: <https://en.wikipedia.org/wiki/OpenCV>
- [84] « Pandas software, » [En ligne]. Available: [https://en.wikipedia.org/wiki/Pandas_\(software\)](https://en.wikipedia.org/wiki/Pandas_(software))
- [85] « SciPy, » Visité en Février 2022, [En ligne]. Available: <https://en.wikipedia.org/wiki/SciPy>
- [86] « Scikit-learn, » Visité en Février 2022, [En ligne]. Available: <https://en.wikipedia.org/wiki/Scikit-learn>
- [87] « Tensorflow, » Visité en Février 2022, [En ligne]. Available: <https://en.wikipedia.org/wiki/TensorFlow>
- [88] « Matplotlib, » Visité en Février 2022, [En ligne]. Available: <https://matplotlib.org>.
- [89] « Jupyter Notebook, » Visité en Février 2022, [En ligne]. Available: <https://jupyter.org>.
- [90] « Coffe, » Visité en Février 2022, [En ligne]. Available: <https://caffe.berkeleyvision.org/>
- [91] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li et L. Fei-Fei, « ImageNet: A Large-Scale Hierarchical Image Database, » IEEE Computer Vision and Pattern Recognition (CVPR), 2009.
- [92] H. M et S. M.N, « A Review on Evaluation Metrics for Data Classification Evaluations, » International Journal of Data Mining & Knowledge Management Process, vol. 5, no. 2, pp. 01–11, Mar. 2015. [Online]. Available: <http://www.airconline.com/ijdkp/V5N2/5215ijdkp01.pdf>
- [93] S. -H. Tsang, « Review: MobileNetV1 — Depthwise Separable Convolution (Light Weight Model), » 2018, Octobre 14. [En ligne]. Available: https://miro.medium.com/max/1400/1*Voah8cvrs7gnTDf6acRvDw.png