



REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



Université Kasdi Merbah Ouargla

Faculté des nouvelles technologies de  
L'information et de la Communication

Département d'électronique et de la télécommunication

## MASTER

**Domain:** Electronique

**Option:** Electronique des Systèmes embarqués

**Proposé par:** BEN KRIMA ABD ERRAHIM et CHERFAOUI ABDENOUR et

IBRAHIM SOHIL SADY SAFI

- THÈME -

# RECONNAISSANCE DES ÉMOTIONS À L'AIDE DE L'APPRENTISSAGE PROFOND

**Date d'évaluation: 10/06/2023**

**Devant le Jury:**

DR. Chlaoua Rachid	MCA	Présidente	UKM Ouargla
DR. BENSID KHALED	MCB	Superviseur	UKM Ouargla
DR. KORICHI Maarouf	MCB	Examineur	UKM Ouargla



# *Dedicace*

A ma famille, soutien et encouragement.

A mes amis, qui ont été une source constante d'inspiration et de motivation.

Et à mes professeurs, qui m'ont mis au défi de grandir et d'apprendre.

Cette thèse vous est dédiée.

Et sincères remerciements au Dr. Khaled BN-Sid, qui nous a soutenus en permanence

# Summaries

Cette thèse vous est dédiée.....	3
<b>Introduction générale</b> .....	1
Chapitre I.....	3
<b>I.1. Introduction :</b> .....	4
<b>I.2. Biométrie :</b> .....	5
I.2.1. Définition:.....	5
I.2.2. Modalités biométriques :.....	<b>Error! Bookmark not defined.</b>
I.2.3. Modalités physiques :.....	<b>Error! Bookmark not defined.</b>
<b>I.2.3. Modalités Comportementales :</b> .....	10
<b>I.3. Émotion :</b> .....	12
I.3.1. Définition :.....	13
I.3.2. Le modèle d'émotion de base :.....	13
I.3.3. Le Modèle Dimensionnel :.....	13
I.3.4. Reconnaissance des émotions :.....	14
I.3.5. Caractéristique de l'émotion :.....	14
<b>I.4. Conclusion</b> .....	15
Chapitre II.....	16
<b>II.1. Introduction</b> .....	17
<b>II.2. Apprentissage automatique (le machine learning) :</b> .....	17
II.2.1. Identification :.....	18
II.2.2. Types d'apprentissage automatique :.....	18
<b>II.2.2.1. Apprentissage automatique supervisé :</b> .....	19
<b>II.2.2.2. Apprentissage automatique non supervisé :</b> .....	20
<b>II.2.2.3. Apprentissage automatique semi-supervisé :</b> .....	20
<b>II.3. Algorithmes de classification d'apprentissage automatique :</b> .....	21
II.3.1. K-Plus Proches Voisins (K-NN) :.....	21
II.3.2. Machine à Vecteur Support (SVM) :.....	22
II.3.3. Régression Logistique :.....	23
II.3.4. Arbre décisionnel :.....	24
<b>II.4 Apprentissage profond :</b> .....	25

II.1 Identification :	25
II.2 Types de méthode d'apprentissage en profondeur :	25
II.3. Comparaison du deep learning avec le machine learning :	26
Chapitre III	27
<b>III.1. Introduction :</b>	<b>28</b>
<b>III.2. Deep learning for Emotion Recognition :</b>	<b>29</b>
<b>III.3. System proposé :</b>	<b>30</b>
<b>III.4. Expériences et analyse des résultats :</b>	<b>31</b>
<b>III.4.1 Data :</b>	<b>32</b>
<b>III.4.1.1. Les bibliothèques :</b>	<b>32</b>
<b>III.4.1.2. Les classes :</b>	<b>33</b>
<b>III.4.1.3. L'analyse des données :</b>	<b>34</b>
<b>III.4.2 Matrice :</b>	<b>35</b>
<b>III.4.2.1 Matrice de Confusion :</b>	<b>35</b>
<b>III.4.2.2 Accuracy :</b>	<b>37</b>
<b>III.4.3 Results :</b>	<b>39</b>
<b>III.4.3.1. Epochs :</b>	<b>40</b>
<b>III.4.3.2. Pour 50 epochs :</b>	<b>41</b>
<b>III.4.3.3. Pour 200 epochs :</b>	<b>44</b>
<b>III.4.3.4. Pour 500 epochs :</b>	<b>46</b>
III.4.5. Comparaison :	49
<b>III.5. Conclusion</b>	<b>52</b>
<b>Conclusion générale</b>	<b>54</b>

## Liste de Figure

<b>Figure I.1:</b> Exemples de traits biométriques. ....	5
<b>Figure I.2 :</b> Un échantillon d'empreintes digitales.....	6
<b>Figure I.3:</b> image d'empreinte palmaire. ....	7
<b>Figure I.4:</b> Identification à l'aide de la géométrie de la main. ....	7
<b>Figure I.5:</b> La reconnaissance faciale. ....	8
<b>Figure I.6:</b> Reconnaissance de l'oreille. ....	8
<b>Figure I.7 :</b> Identification par ADN. ....	9
<b>Figure I.8 :</b> Reconnaissance de l'iris. ....	9
<b>Figure I.9:</b> Frappes. ....	10
<b>Figure I.10:</b> Identification à l'aide de la démarche. ....	10
<b>Figure I.11:</b> Reconnaissance des signatures. ....	11
<b>Figure I.12:</b> Reconnaissance vocale. ....	11
<b>Figure I.13 :</b> émotions de base.....	12
<b>Figure II.14:</b> Deep Learning comme une sorte d'apprentissage automatique. ....	18
<b>Figure II.15:</b> Diagramme KNN. ....	20
<b>Figure II.16:</b> Représentation hyperplane de classification de SVM. ....	22
<b>Figure II.17 :</b> Données séparables linéairement B : Données séparables non linéairement. ....	23
<b>Figure II.18 :</b> Exemple de transformation de fonction logistique. ....	24
<b>Figure II.19:</b> représentation de l'arbre de décision. ....	24
<b>Figure II.20:</b> Différences entre l'apprentissage en profondeur et l'apprentissage automatique. ....	25
<b>Figure II.21 :</b> Neurone biologique et neurone artificiel. ....	26
<b>Figure II.22:</b> Perceptron simple. ....	26
<b>Figure II.23:</b> un réseau de neurones convolutifs. ....	27
<b>Figure II.24:</b> Processus de classification CNN. ....	27
<b>Figure II.25:</b> Architecture et noyau CNN de base. ....	28
<b>Figure II.26:</b> Réseau neuronal récurrent simple. ....	29
<b>Figure II.27:</b> Architecture de LSTM. ....	30
<b>Figure II.28:</b> Architecture de l'apprentissage par transfert. ....	31
<b>Figure II.29:</b> Architecture VGG. ....	32
<b>Figure II.30:</b> Apprentissage résiduel d'un bloc de construction. ....	33
<b>Figure II.31:</b> L'architecture de notre réseau de base EfficientNet. ....	34
<b>Figure II.32:</b> Comparaison de l'apprentissage en profondeur avec l'apprentissage automatique. ....	35
<b>Figure III.33:</b> Training process of CNN model for facial emotion recognition.....	<b>Error!</b>
<b>Bookmark not defined.</b>	
<b>Figure III.34 :</b> Modèle de reconnaissance des émotions loss et accuracy graphiques Pour 50 epochs.....	46
<b>Figure III.35 :</b> Matrice de confusion des classes d'émotions Pour 50 epochs. ....	47
<b>Figure III.36:</b> Les résultats de l'exécution du programme pour 50 epochs. ....	47

<b>Figure III.37:</b> Modèle de reconnaissance des émotions loss et accuracy graphiques Pour 200 epochs.....	48
<b>Figure III.38:</b> Matrice de confusion des classes d'émotions Pour 200 epochs. ....	49
<b>Figure III.39 :</b> Les résultats de l'exécution du programme pour 200 epochs.....	49
<b>Figure III.40:</b> Modèle de reconnaissance des émotions loss et accuracy graphiques Pour 500 epochs.....	50
<b>Figure III.41:</b> Matrice de confusion des classes d'émotions Pour 500 epochs. ....	51
<b>Figure III.42:</b> Les résultats de l'exécution du programme pour 500 epochs.....	51

## Introduction générale

L'interaction émotionnelle est un phénomène psychologique courant dans la vie quotidienne de l'homme. La reconnaissance précise des émotions est la prémisse d'une communication, d'une interaction et d'une prise de décision humaines efficaces. Avec le développement de la technologie des mégadonnées et de l'intelligence artificielle, la recherche de systèmes de reconnaissance des émotions est devenue un projet typique à la fois dans le milieu universitaire et dans l'industrie.

Grâce à ces plateformes, les humains expriment leurs pensées, se connectent et communiquent entre eux. De cette façon, une grande quantité de données est générée, y compris des idées publiées, des sentiments et même des photos et des vidéos. Ces ressources contiennent de riches informations émotionnelles et fournissent une base de données pour la recherche liée aux émotions. D'autre part, les demandes des gens en matière de soins de santé mentale et de gestion émotionnelle augmentent continuellement, aidant à mieux se comprendre et à trouver un moyen plus efficace d'apprendre, de travailler et de vivre. Ces faits ont accéléré le besoin d'une reconnaissance fine des émotions.

La tâche de reconnaissance des émotions a suscité un intérêt considérable de la part de la communauté des chercheurs, visant à souligner les différences subtiles dans le contenu émotionnel. La reconnaissance textuelle des émotions (TER) identifie automatiquement les émotions humaines dans leur expression textuelle, telles que heureux, triste, en colère, craintif, excité, ennuyé ou autre. Cette tâche peut être facilement accomplie par les humains en fonction de leurs sentiments subjectifs. Néanmoins, pour les systèmes TER automatisés, les méthodologies de calcul doivent être continuellement développées et optimisées pour obtenir une prédiction plus précise des émotions.

L'apprentissage en profondeur est une sous-discipline de l'automatisation qui permet d'apprendre à partir de problèmes en utilisant des prix de neurones profonds artificiels. Les réseaux sont des modèles informatiques constitués de plusieurs couches de neurones, permettant une représentation hiérarchique des données capture des caractéristiques complexes et implémente des techniques de classification, de régression, de reconnaissance de formes et de création de contenu. [2]



L'objectif que nous avons essayé d'atteindre dans est d'essayer d'identifier les principales émotions à partir des images de nos visages en utilisant l'apprentissage en profondeur, et cela a été fait en utilisant sept sentiments, et ils sont (la colère, la tristesse, le bonheur, dégoût, étonnement, jouissance, état normal).

Cette note est divisée en trois chapitres principaux :

- Le premier chapitre contient une définition générale de la biométrie et des méthodes d'identification perceptive, physique et même biologique
- Le deuxième chapitre traite des définitions et des classifications de l'apprentissage automatique, de l'apprentissage en profondeur et de ses types, avec une explication détaillée de la méthode sur laquelle nous nous sommes appuyés pour l'apprentissage en profondeur, à savoir le réseau de neurones convolutifs CNN.
- Le dernier chapitre contient une analyse des résultats obtenus lors de l'exécution du programme et une tentative de comprendre ce qui s'est passé

# Chapitre I

---

---

# Reconnaissance des émotions

---

---

**I.1. Introduction :**

Les émotions faciales et leur analyse jouent un rôle essentiel dans la communication non verbale. Elle rend la communication orale plus efficace et propice à la compréhension des concepts.

Il est également propice à la détection de l'attention humaine, telle que le comportement, l'état mental, la personnalité, la tendance à la criminalité, les mensonges, etc. Indépendamment du sexe, de la nationalité, de la culture et de la race, la plupart des gens peuvent facilement reconnaître les émotions faciales.

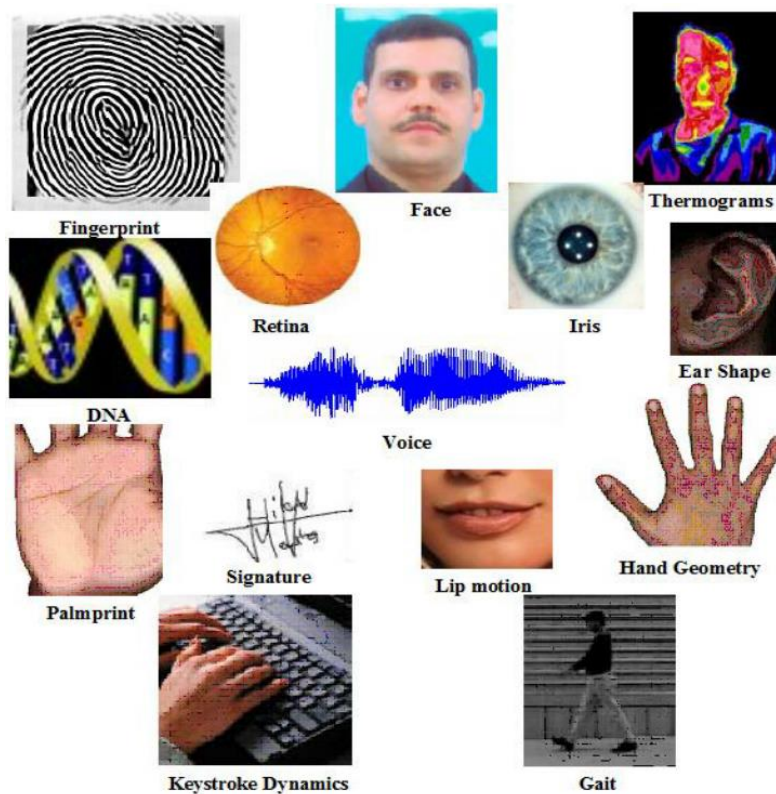
Cependant, une tâche difficile est l'automatisation de la détection et de la classification des émotions faciales. La communauté des chercheurs utilise quelques sentiments de base, tels que la peur, l'agressivité, la colère et le plaisir. Cependant, faire la différence entre de nombreux sentiments est très difficile pour les machines. De plus, les machines doivent être suffisamment entraînées pour comprendre spécifiquement l'environnement environnant, les intentions d'un individu.

Le principal problème est de classer les émotions des gens en fonction des variations de sexe, d'âge, de race, d'origine ethnique et de qualité d'image ou de vidéos. Il est nécessaire de fournir un système capable de reconnaître les émotions faciales avec des connaissances similaires à celles que possèdent les humains.

## I.2. Biométrie :

### I.2.1. Définition:

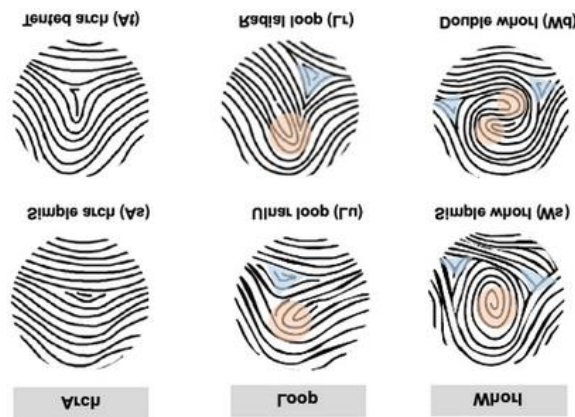
La biométrie est la science qui permet d'établir l'identité d'un individu sur la base des attributs physiques, chimiques ou comportementaux de la personne. La pertinence de la biométrie dans la société moderne a été renforcée par le besoin de systèmes de gestion d'identité à grande échelle dont la fonctionnalité repose sur la détermination précise de l'identité d'un individu dans le contexte de plusieurs applications différentes. Des exemples de ces applications incluent le partage de ressources informatiques en réseau, l'octroi d'accès à des installations nucléaires, la réalisation de transactions financières à distance ou l'embarquement sur un vol commercial. La prolifération des services Web (par exemple, les services bancaires en ligne) et le déploiement de centres de service à la clientèle décentralisés (par exemple, cartes de crédit) ont encore souligné le besoin de systèmes de gestion d'identité fiables pouvant accueillir un grand nombre d'individus. [1]



**Figure I.1:** Exemples de traits biométriques.

### I.3. Empreintes :

Une empreinte digitale est le motif de crêtes et de vallées à la surface d'un bout de doigt dont la formation est déterminée au cours des sept premiers mois du développement fœtal. C'est la méthode biométrique la plus utilisée. Les systèmes standard qui reposent sur cette méthode sont très courants et peuvent être trouvés dans de nombreux endroits. Peut être utilisé pour l'identification et la vérification.

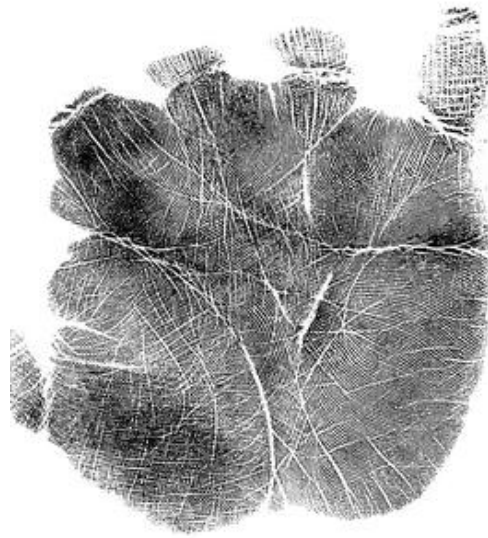


**Figure I.2 :** Un échantillon d'empreintes digitales.

#### I.3.1. Empreintes palmaires :

Les systèmes de reconnaissance d'empreintes palmaires utilisent un appareil de numérisation ou une application basée sur un appareil photo, ainsi qu'un logiciel associé qui traite les données d'image d'une photographie de la paume d'un individu et les compare à un enregistrement stocké pour cette personne. Les empreintes palmaires sont des homologues des empreintes digitales, y compris des détails similaires. Comme c'est le cas avec la numérisation d'empreintes digitales, les scanners de paume utilisent des méthodes optiques, thermiques ou tactiles pour faire ressortir les détails du motif des zones surélevées (appelées crêtes) et des branches (appelées bifurcations) dans une image d'une paume humaine, ainsi que d'autres détails.

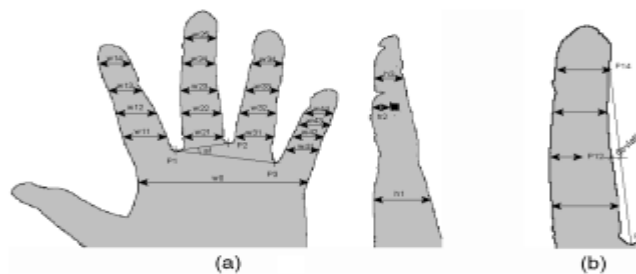
Compris les cicatrices, les plis et la texture. Ces trois méthodes reposent respectivement sur l'analyse de la lumière visible, l'analyse des émissions de chaleur et l'analyse de la pression.[2]



**Figure I.3:** image d'empreinte palmaire.

### I.3.2. Géométrie de la main :

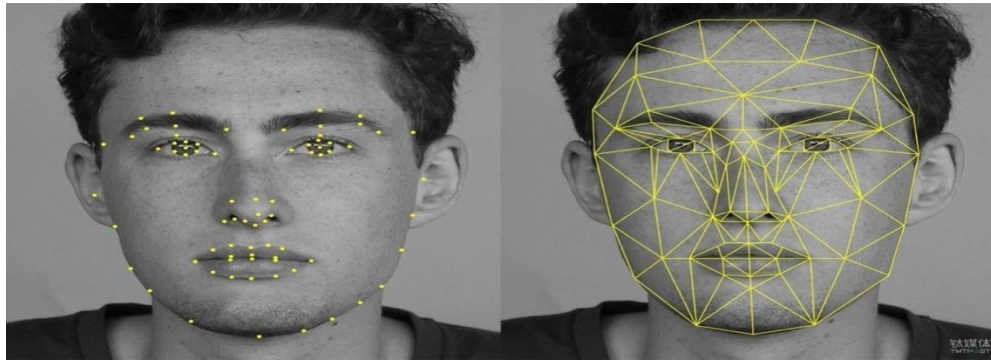
Les systèmes biométriques à géométrie de la main intègrent les principales caractéristiques de la géométrie des doigts, mais incluent également les surfaces de la main elle-même et son profil latéral. Les images sont prises alors que la main est maintenue paume vers le bas sur une plaque de support et maintenue en position par l'utilisation de chevilles de guidage. La longueur, la largeur, l'épaisseur et la surface de la main de l'individu sont mesurées et enregistrées. Plusieurs caractéristiques et mesures sont extraites au cours de ce processus. Plusieurs images de la même main sont souvent prises afin de produire un seul modèle stocké suffisamment détaillé à des fins d'identification. Ces images et les données à l'appui sont ensuite archivées dans une base de données et sont utilisées pour authentifier l'identité de l'enrôlé lors de rencontres ultérieures lorsque les mains du sujet sont à nouveau photographiées et comparées aux images de référence pour confirmer ou rejeter la revendication d'identité.



**Figure I.4:** Identification à l'aide de la géométrie de la main.

### I.3.3. Visage:

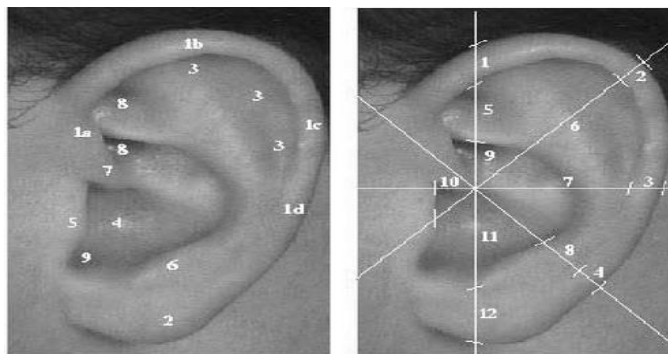
Cette technologie biométrique utilise des images du visage d'une personne pour vérifier son identité, elle est considérée comme une méthode robuste pour reconnaître les individus, cependant, un visage lui-même en tant que modalité peut changer avec l'âge, ou avec le maquillage et la chirurgie plastique, et même différentes expressions faciales peuvent avoir un effet sur le processus de reconnaissance. Il existe deux types de systèmes de reconnaissance faciale, ceux qui reconnaissent une personne dans un environnement connu, principalement utilisés pour l'authentification, et d'autres systèmes qui reconnaissent un individu parmi un groupe de personnes dans un environnement aléatoire.[1]



**Figure I.5:**La reconnaissance faciale.

### I.3.4. Oreille :

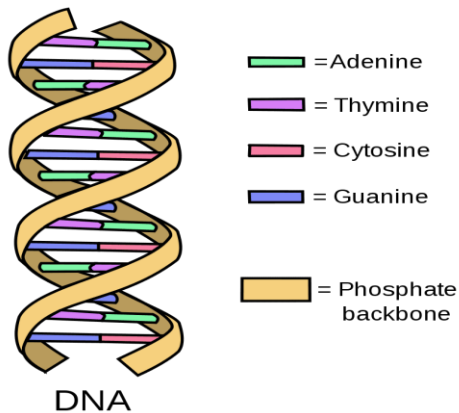
C'est une méthode de reconnaissance assez nouvelle, il a été suggéré que la forme de l'oreille est distincte d'une personne à l'autre. Une approche utilisée dans la reconnaissance de l'oreille consiste à vérifier la ligne max, qui est une ligne dont les deux extrémités se trouvent sur les bords de l'oreille.[1]



**Figure I.6:** Reconnaissance de l'oreille.

### I.3.5. ADN :

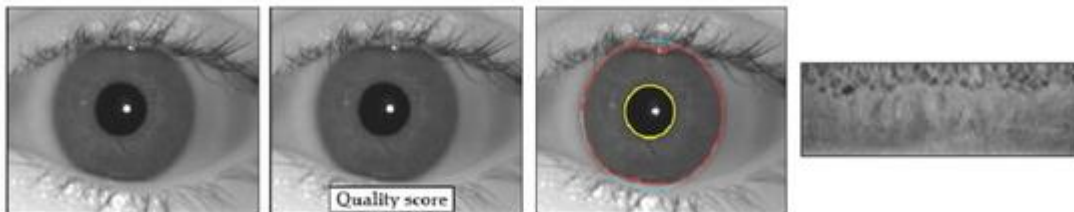
c'est l'abréviation de l'acide nucléique désoxyribose, il est unique pour chaque individu et il peut être trouvé dans chaque cellule du corps, ce qui en fait une modalité biométrique très fiable lorsqu'une identification positive est efficace car elle nécessite beaucoup de tests et de ressources, il est principalement utilisé dans les domaines juridiques et de l'application de la loi.[3]



**Figure I.7 :** Identification par ADN.

### I.3.6. Iris :

La reconnaissance de l'iris est une méthode automatisée d'identification biométrique, prenant des modèles uniques dans une région en forme d'anneau entourant la pupille de chaque œil. Il s'agit d'une méthode d'identification extrêmement fiable et précise avec des taux de fausses correspondances très faibles. Cependant, il présente certains inconvénients, tels que la nécessité d'un équipement matériel spécialisé.[4]



**Figure I.8 :** Reconnaissance de l'iris.



## I.4. Modalités Comportementales :

Voici les modalités biométriques qui s'appuient sur le comportement d'un individu pour établir son identité :

### I.4.1. frapper le clavier :

C'est l'une des dernières technologies en biométrie, la dynamique de frappe fonctionne en analysant la façon dont un individu tape, en prenant comme facteurs le temps et la vitesse de frappe de l'utilisateur, mais ce n'est pas trop stable, car l'humeur de la personne peut l'affecter et aussi le fait que la façon de taper change à mesure que la personne s'habitue de plus en plus à taper.[5]



Figure I.9: Frappes.

### I.4.2. Démarche :

La démarche fait référence à la manière dont une personne marche et est l'un des rares traits biométriques qui peuvent être utilisés pour reconnaître les personnes à distance. Par conséquent, ce trait est très approprié dans les scénarios de surveillance où l'identité d'un individu peut être subrepticement établie. La plupart des algorithmes de reconnaissance de la marche tentent d'extraire la silhouette humaine afin d'en déduire les attributs spatio-temporels d'un individu en mouvement. Par conséquent, la sélection d'un bon modèle pour représenter le corps humain est essentielle au fonctionnement efficace d'un système de reconnaissance de la marche. [1]

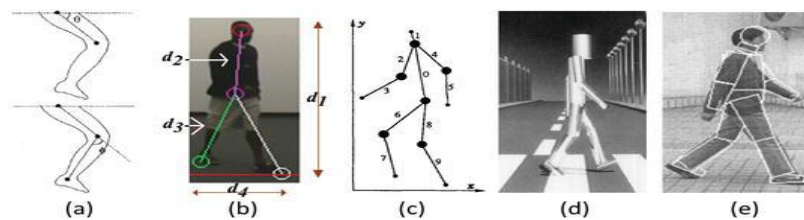


Figure I.10: Identification à l'aide de la démarche.

### I.4.3. Signature :

Appartenant à la famille des produits biométriques, la vérification de signature biométrique authentifie les signataires en mesurant leurs signatures manuscrites. La signature contient des données biométriques uniques, telles que le rythme d'écriture, l'accélération et la pression. Contrairement à d'autres méthodes de capture de signature électronique, la vérification de signature biométrique ne traite pas la signature comme une image graphique. Avec des images graphiques, telles que les signatures numérisées que nous attachons souvent à nos documents, il n'est pas possible de détecter la dynamique au sein de la signature de chaque individu et, par conséquent, les signatures peuvent être facilement copiées. En revanche, la vérification de signature biométrique mesure exactement comment la signature est signée. [6]

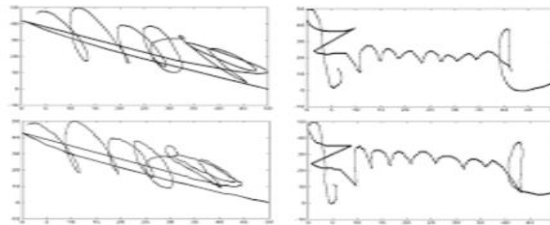


Figure I.11: Reconnaissance des signatures.

### I.4.4. Voix :

La reconnaissance vocale est également appelée reconnaissance du locuteur. Au moment de l'inscription, l'utilisateur doit prononcer un mot ou une phrase dans un microphone. Ceci est nécessaire pour acquérir un échantillon de la parole d'un candidat.

Le signal électrique du microphone est converti en signal numérique par un convertisseur analogique-numérique (ADC). Il est enregistré dans la mémoire de l'ordinateur sous forme d'échantillon numérisé. L'ordinateur compare ensuite et tente de faire correspondre la voix d'entrée du candidat avec l'échantillon de voix numérisé stocké et identifie le candidat. [7]

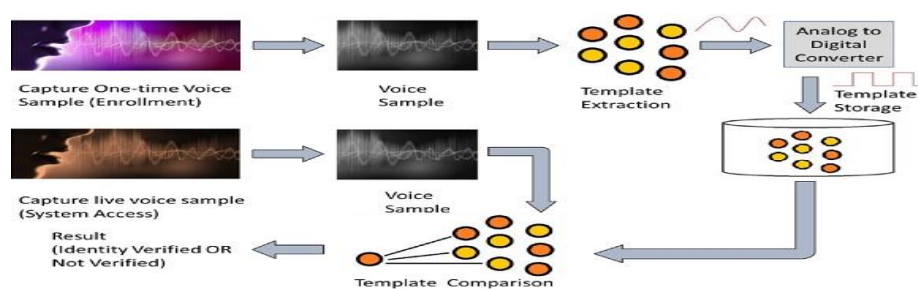


Figure I.12: Reconnaissance vocale.

## I.4.5. Émotion :

### I.4.5.1. Définition :

Les émotions sont un processus, un type particulier d'évaluation automatique influencé par notre passé évolutif et personnel, dans lequel nous sentons que quelque chose d'important pour notre bien-être se produit, et un ensemble de changements psychologiques et de comportements émotionnels commence à faire face à la situation.

### I.4.5.2. Le modèle d'émotion de base :

Les émotions de base ont évolué pour gérer les tâches fondamentales de la vie, par exemple, la peur et la colère peuvent aider à la survie en incitant un organisme à fuir pour sa sécurité ou à se battre pour se défendre. Les éléments des émotions de base peuvent être combinés pour former des émotions complexes ou composées. Même si de nombreux psychologues ont accepté la théorie des émotions de base, il n'y a pas de consensus sur le nombre précis d'émotions de base. Robert Plutchik a proposé huit émotions primaires : la colère, la peur, la tristesse, le dégoût, la surprise, l'anticipation, la confiance et la joie, et les a disposées dans une roue chromatique. Ekman a proposé sept émotions de base : la peur, la colère, la joie, la tristesse, le mépris, le dégoût et la surprise ; mais il est passé à six émotions de base : la peur, la colère, la joie, la tristesse, le dégoût et la surprise. Cependant, une étude récente a révélé que le dégoût et la colère partageaient un nez ridé similaire, et que la peur et la surprise partageaient des sourcils levés.



Figure I.13 : émotions de base.

### I.4.5.3. Le Modèle Dimensionnel :

Les émotions dans un cadre dimensionnel peuvent être cartographiées par deux ou trois variables, telles que la valence, l'excitation et l'énergie ou le contrôle. La dimension de valence représente généralement le degré positif ou négatif de l'émotion, et la gamme va des sentiments inconfortables aux sentiments confortables. La dimension d'excitation représente le degré d'excitation de l'émotion, et elle varie de faible à élevée. Alors que la dimension d'énergie ou de contrôle représente le degré d'énergie ou de contrôle sur l'émotion

### I.4.5.4. Reconnaissance des émotions :

La reconnaissance des émotions est le processus d'identification des émotions humaines. Cela a beaucoup à voir avec la lecture précise des expressions faciales, des gestes non verbaux et même du langage corporel - des caractéristiques de communication qui peuvent souvent parler plus fort que les mots ne le pourraient jamais. La reconnaissance des émotions est une technique utilisée dans les logiciels qui permet à un programme de "lire" les émotions sur un visage humain à l'aide d'un traitement d'image avancé.

### I.4.5.5. Caractéristique de l'émotion :

#### **Colère :**

Sourcils tirés vers le bas, paupières supérieures relevées, paupières inférieures relevées, fusion des lèvres retroussées, les lèvres peuvent être resserrées.

#### **Craindre :**

Avec les caractéristiques sous les sourcils relevés et rapprochés, les paupières supérieures relevées, la bouche tendue.

#### **Dégoûter :**

Les sourcils baissés, le nez plissé, la lèvre supérieure relevée, les lèvres lâches.

#### **Bonheur :**

Muscle autour de l'œil resserré, rides de la patte d'oie autour des yeux, joues relevées, coins des lèvres relevés en diagonale.

**✚ Tristesse :**

Coins intérieurs des sourcils levés, paupières lâches, coins des lèvres rabattus.

**✚ Surprendre :**

Sourcils entiers relevés, paupières relevées, bouche ouverte.

Il existe d'autres émotions telles que la honte, la fierté, la jalousie et la culpabilité. Bien que ces émotions soient importantes, elles ne sont toujours pas considérées comme faisant partie de l'ensemble des émotions de base. Ces caractéristiques sont des points visuellement saillants dans les régions faciales, appelées repères faciaux (FL). Ekman l'utilise pour développer le système de codage d'action faciale (FACS).

**✚ Système de codage des actions faciales (FACS) :**

Le système de codage des actions faciales (FACS) est le système le plus reconnu, le plus complet et le plus crédible pour mesurer les mouvements du visage discernables visuellement. Il a été développé par les psychologues Paul Ekman et Wallace V. Friesen dans le but de mesurer objectivement les expressions faciales pour les enquêtes en sciences du comportement [10].

Avant le développement du FACS, un système standard de mesure faciale était inexistant; cela a été très problématique pour la collecte de données scientifiques fiables dans le doma

## I.5. Conclusion

La reconnaissance des émotions faciales est une tâche complexe mais importante dans le domaine de la communication non verbale. Elle permet de mieux comprendre les concepts et d'améliorer l'efficacité de la communication orale. De plus, l'analyse des émotions faciales peut être utilisée pour détecter des comportements spécifiques, des états mentaux, des tendances à la criminalité, etc. La capacité à reconnaître les émotions faciales est présente chez la plupart des individus, indépendamment de leur sexe, nationalité, culture ou race.

Cependant, l'automatisation de la détection et de la classification des émotions faciales reste un défi. La communauté de recherche a identifié quelques émotions de base, telles que la peur, l'agressivité, la colère et le plaisir. Toutefois, différencier de nombreuses émotions est une tâche difficile pour les machines. De plus, les machines doivent être entraînées de manière spécifique pour comprendre l'environnement, les intentions individuelles et autres facteurs contextuels.

Le principal défi réside dans la classification précise des émotions en fonction des variations de sexe, d'âge, de race, d'origine ethnique et de qualité des images ou vidéos. Il est essentiel de développer des systèmes capables de reconnaître les émotions faciales avec une précision similaire à celle des humains.

En conclusion, la reconnaissance des émotions faciales est un domaine de recherche prometteur, qui présente des applications variées. Malgré les défis techniques, les progrès réalisés dans la conception de systèmes de reconnaissance des émotions faciales nous rapprochent de la création de systèmes plus intelligents et plus capables de comprendre les aspects non verbaux de la communication humaine.

## Chapitre II

---

---

# Méthode d'apprentissage en profondeur

---

---

## II.1. Introduction

L'apprentissage automatique et l'apprentissage profond sont deux domaines de l'intelligence artificielle qui ont révolutionné la façon dont les machines peuvent apprendre à partir des données. Dans ce chapitre, nous explorerons ces deux approches et discuterons de leurs différentes formes.

L'apprentissage automatique, également connu sous le nom de machine learning, est une discipline qui permet aux ordinateurs d'apprendre et d'améliorer leurs performances sans être explicitement programmés. Il repose sur l'idée que les machines peuvent apprendre des modèles et des relations à partir des données et utiliser ces connaissances pour prendre des décisions ou effectuer des prédictions. L'apprentissage automatique comprend diverses techniques telles que les arbres de décision, les machines à vecteurs de support (SVM), les réseaux de neurones artificiels, etc.

L'apprentissage profond, quant à lui, est une branche de l'apprentissage automatique qui se concentre sur l'utilisation de réseaux de neurones artificiels profonds pour résoudre des problèmes complexes. Les réseaux de neurones profonds sont des modèles computationnels composés de multiples couches de neurones, permettant ainsi une représentation hiérarchique des données. Ils sont capables d'apprendre des caractéristiques abstraites et de capturer des relations complexes entre les variables d'entrée. L'apprentissage profond a connu des avancées significatives avec l'introduction de architectures telles que les réseaux de neurones convolutionnels (CNN), les réseaux de neurones récurrents (RNN) et les réseaux de neurones générateurs adverses (GAN).

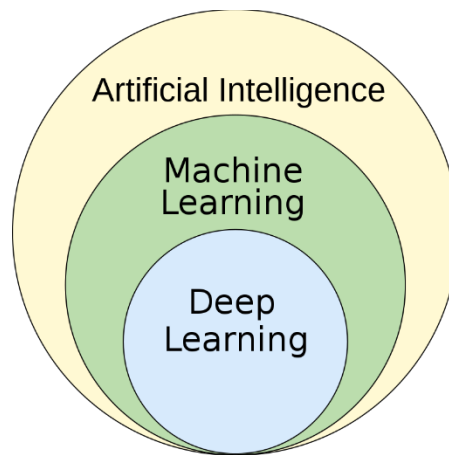
Au cours de ce chapitre, nous examinerons de près ces différentes formes d'apprentissage automatique et d'apprentissage profond, en soulignant leurs forces, leurs applications et leurs limitations. Nous discuterons également des considérations pratiques liées à l'entraînement et à l'évaluation de ces modèles, ainsi que des défis spécifiques rencontrés lors de leur mise en œuvre dans des scénarios réels. En comprenant ces concepts fondamentaux, nous pourrions mieux apprécier le potentiel et les implications de l'apprentissage automatique et de l'apprentissage profond dans le domaine de l'intelligence artificielle.



## II.2. Apprentissage automatique (le machine learning) :

### II.2.1. Identification :

L'apprentissage automatique est une branche de intelligence artificielle (IA) et l'informatique qui se concentre sur l'utilisation de données et d'algorithmes pour imiter la façon dont les humains apprennent, améliorant progressivement sa précision. [8]



**Figure II.14:** Deep Learning comme une sorte d'apprentissage automatique.

Le machine learning est une technique de modélisation qui fait intervenir des données, c'est-à-dire programmer des machines pour apprendre à partir de données, afin de construire des modèles représentatifs des classes que l'ordinateur délivre après apprentissage à partir des données. Ces dernières peuvent être des documents, des images, des sons, des vidéos, etc.[8]

L'apprentissage automatique est un élément important du domaine en plein essor de la science des données. Grâce à l'utilisation de méthodes statistiques, les algorithmes sont formés pour effectuer des classifications ou des prédictions, et pour découvrir des informations clés dans les projets d'exploration de données. Ces informations guident ensuite la prise de décision au sein des applications et des entreprises, impactant idéalement les mesures de croissance clés. À mesure que les données volumineuses continue de se développer et de se développer, la demande du marché pour les scientifiques des données augmentera.[8]

### II.2.2. Types d'apprentissage automatique :

Les problèmes qui ne peuvent pas être résolus uniquement par des méthodes numériques sont résolus à l'aide de l'apprentissage automatique. Bien qu'il existe d'autres techniques d'apprentissage automatique, l'apprentissage automatique supervisé et non supervisé sont deux des plus fréquemment utilisés. Voici une liste des types les plus courants.

#### II.2.2.1. Apprentissage automatique supervisé :

L'apprentissage supervisé, également connu sous le nom d'apprentissage automatique supervisé, est une sous-catégorie de apprentissage automatique et intelligence artificielle. Il est défini par son utilisation d'ensembles de données étiquetés pour former des algorithmes qui classent les données ou prédisent les résultats avec précision. Au fur et à mesure que les données d'entrée sont introduites dans le modèle, il ajuste ses pondérations jusqu'à ce que le modèle ait été ajusté de manière appropriée, ce qui se produit dans le cadre du processus de validation croisée. L'apprentissage supervisé aide les organisations à résoudre une variété de problèmes réels à grande échelle, tels que la classification du spam dans un dossier distinct de votre boîte de réception.

#### II.2.2.2. Apprentissage automatique non supervisé :

Apprentissage non supervisé, également appelé apprentissage automatique non supervisé, utilise des algorithmes d'apprentissage automatique pour analyser et regrouper des ensembles de données non étiquetés. Ces algorithmes découvrent des modèles ou des regroupements de données cachés sans intervention humaine. Sa capacité à découvrir les similitudes et les différences dans les informations en fait la solution idéale pour l'analyse exploratoire des données, les stratégies de vente croisée, la segmentation des clients et la reconnaissance d'image.

#### II.2.2.3. Apprentissage automatique semi-supervisé :

L'apprentissage automatique semi-supervisé est une combinaison de supervisé et apprentissage automatique non supervisé méthodes. Avec plus commun apprentissage automatique supervisé méthodes, vous former un apprentissage automatique algorithme sur un ensemble de données « étiqueté » dans lequel chaque enregistrement comprend les informations sur les résultats. Cela permet à l'algorithme de déduire des modèles et d'identifier

les relations entre votre cible variable et le reste de l'ensemble de données en fonction des informations dont il dispose déjà. En revanche, apprentissage automatique non supervisé algorithmes apprennent à partir d'un ensemble de données sans la variable de résultat. Dans l'apprentissage semi-supervisé, un algorithme apprend à partir d'un ensemble de données qui comprend à la fois des données étiquetées et non étiquetées, généralement pour la plupart non étiquetées.

### II.3. Algorithmes de classification d'apprentissage automatique :

Dans l'apprentissage automatique et les statistiques, la classification est une approche d'apprentissage supervisé dans laquelle le programme informatique apprend à partir des données d'entrée, puis utilise cet apprentissage pour classer de nouvelles observations. Cet ensemble de données peut simplement être bi-classe (comme identifier si la personne est un homme ou une femme ou si le courrier est un spam ou non) ou il peut être multi-classe. Voici quelques exemples pratiques de problèmes de classification : la reconnaissance vocale, la reconnaissance de l'écriture manuscrite, l'identification biométrique, la classification des documents, etc.[10]

#### II.3.1. K-Plus Proches Voisins (K-NN) :

L'algorithme des k plus proches voisins, également connu sous le nom de KNN ou k-NN, est un classificateur d'apprentissage supervisé non paramétrique, qui utilise la proximité pour effectuer des classifications ou des prédictions sur le regroupement d'un point de données individuel. Bien qu'il puisse être utilisé pour des problèmes de régression ou de classification, il est généralement utilisé comme algorithme de classification, en partant de l'hypothèse que des points similaires peuvent être trouvés les uns à côté des autres.

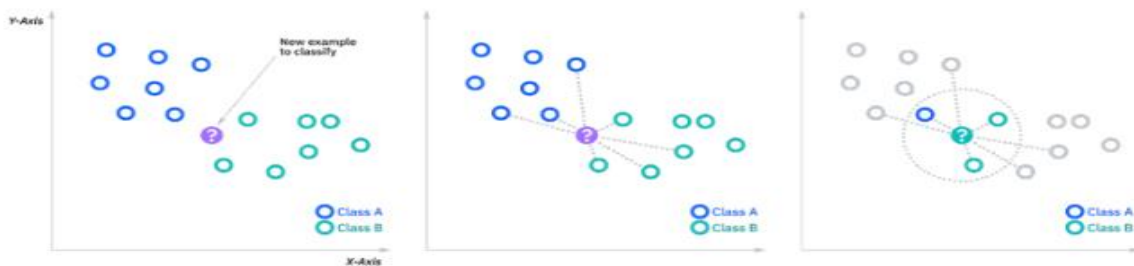


Figure II.15:Diagramme KNN.

Les problèmes de régression utilisent un concept similaire à celui de problème de classification, mais dans ce cas, la moyenne des  $k$  plus proches voisins est prise pour faire une prédiction sur une classification. La principale distinction ici est que la classification est utilisée pour les valeurs discrètes, tandis que la régression est utilisée avec les valeurs continues. Cependant, avant qu'une classification puisse être faite, la distance doit être définie. La distance euclidienne est la plus couramment utilisée, que nous aborderons plus en détail ci-dessous.

- ✚ **Distance euclidienne (p=2):** Il s'agit de la mesure de distance la plus couramment utilisée, et elle est limitée aux vecteurs à valeurs réelles. En utilisant la formule ci-dessous, il mesure une ligne droite entre le point de requête et l'autre point mesuré.

$$d(x, y) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2}$$

- ✚ **Manhattan distance (p=1):** Il s'agit également d'une autre mesure de distance populaire, qui mesure la valeur absolue entre deux points. Elle est également appelée distance en taxi ou distance d'un pâté de maisons, car elle est généralement visualisée avec une grille, illustrant comment on peut naviguer d'une adresse à une autre via les rues de la ville.

$$d(x, y) = \sqrt{\sum_{i=1}^m |x_i - y_i|}$$

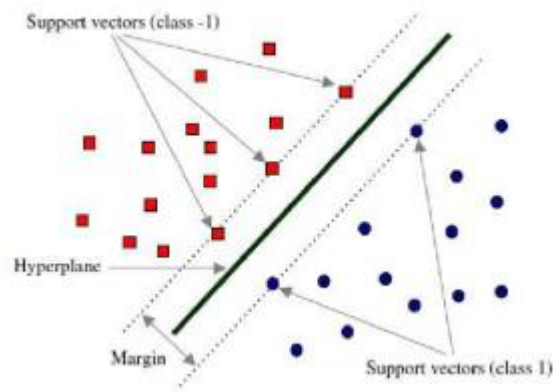
- ✚ **Distance de Hamming :** Cette technique est généralement utilisée avec des vecteurs booléens ou de chaîne, identifiant les points où les vecteurs ne correspondent pas. En conséquence, il a également été appelé la métrique de chevauchement. Ceci peut être représenté par la formule suivante :

$$d(x, y) = \left( \sum_{i=1}^k |x_i - y_i| \right)$$

### II.3.2. Machine à Vecteur Support (SVM) :

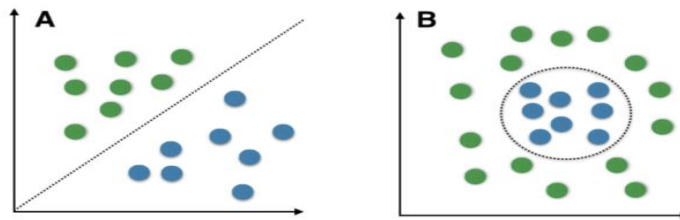
Il est un algorithme d'apprentissage automatique supervisé relativement simple utilisé pour la classification et/ou la régression. Il est plus préféré pour la classification mais est parfois très utile pour la régression également. Fondamentalement, SVM trouve un hyper-plan qui crée une frontière entre les types de données. Dans l'espace à 2 dimensions, cet hyperplan n'est rien d'autre qu'une ligne. Dans SVM, nous traçons chaque élément de données dans l'ensemble de données dans un espace à N dimensions, où N est le nombre de caractéristiques/attributs dans les données. Ensuite, trouvez l'hyperplan optimal pour séparer les données. Donc, par là, vous devez avoir compris que, par nature, SVM ne peut effectuer qu'une classification binaire (c'est-à-dire choisir entre deux classes). Cependant, il existe différentes techniques à utiliser pour les problèmes multi-classes.

✚ **En gros SVM:** est un algorithme de classification des problèmes binaires, bien sûr lorsque nous traitons des classes séparables linéairement, la figure.I.3 montre un hyperplan créé par le svm pour diviser l'espace des caractéristiques en segments de classe, l'hyperplan aura la plus grande marge possible entre les deux Des classes. Cette maximisation des marges est le concept essentiel des SVM.



**Figure II.16:** Représentation hyperplane de classification de SVM.

les points de données des deux classes qui sont parallèles au vecteur définissant l'hyperplan. L'algorithme tire son nom d'eux. L'approche peut également être utilisée pour résoudre des problèmes où les exemples ne peuvent pas être séparés de manière linéaire. La figure.I.4 montre clairement que les données ne peuvent pas être séparées linéairement ou qu'il n'y a pas de limite de décision linéaire (une seule ligne séparant les deux balises).



**Figure II.17 :**Données séparables linéairement B : Données séparables non linéairement.

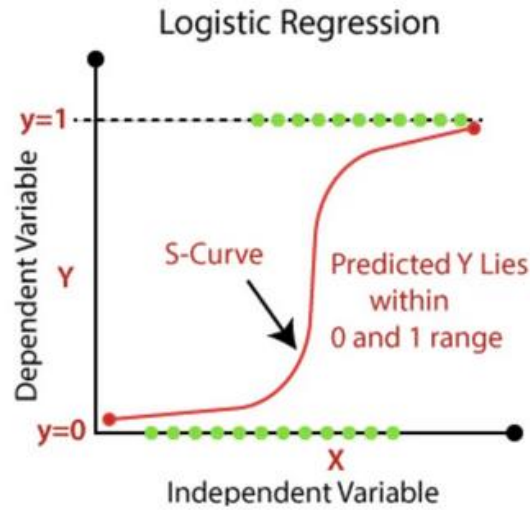
Nous utilisons Kernelized SVM pour les données non linéairement séparables. Supposons que nous disposions de données non linéairement séparables dans une dimension. Nous pouvons transformer ces données en deux dimensions et les données deviendront linéairement séparables en deux dimensions. Cela se fait en mappant chaque point de données 1-D à une paire ordonnée 2-D correspondante. Ainsi, pour toutes les données non linéairement séparables dans n'importe quelle dimension, nous pouvons simplement mapper les données sur une dimension supérieure, puis les rendre linéairement séparables. Il s'agit d'une transformation très puissante et générale.[12]

### II.3.3. Régression Logistique :

Ce type de modèle statistique (également appelé modèle logit) est souvent utilisé pour la classification et l'analyse prédictive. La régression logistique estime la probabilité qu'un événement se produise, tel que voter ou ne pas voter, sur la base d'un ensemble de données donné de variables indépendantes. Puisque le résultat est une probabilité, la variable dépendante est bornée entre 0 et 1. Dans la régression logistique, une transformation logit est appliquée sur la cote, c'est-à-dire la probabilité de succès divisée par la probabilité d'échec. Ceci est également connu sous le nom de log odds, ou le logarithme naturel des odds, et cette fonction logistique est représentée par les formules suivantes :

$$\log(p_i) = \frac{1}{(1 + \exp(-p_i))}$$

$$\ln(\frac{p_i}{1-p_i}) = \text{Beta}_0 + \text{Beta}_1 * X_1 + \dots + B_k * K_k$$



**Figure II.18 :**Exemple de transformation de fonction logistique.

Dans cette équation de régression logistique, logit (pi) est la variable dépendante ou de réponse et x est la variable indépendante. Le paramètre bêta, ou coefficient, dans ce modèle est généralement estimé via l'estimation du maximum de vraisemblance (MLE).[13]

II.3.4. Arbre décisionnel :

C'est une façon d'afficher un algorithme qui ne contient que des instructions de contrôle conditionnelles. Les arbres de décision sont couramment utilisés dans recherche opérationnelle, précisément dans analyse décisionnelle, pour aider à identifier une stratégie la plus susceptible d'atteindre un objectif, mais sont également un outil populaire dans apprentissage automatique. [14]



**Figure II.19:**représentation de l'arbre de décision.

II.3.5 Classificateur naïf de Bayes :

Un classificateur Naive Bayes est un modèle d'apprentissage automatique probabiliste utilisé pour la tâche de classification. Le nœud du classificateur est basé sur le théorème de Bayes.

$$P\left(\frac{A}{B}\right) = \frac{P\left(\frac{A}{B}\right)P(A)}{P(B)}$$

En utilisant le théorème de Bayes, nous pouvons trouver la probabilité que A se produise, étant donné que B s'est produit. Ici, B est la preuve et A est l'hypothèse. L'hypothèse faite ici est que les prédicteurs/caractéristiques sont indépendants. C'est la présence d'une caractéristique particulière n'affecte pas l'autre. C'est pourquoi on l'appelle naïf.[15]

II.4 Apprentissage profond :

II.1 Identification :

L'apprentissage en profondeur est un sous-ensemble de apprentissage automatique, qui est essentiellement un réseau de neurones à trois couches ou plus. Ces réseaux de neurones tentent de simuler le comportement du cerveau humain, bien que loin de correspondre à ses capacités, lui permettant «d'apprendre» à partir de grandes quantités de données. Alors qu'un réseau de neurones avec une seule couche peut toujours faire des prédictions approximatives, des couches cachées supplémentaires peuvent aider à optimiser et à affiner la précision.

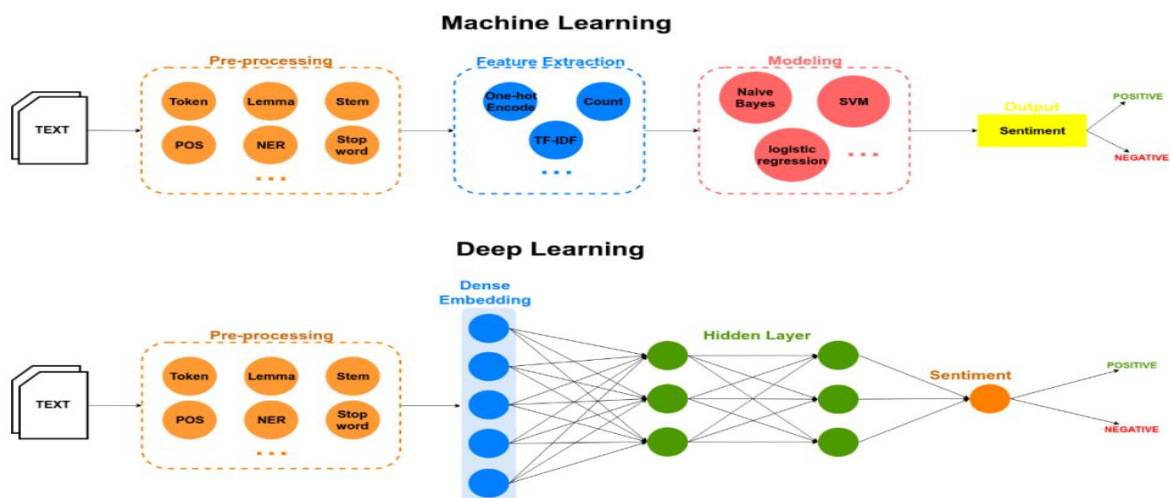


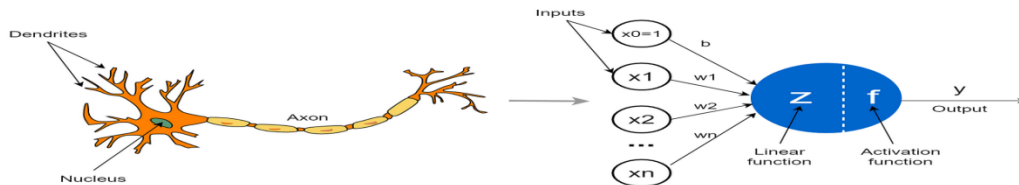
Figure II.20: Différences entre l'apprentissage en profondeur et l'apprentissage automatique.



## II.2 Types de méthode d'apprentissage en profondeur :

### II.2.1 Réseaux de neurones :

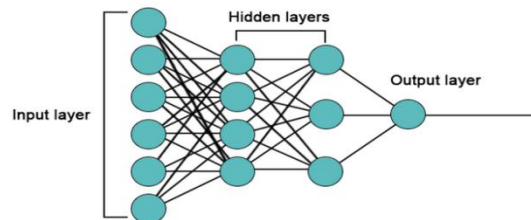
Les réseaux de neurones, également appelés réseaux de neurones artificiels (ANN) ou réseaux de neurones simulés (SNN), sont un sous-ensemble de apprentissage automatique et sont au cœur de l'apprentissage en profondeur algorithmes. Leur nom et leur structure sont inspirés du cerveau humain, imitant la façon dont les neurones biologiques se signalent les uns aux autres.[18]. Les réseaux de neurones s'appuient sur les données d'entraînement pour apprendre et améliorer leur précision au fil du temps. Cependant, une fois que ces algorithmes d'apprentissage sont affinés pour plus de précision, ils deviennent de puissants outils en informatique et intelligence artificielle, nous permettant de classer et de regrouper les données à grande vitesse. Les tâches de reconnaissance vocale ou de reconnaissance d'images peuvent prendre quelques minutes au lieu d'heures par rapport à l'identification manuelle par des experts humains. L'algorithme de recherche de Google est l'un des réseaux de neurones les plus connus.



**Figure II.21 :**Neurone biologique et neurone artificiel.

### II.2.2 Réseaux de Neurones Artificiels :

Les réseaux de neurones artificiels (ANN) sont composés d'une couche de nœuds, contenant une couche d'entrée, une ou plusieurs couches cachées et une couche de sortie. Chaque nœud, ou neurone artificiel, se connecte à un autre et a un poids et un seuil associés. Si la sortie d'un nœud individuel est supérieure à la valeur de seuil spécifiée, ce nœud est activé, envoyant des données à la couche suivante du réseau. Sinon, aucune donnée n'est transmise à la couche suivante du réseau.



**Figure II.22:**Perceptron simple.

II.2.3 Réseaux de neurones convolutifs (CNN) :

Les réseaux de neurones sont un sous-ensemble de l'apprentissage automatique et ils sont au cœur des algorithmes d'apprentissage en profondeur. Ils sont composés de couches de nœuds, contenant une couche d'entrée, une ou plusieurs couches cachées et une couche de sortie. Chaque nœud se connecte à un autre et a un poids et un seuil associés. Si la sortie d'un nœud individuel est supérieure à la valeur de seuil spécifiée, ce nœud est activé, envoyant des données à la couche suivante du réseau. Sinon, aucune donnée n'est transmise à la couche suivante du réseau.

Le chiffremontre une matrice d'intégration d'entrée traitée par quatre couches de convolution et deux couches de regroupement maximum. Les deux premières couches de convolution ont 64 et 32 filtres, qui sont utilisés pour former différentes fonctionnalités ; celles-ci sont suivies d'une couche de mise en commun maximale, qui est utilisée pour réduire la complexité de la sortie et pour empêcher le sur ajustement des données. Les troisième et quatrième couches de convolution ont respectivement 16 et 8 filtres, qui sont également suivis d'une couche de regroupement max. La couche finale est une couche entièrement connectée qui réduira le vecteur de hauteur 8 à un vecteur de sortie de un, étant donné qu'il y a deux classes à prédire (Positif, Négatif).

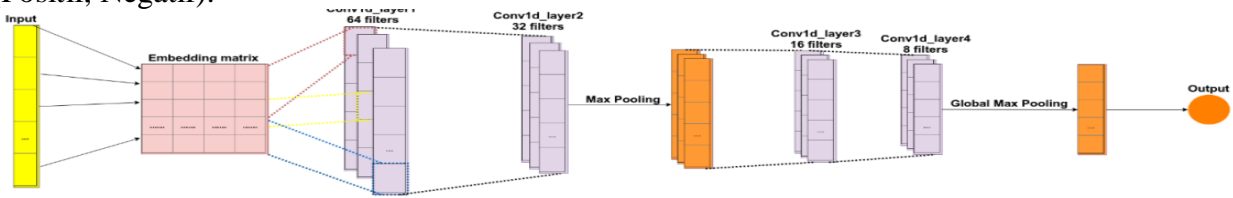


Figure II.23: un réseau de neurones convolutifs.

II.2.4 Couche de convolution :

La couche convolutive est le bloc de construction central d'un CNN, vérifiant si la caractéristique est présente. Ce processus est connu sous le nom de convolution.

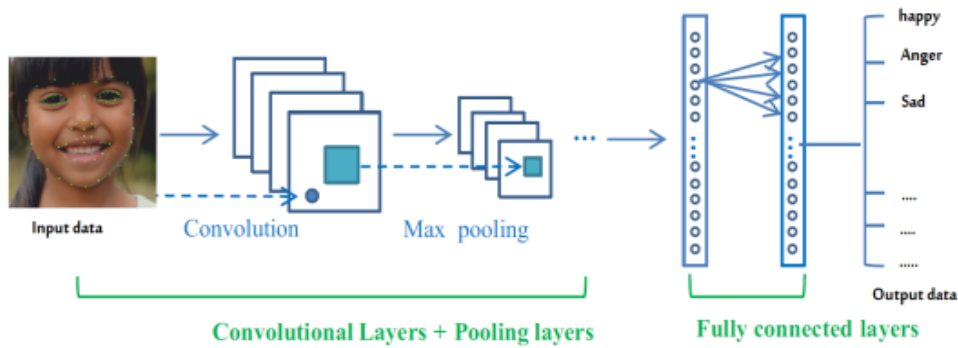


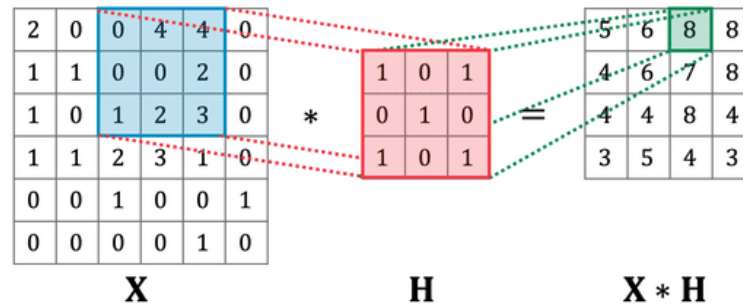
Figure II.24: Processus de classification CNN.

## II.2.5 Couche mutualisation :

Le regroupement des couches, également connu sous le nom de sous-échantillonnage, permet de réduire la dimensionnalité, réduisant ainsi le nombre de paramètres dans l'entrée. Le noyau applique une fonction d'agrégation aux valeurs dans le champ réceptif, remplissant le tableau de sortie. Il existe deux grands types de mutualisation :

- ✚ **Mise en commun maximale:** Lorsque le filtre se déplace sur l'entrée, il sélectionne le pixel avec la valeur maximale à envoyer au tableau de sortie.
- ✚ **Mise en commun moyenne:** Lorsque le filtre se déplace sur l'entrée, il calcule la valeur moyenne dans le champ récepteur à envoyer au tableau de sortie.

Bien que de nombreuses informations soient perdues dans la couche de mise en commun, cela présente également un certain nombre d'avantages pour le CNN. Ils contribuent à réduire la complexité, à améliorer l'efficacité et à limiter le risque de surajustement.



**Figure II.25:** Architecture et noyau CNN de base.

## II.2.6 Couche entièrement connectée :

Les valeurs de pixel de l'image d'entrée ne sont pas directement connectées à la couche de sortie dans des couches partiellement connectées. Cependant, dans la couche entièrement connectée, chaque nœud de la couche de sortie se connecte directement à un nœud de la couche précédente.

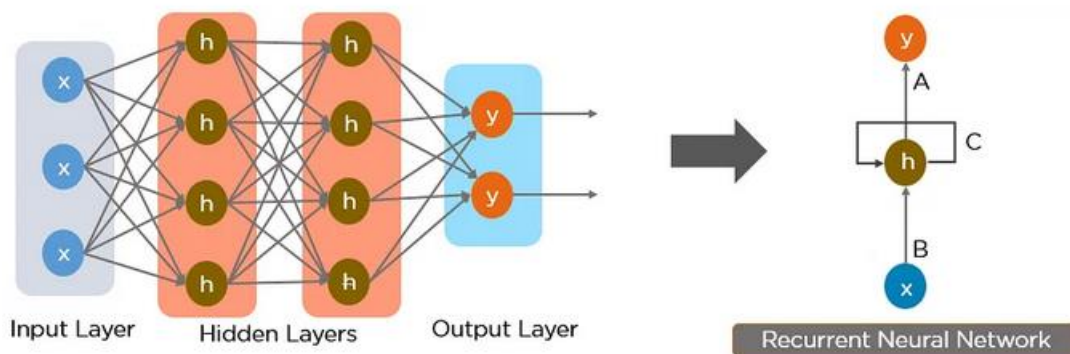
Cette couche effectue la tâche de classification basée sur les caractéristiques extraites à travers les couches précédentes et leurs différents filtres. Alors que les couches convolutionnelles et de regroupement ont tendance à utiliser les fonctions ReLu, les couches FC exploitent généralement une fonction d'activation softmax pour classer les entrées de manière appropriée, produisant une probabilité de 0 à 1.

## II.2.7 Réseaux de Neurones Récurrents (RNN) ;

Un réseau neuronal récurrent (RNN) est un type de réseau neuronal artificiel qui utilise des données séquentielles ou des données de séries chronologiques. Ces algorithmes d'apprentissage en profondeur sont couramment utilisés pour les problèmes ordinaux ou temporels, tels que la traduction linguistique, le traitement du langage naturel (nlp), la reconnaissance vocale et le sous-titrage d'images; Ils se distinguent par leur "mémoire" car ils prennent des informations d'entrées précédentes pour influencer l'entrée et la sortie actuelles.

Alors que les réseaux de neurones profonds traditionnels supposent que les entrées et les sorties sont indépendantes les unes des autres, la sortie des réseaux de neurones récurrents dépend des éléments antérieurs de la séquence. [19]

La fonction d'activation détermine si un neurone doit être activé. Les fonctions non linéaires convertissent généralement la sortie d'un neurone donné en une valeur comprise entre 0 et 1 ou -1 et 1.



**Figure II.26:**Réseau neuronal récurrent simple.

Les nœuds des différentes couches du réseau neuronal sont compressés pour former une seule couche de réseaux neuronaux récurrents. A, B et C sont les paramètres du réseau.[19]

II.2.8 Mémoire longue à court terme :

Est un réseau neuronal séquentiel d'apprentissage en profondeur qui permet aux informations de persister. Il s'agit d'un type spécial de réseau neuronal récurrent capable de gérer le problème de gradient de fuite auquel est confronté RNN. LSTM a été conçu par Hochreiter et Schmidhuber qui résout le problème causé par les rnns traditionnels et les algorithmes d'apprentissage automatique. [20] Chaque cellule LSTM calcule de nouvelles valeurs d'état caché et d'état de cellule, la formulation mathématique étant donnée ci-dessous.

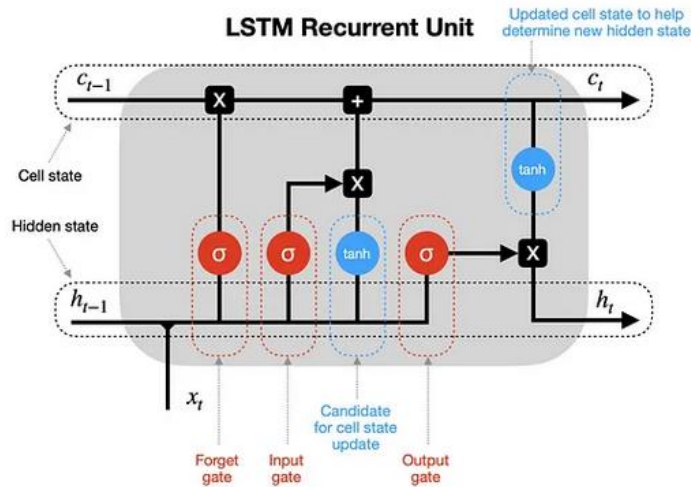


Figure II.27: Architecture de LSTM.

$$f_t = \alpha(x_t W_f + H_{t-1} U_f)$$

$$i_t = \alpha(x_t W_i + H_{t-1} U_i)$$

$$o_t = \alpha(x_t W_o + H_{t-1} U_o)$$

$$H_t = \tanh(x_t W_g + H_{t-1} U_g)$$

$$C_t = \alpha(c_{t-1} * f_t + i_t * H_t)$$

$$H_t = \tanh(C_t) * o_t$$

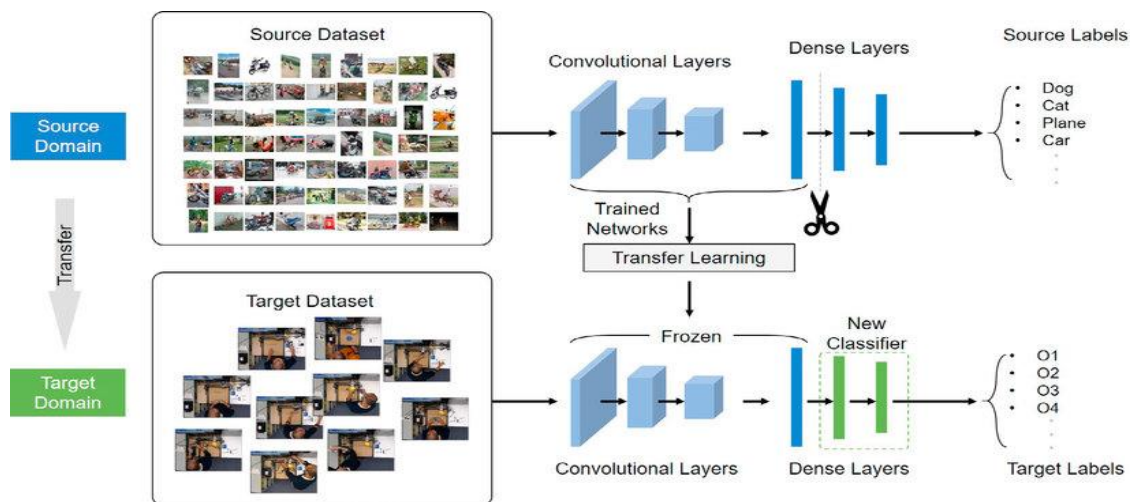
Où  $i, f$  et  $o$  sont entrés, oubliés et sortis respectivement.  $H$  et  $C$  sont respectivement l'état caché et l'état de la mémoire. De nombreuses recherches ont combiné deux ou trois modèles de réseaux de neurones, tels que CNN, RNN ou d'autres types d'ANN, pour bénéficier de leurs avantages, et ils ont atteint une bonne précision.

## II.2.9 Réseau Convolutif Temporel (TCN) :

Le réseau convolutif temporel (TCN) est un cadre qui utilise des convolutions et des dilatations occasionnelles afin qu'il soit adaptatif pour les données séquentielles avec sa temporalité et ses grands champs récepteurs. TCN est particulièrement utile pour extraire des fonctionnalités de haut niveau dans les données structurées.

## II.2.10 Apprentissage par transfert :

L'apprentissage par transfert, utilisé dans l'apprentissage automatique, est la réutilisation d'un modèle pré-entraîné sur un nouveau problème. Dans l'apprentissage par transfert, une machine exploite les connaissances acquises lors d'une tâche précédente pour améliorer la généralisation sur une autre. [21] Dans l'apprentissage par transfert, nous essayons de transférer autant de connaissances que possible de la tâche précédente sur laquelle le modèle a été formé à la nouvelle tâche à accomplir. Cette connaissance peut prendre différentes formes selon le problème et les données.



**Figure II.28:** Architecture de l'apprentissage par transfert.

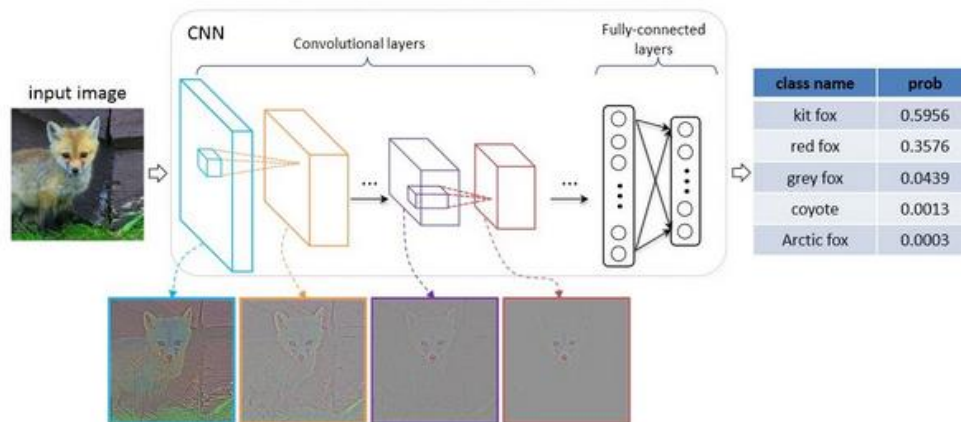
En vision par ordinateur, par exemple, les réseaux de neurones tentent généralement de détecter les bords dans les couches précédentes, les formes dans la couche intermédiaire et certaines caractéristiques spécifiques à une tâche dans les couches ultérieures. Dans l'apprentissage par transfert, les couches précoces et intermédiaires sont utilisées et nous ne réentraînon que les dernières couches. Il aide à tirer parti des données étiquetées de la tâche sur laquelle il a été initialement formé.

Dans l'exemple suivant, un modèle est formé pour reconnaître un sac à dos sur une image, qui sera utilisée pour identifier des lunettes de soleil. Dans les couches précédentes, le modèle a appris à reconnaître les objets, c'est pourquoi nous allons simplement recycler les dernières couches jusqu'à ce qu'il apprenne ce qui sépare les lunettes de soleil des autres objets.

### II.2.11 VGG :

VGG est un modèle innovant de reconnaissance d'objets qui prend en charge jusqu'à 19 couches. Construit comme un CNN profond, VGG surpasse également les lignes de base sur de nombreuses tâches et ensembles de données en dehors d'ImageNet. VGG est aujourd'hui encore l'une des architectures de reconnaissance d'images les plus utilisées. [22]

- ✚ **VGG16** : Le modèle VGG, ou VGGNet, qui prend en charge 16 couches est également appelé VGG16, qui est un modèle de réseau neuronal convolutionnel proposé par A. Zisserman et K. Simonyan de l'Université d'Oxford.
- ✚ **VGG19** : Le concept du modèle VGG19 (également VGGNet-19) est le même que celui du VGG16, sauf qu'il prend en charge 19 couches. Le « 16 » et le « 19 » représentent le nombre de couches de poids dans le modèle (couches convolutives).



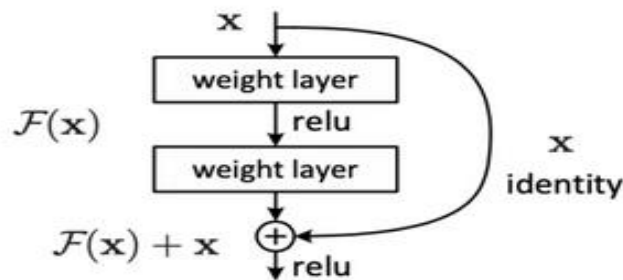
**Figure II.29:** Architecture VGG.

Les VGG sont basés sur les caractéristiques les plus essentielles des réseaux de neurones convolutifs (CNN). Le graphique montre le concept de base du fonctionnement d'un CNN.



## II.2.12 ResNet :

Un réseau neuronal résiduel (ResNet) est un modèle d'apprentissage en profondeur dans lequel les couches de poids apprennent des fonctions résiduelles en référence aux entrées de la couche. Un réseau résiduel est caractérisé comme un réseau avec des connexions de saut qui effectuent des mappages d'identité, fusionnés avec les sorties de couche par addition. Les réseaux résiduels ont permis aux modèles d'apprentissage en profondeur avec des dizaines ou des centaines de couches de s'entraîner facilement et d'approcher une meilleure précision lorsqu'ils vont plus loin. L'identité ignore les connexions, souvent appelées "connexions résiduelles". [23]



**Figure II.30:** Apprentissage résiduel d'un bloc de construction.

L'image ci-dessus montre un bloc résiduel typique. Cela peut être exprimé en code Python en utilisant l'expression  $\text{output} = F(x) + x$  où  $x$  est une entrée du bloc résiduel et la sortie de la couche précédente, et  $F(x)$  fait partie d'un CNN composé de plusieurs blocs convolutionnels. Cette technique lisse le flux de gradient lors de la rétro propagation, permettant au réseau de passer à 50, 100 ou même 150 couches. Ignorer une connexion n'ajoute pas de charge de calcul supplémentaire au réseau.

✚ **DenseNet** : est une autre variante populaire de ResNet, qui tente de résoudre le problème de la disparition des gradients en créant davantage de connexions. Les auteurs de DenseNet ont assuré le maximum de flux d'informations entre les couches du réseau en connectant chaque couche directement à toutes les autres. Ce modèle préserve les capacités d'anticipation en permettant à chaque couche d'obtenir des entrées supplémentaires de ses couches précédentes et de transmettre la carte d'entités aux couches suivantes.



✚ **Large ResNet:** est une version plus récente et améliorée du modèle original Deep ResNet. Il permet aux utilisateurs d'élargir ou de réduire le réseau sans affecter les performances. Il ne repose pas sur des augmentations de profondeur pour améliorer la précision du réseau.

II.2.13 EfficientNet:

EfficientNet est une architecture de réseau neuronal convolutif et une méthode de mise à l'échelle qui met uniformément à l'échelle toutes les dimensions de profondeur, largeur, résolution à l'aide d'un coefficient composé.

L'efficacité de la mise à l'échelle du modèle dépend également fortement du réseau de référence. Pour améliorer encore les performances, ce qui optimise à la fois la précision et l'efficacité (FLOPS). L'architecture résultante utilise la convolution de goulot d'étranglement inversé mobile (MBConv), similaire à MobileNetV2 et MnasNet, mais est légèrement plus grande en raison d'un budget FLOP accru. Nous augmentons ensuite le réseau de base pour obtenir une famille de modèles, appelés EfficientNets.

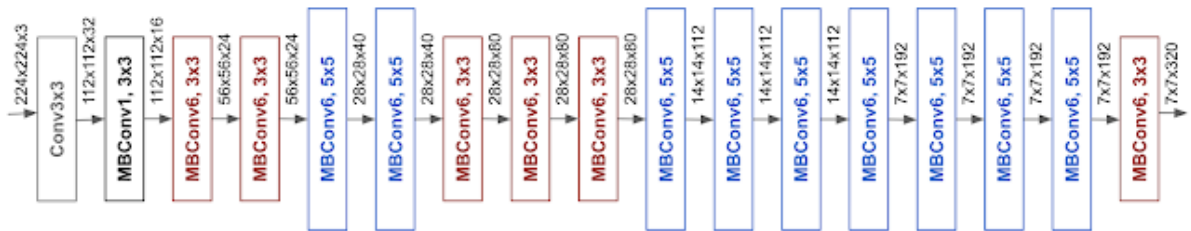
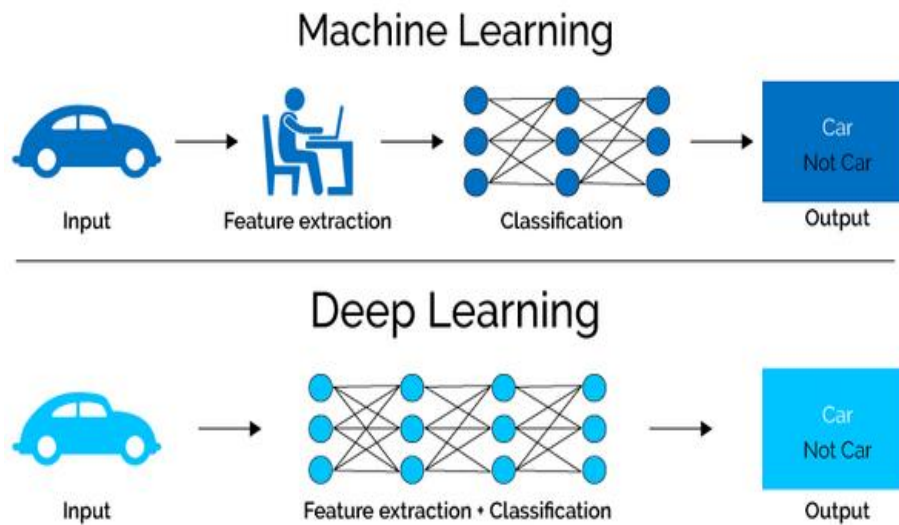


Figure II.31: L'architecture de notre réseau de base EfficientNet.

En général, les modèles EfficientNet atteignent à la fois une plus grande précision et une meilleure efficacité par rapport aux CNN existants, réduisant la taille des paramètres et les FLOPS d'un ordre de grandeur.

### II.3. Comparaison du deep learning avec le machine learning :

Étant donné que l'apprentissage en profondeur et l'apprentissage automatique ont tendance à être utilisés de manière interchangeable, il convient de noter les nuances entre les deux. L'apprentissage automatique, l'apprentissage en profondeur et les réseaux de neurones sont tous des sous-domaines de l'intelligence artificielle. Cependant, les réseaux de neurones sont en fait un sous-domaine de l'apprentissage automatique, et l'apprentissage en profondeur est un sous-domaine des réseaux de neurones.



**Figure II.32:** Comparaison de l'apprentissage en profondeur avec l'apprentissage automatique.

**II.4. conclusion :**

En conclusion, l'apprentissage automatique et l'apprentissage profond sont des domaines en constante évolution de l'intelligence artificielle qui ont ouvert de nouvelles perspectives passionnantes pour résoudre des problèmes complexes. L'apprentissage automatique offre des techniques puissantes pour extraire des informations à partir des données, tandis que l'apprentissage profond permet une représentation hiérarchique et une compréhension plus profonde des modèles sous-jacents.

Nous avons également exploré différentes formes d'apprentissage automatique et d'apprentissage profond, telles que les arbres de décision, les SVM, les réseaux de neurones artificiels, les CNN, les RNN et les GAN. Chaque forme a ses avantages et ses limites, et leur choix dépendra du problème spécifique et des données disponibles.

Il est important de noter que l'apprentissage automatique et l'apprentissage profond ne sont pas des solutions miracles, mais plutôt des outils puissants qui nécessitent une compréhension approfondie et une expertise pour être utilisés efficacement. L'entraînement et l'évaluation rigoureux des modèles sont essentiels pour garantir des performances optimales et éviter les biais indésirables. Dans les prochaines sections de ce chapitre, nous plongerons plus en profondeur dans les détails techniques de chaque forme d'apprentissage, en explorant des exemples concrets d'applications et en discutant des défis spécifiques liés à leur mise en œuvre.

En somme, l'apprentissage automatique et l'apprentissage profond continuent de redéfinir la façon dont les machines peuvent comprendre et interagir avec le monde qui les entoure. En développant une compréhension solide de ces domaines, nous serons en mesure de relever de nouveaux défis et de tirer parti de leur potentiel pour créer des systèmes intelligents et novateurs.

## Chapitre III

---

---

# Résultat et discussions

---

---

### III.1. Introduction :

L'expression faciale des humains est un moyen majeur de transmettre l'émotion car c'est le signal le plus puissant, le plus naturel et le plus universel pour exprimer l'état émotionnel des humains. Cependant, l'expression faciale des humains a des modèles similaires, et il est très déroutant de reconnaître l'expression à l'œil nu.

L'analyse du visage et de l'expression à l'aide d'ordinateurs est un nouveau domaine. L'analyse des émotions consiste à faire correspondre un visage à une émotion. L'objectif est donc de lire les sentiments d'une personne sur son visage. Les systèmes automatisés d'analyse de l'expression du visage facilitent l'interaction homme-machine. Mais ce n'est pas un processus facile. De nombreuses caractéristiques des expressions faciales peuvent désormais être récupérées et évaluées pour une bonne analyse des sentiments à l'aide d'apprentissage en profondeur et de réseaux de neurones convolutifs (CNN)

### III.2. Deep learning for Emotion Recognition :

Dans cette étude, un modèle DCNN pré-entraîné a été utilisé en utilisant la technique de transfert d'apprentissage (TL) pour la reconnaissance des émotions, ce qui constitue la contribution principale de l'étude. Une étude antérieure de Mahendran et Vedaldi a montré une visualisation de ce que les couches CNN apprennent visuellement. Dans la première couche du CNN, des caractéristiques de base telles que les contours et les angles de l'image sont capturées. Dans la couche suivante, le réseau est capable de découvrir des caractéristiques plus complexes telles que les textures et les formes, et la couche supérieure suit le même mécanisme pour apprendre des motifs encore plus complexes. Étant donné que les caractéristiques de base sont similaires dans toutes les images, les tâches liées à la reconnaissance des émotions dans les couches inférieures du modèle DCNN sont similaires à d'autres opérations basées sur les images, telles que la classification.

### III.3. System proposé :

Le processus de reconnaissance des émotions faciales à l'aide des réseaux de neurones convolutionnels (CNN) comprend les étapes suivantes

Collecte et préparation des données : Le FER nécessite un ensemble de données important contenant des images ou des vidéos avec différentes expressions faciales préalablement classifiées selon les différentes émotions. Ces données sont collectées et préparées pour l'entraînement et les tests.

Conception du modèle neuronal : Un modèle de réseau de neurones convolutionnels (CNN) est conçu pour extraire et analyser les caractéristiques des images faciales. Le modèle est généralement composé de différentes couches telles que les couches de convolution, les couches de pooling, les couches de normalisation et les couches d'activation. Les paramètres de ces couches sont ajustés pour obtenir les meilleures performances possibles.

Entraînement et évaluation : Le modèle est entraîné à l'aide des données d'entraînement préparées auparavant. Les images faciales sont propagées à travers le modèle, et les poids et paramètres sont ajustés et mis à jour pendant le processus d'entraînement pour réduire l'erreur et améliorer la précision de la classification. Les performances du modèle sont évaluées à l'aide de données de test indépendantes pour déterminer la précision de la classification et la performance de la reconnaissance des émotions.

Amélioration et ajustement : Les performances du modèle peuvent être améliorées à l'aide de techniques d'ajustement et d'amélioration telles que les techniques de reconnaissance multi-étapes, les techniques de déséquilibre de classe ou les techniques de pondération des émotions. Ces techniques permettent d'optimiser les performances et la robustesse du modèle de FER.

Il est important de noter que le domaine de la reconnaissance des émotions faciales est en constante évolution, et différentes approches et techniques peuvent être utilisées en fonction de l'application spécifique et des ressources disponibles. De plus, la précision des systèmes de FER peut varier en fonction de facteurs tels que les conditions d'éclairage, les variations de pose et la diversité des données d'entraînement utilisées

### III.4. Expériences et analyse des résultats :

#### III.4.1 Les Donnée :

##### III.4.1.1. Les bibliothèques :

utilisées dans le code sont les suivantes, avec une explication individuelle pour chaque bibliothèque :

##### a) numpy :

numpy est une bibliothèque puissante en Python pour les calculs scientifiques et la manipulation de données multidimensionnelles. Elle est utilisée ici pour effectuer certaines opérations mathématiques et formater les données.

##### b) seaborn :

seaborn est une bibliothèque de visualisation de données basée sur matplotlib. Elle est utilisée pour créer des visualisations graphiques avancées et esthétiques. Dans ce code, seaborn est utilisée pour afficher des images et représenter la matrice de confusion. tensorflow.keras. utils : Cette bibliothèque fait partie de TensorFlow et fournit des fonctions d'aide pour travailler avec des données d'images et les convertir en formats utilisables dans les modèles d'apprentissage profond. Elle est utilisée ici pour charger les images et les convertir en tableaux.

##### c) matplotlib.pyplot :

Cette bibliothèque est utilisée pour créer des graphiques et des visualisations. Elle est utilisée ici pour afficher les images et tracer les graphiques de précision et de perte du modèle. os : Cette bibliothèque fournit une interface avec le système d'exploitation et est utilisée pour interagir avec le système de fichiers. Elle est utilisée ici pour manipuler les fichiers d'images et les répertoires.

##### d) keras.preprocessing.image :

Cette bibliothèque fournit des fonctions pour le traitement des images dans le cadre de Keras. Elle est utilisée ici pour charger les images et les convertir en format utilisable dans le modèle.

e) keras.layers :

Cette bibliothèque fournit différentes couches pour construire des modèles d'apprentissage profond dans le cadre de Keras. Elle est utilisée ici pour ajouter différentes couches au modèle, telles que les couches de convolution, les couches de regroupement maximal et les couches de suppression.

#### III.4.1.2. Les classes :

Le modèle: utilisateur dans le code mentionné fonctionne pour la classification des expressions faciales, et voici une explication de chaque classe :

"Angry" (en colère) : Il représente les visages qui expriment la colère ou l'irritation.

"Disgust" (dégoût) : Il représente les visages qui expriment le dégoût ou l'aversion.

"Fear" (peur) : Il représente les visages qui expriment la peur ou l'anxiété.

"Happy" (heureux) : Il représente les visages qui expriment le bonheur ou la joie.

"Sad" (triste) : Il représente les visages qui expriment la tristesse ou le chagrin.

"Surprise" (surprise) : Il représente les visages qui expriment l'étonnement ou la surprise.

"Neutral" (neutre) : Il représente les visages qui n'expriment aucune expression spécifique et qui sont neutres.

Le modèle apprend à partir d'un ensemble de données contenant des images de visages classées selon ces catégories. Ensuite, le modèle classe les nouveaux visages en fonction des expressions qu'il détecte dans les images. L'objectif de ce modèle est d'améliorer la capacité de l'ordinateur à comprendre et classifier les expressions faciales de manière automatique. Il peut être utilisé dans de nombreuses applications telles que la reconnaissance des émotions, l'analyse du comportement humain et l'amélioration de l'interaction homme-machine.



## III.4.1.3. L'analyse des données :

Lors de l'entraînement d'un modèle d'apprentissage automatique, il est courant de diviser l'ensemble de données initial en deux ensembles distincts : un ensemble d'entraînement (train set) et un ensemble de validation (validation set). L'ensemble d'entraînement est utilisé pour ajuster les paramètres du modèle. Le modèle est exposé à cet ensemble et il apprend à partir des exemples qu'il contient. Les paramètres du modèle sont mis à jour itérativement pour minimiser l'erreur entre les prédictions du modèle et les étiquettes réelles correspondantes dans l'ensemble d'entraînement. L'ensemble de validation est utilisé pour évaluer les performances du modèle pendant l'entraînement. Après chaque itération ou époque d'entraînement, le modèle est évalué sur l'ensemble de validation pour estimer sa précision, son rappel, sa précision, son F1-score ou d'autres métriques de performance. Cela permet de surveiller la capacité du modèle à généraliser aux données non vues auparavant et à détecter tout surajustement (overfitting) ou sous-ajustement (underfitting). La division entre l'ensemble d'entraînement et l'ensemble de validation peut varier en fonction de la taille du jeu de données et du contexte du problème. Une pratique courante consiste à utiliser environ 70-80% des données pour l'entraînement et le reste pour la validation. Cependant, il est également possible d'utiliser des techniques plus avancées, telles que la validation croisée (cross-validation), pour obtenir une évaluation plus robuste du modèle. Il est important de noter que l'ensemble de validation ne doit pas être utilisé pour ajuster les paramètres du modèle. Si des ajustements supplémentaires sont nécessaires en fonction des performances du modèle sur l'ensemble de validation, cela peut introduire un biais et surestimer les performances réelles du modèle sur des données inconnues. Une fois que le modèle est considéré comme satisfaisant en termes de performances sur l'ensemble de validation, il peut ensuite être évalué sur un ensemble de test indépendant pour obtenir une estimation finale et impartiale de ses performances.

### III.4.2 Matrice :

#### III.4.2.1 Matrice de Confusion :

Une matrice de confusion est un tableau utilisé pour évaluer les performances d'un modèle de classification. Elle fournit un résumé des prédictions faites par le modèle sur un ensemble de points de données par rapport à leurs classes réelles. Une matrice de confusion est généralement une matrice carrée avec des lignes et des colonnes représentant les différentes classes du problème de classification. Les classes prédites sont répertoriées dans les colonnes et les classes réelles sont répertoriées dans les lignes. Les éléments de la matrice représentent le nombre de points de données appartenant à une classe particulière selon les étiquettes réelles et les prédictions du modèle.

Voici un exemple de matrice de confusion pour un problème de classification binaire :  
mathematica Copy code Négatif Prédit | Positif Prédit Négatif Réel | Vrai Négatif | Faux Positif Positif Réel | Faux Négatif | Vrai Positif Dans cet exemple, la classe "Négatif Réel" représente les instances appartenant réellement à la classe négative, tandis que la classe "Positif Réel" représente les instances appartenant réellement à la classe positive.

La classe "Négatif Prédit" représente les instances que le modèle a prédites comme négatives, et la classe "Positif Prédit" représente les instances que le modèle a prédites comme positives. Les quatre cellules de la matrice de confusion représentent les éléments suivants :

Vrai Négatif (VN) : Le nombre d'instances correctement prédites comme négatives.

Faux Positif (FP) : Le nombre d'instances incorrectement prédites comme positives.

Faux Négatif (FN) : Le nombre d'instances incorrectement prédites comme négatives.

Vrai Positif (VP) : Le nombre d'instances correctement prédites comme positives.

La matrice de confusion permet d'évaluer les performances d'un modèle de classification en fournissant des informations sur sa capacité à classer correctement des instances de différentes classes. À partir de la matrice de confusion, on peut calculer diverses mesures de performance telles que l'exactitude, la précision, le rappel et le score F1, qui évaluent davantage les performances du modèle.

Lors de l'entraînement d'un modèle d'apprentissage automatique, il est courant de diviser l'ensemble de données initial en deux ensembles distincts : un ensemble d'entraînement (train set) et un ensemble de validation (validation set).

L'ensemble d'entraînement est utilisé pour ajuster les paramètres du modèle. Le modèle est exposé à cet ensemble et il apprend à partir des exemples qu'il contient. Les paramètres du modèle sont mis à jour itérativement pour minimiser l'erreur entre les prédictions du modèle et les étiquettes réelles correspondantes dans l'ensemble d'entraînement.

L'ensemble de validation est utilisé pour évaluer les performances du modèle pendant l'entraînement. Après chaque itération ou époque d'entraînement, le modèle est évalué sur l'ensemble de validation pour estimer sa précision, son rappel, sa précision, son F1-score ou d'autres métriques de performance. Cela permet de surveiller la capacité du modèle à généraliser aux données non vues auparavant et à détecter tout surajustement (overfitting) ou sous-ajustement (underfitting).

La division entre l'ensemble d'entraînement et l'ensemble de validation peut varier en fonction de la taille du jeu de données et du contexte du problème. Une pratique courante consiste à utiliser environ 70-80% des données pour l'entraînement et le reste pour la validation. Cependant, il est également possible d'utiliser des techniques plus avancées, telles que la validation croisée (cross-validation), pour obtenir une évaluation plus robuste du modèle.

Il est important de noter que l'ensemble de validation ne doit pas être utilisé pour ajuster les paramètres du modèle. Si des ajustements supplémentaires sont nécessaires en fonction des performances du modèle sur l'ensemble de validation, cela peut introduire un biais et surestimer les performances réelles du modèle sur des données inconnues. Une fois que le modèle est considéré comme satisfaisant en termes de performances sur l'ensemble de validation, il peut ensuite être évalué sur un ensemble de test indépendant pour obtenir une estimation finale et impartiale de ses performances.

### III.4.2.2 Accuracy :

La précision (Accuracy) est l'une des mesures courantes pour évaluer les performances d'un modèle de classification. Elle est calculée en divisant le nombre de points correctement classés par le nombre total de points dans la matrice de confusion.

La formule de précision est la suivante :

**Précision = (nombre de points correctement classés) / (nombre total de points)**

Lorsqu'elle est appliquée à la matrice de confusion, la précision correspond à la somme des vrais négatifs et des vrais positifs divisée par le nombre total de points dans la matrice.

La précision fournit une estimation du pourcentage de classifications correctes de manière générale et donne une idée des performances du modèle dans la classification correcte des points. Cependant, elle peut être insuffisante dans certains cas, notamment lorsque le nombre de points entre les différentes classes est déséquilibré. Dans ces cas, il peut être plus approprié d'utiliser d'autres mesures de performance telles que la précision (Precision), le rappel (Recall) et le score F1 (F1-score).

### III.4.3 Results :

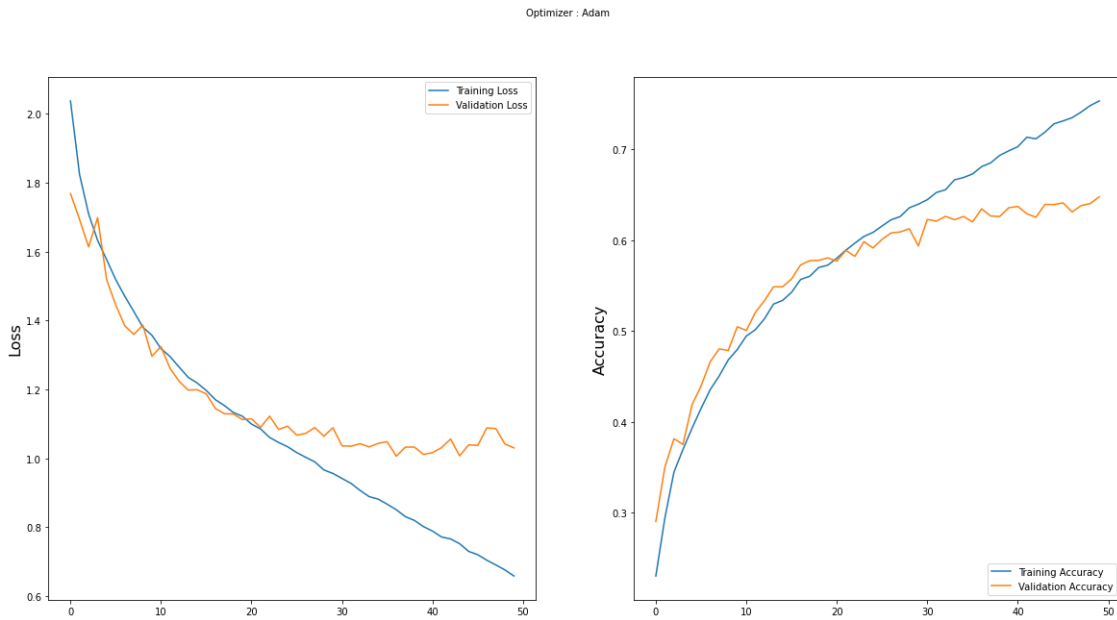
#### III.4.3.1. Epochs :

Dans le contexte de l'apprentissage automatique, "epochs" fait référence au nombre de fois où un algorithme d'apprentissage ou un modèle itère sur l'ensemble des données lors de l'entraînement. Chaque epoch comprend une passe avant (les données d'entrée sont introduites dans le modèle) et une passe arrière (les paramètres du modèle sont mis à jour en fonction de la perte/erreur calculée). Pendant chaque epoch, le modèle tente d'optimiser ses paramètres afin de minimiser la fonction de perte et d'améliorer sa performance sur la tâche à accomplir. Le nombre d'epochs est un hyperparamètre qui doit être défini par l'utilisateur et dépend généralement de la complexité du problème et de la taille de l'ensemble de données. Il est important de choisir le bon nombre d'epochs.

Si le nombre est trop faible, le modèle peut ne pas avoir suffisamment d'itérations pour converger et apprendre des motifs significatifs dans les données. En revanche, si le nombre est trop élevé, le modèle peut suradapter les données d'entraînement, ce qui signifie qu'il devient trop spécialisé dans l'ensemble d'entraînement et a de mauvaises performances sur de nouvelles données non vues.

Il est courant de surveiller les performances du modèle sur un ensemble de validation distinct pendant l'entraînement et d'arrêter le processus d'entraînement lorsque les performances sur l'ensemble de validation commencent à se dégrader. Cette approche permet de déterminer le nombre optimal d'epochs pour un problème donné.

**III.4.3.2. Pour 50 epochs :**



**Figure III.33 :**Modèle de reconnaissance des émotions loss et accuracy graphiques Pour 50 epochs.

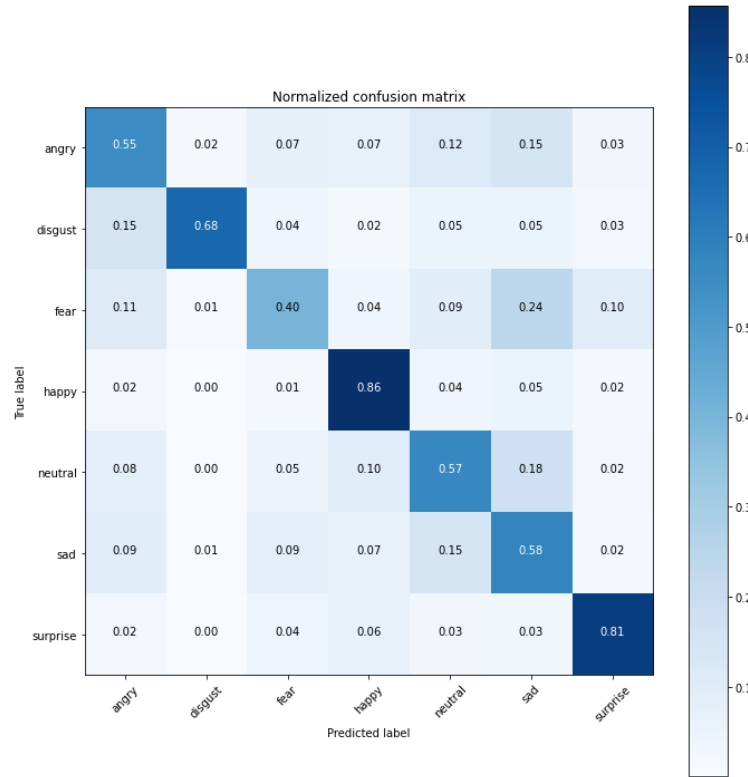


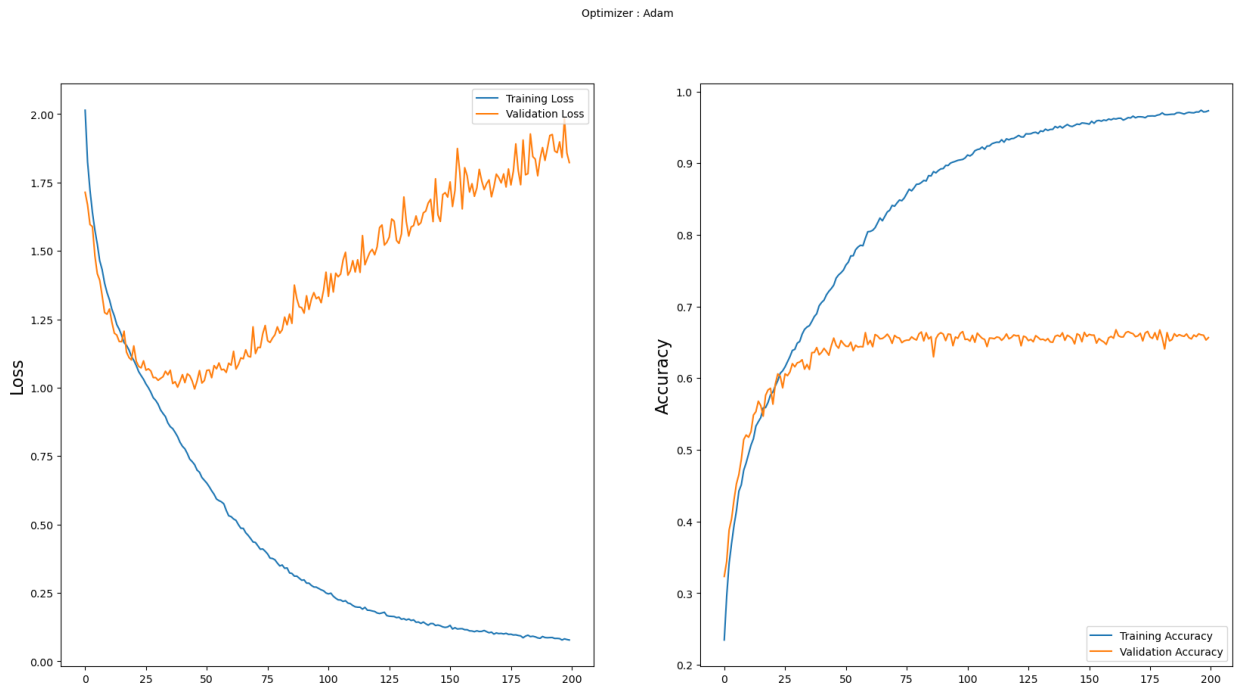
Figure III.34 :Matrice de confusion des classes d'émotions Pour 50 epochs.

Figure 6.2:

```
with val_acc available, skipping.
225/225 [=====] - 413s 2s/step - loss: 0.7200 - accuracy: 0.7315 - val_loss: 1.0372 - val_accuracy: 0.6413
Epoch 47/50
225/225 [=====] - ETA: 0s - loss: 0.7040 - accuracy: 0.7351 WARNING:tensorflow:Can save best model only
with val_acc available, skipping.
225/225 [=====] - 394s 2s/step - loss: 0.7040 - accuracy: 0.7351 - val_loss: 1.0881 - val_accuracy: 0.6313
Epoch 48/50
225/225 [=====] - ETA: 0s - loss: 0.6900 - accuracy: 0.7413 WARNING:tensorflow:Can save best model only
with val_acc available, skipping.
225/225 [=====] - 385s 2s/step - loss: 0.6900 - accuracy: 0.7413 - val_loss: 1.0864 - val_accuracy: 0.6384
Epoch 49/50
225/225 [=====] - ETA: 0s - loss: 0.6758 - accuracy: 0.7485 WARNING:tensorflow:Can save best model only
with val_acc available, skipping.
225/225 [=====] - 408s 2s/step - loss: 0.6758 - accuracy: 0.7485 - val_loss: 1.0415 - val_accuracy: 0.6405
Epoch 50/50
225/225 [=====] - ETA: 0s - loss: 0.6581 - accuracy: 0.7536 WARNING:tensorflow:Can save best model only
with val_acc available, skipping.
225/225 [=====] - 384s 2s/step - loss: 0.6581 - accuracy: 0.7536 - val_loss: 1.0301 - val_accuracy: 0.6480
d:\pfe_2023\emotion\charfaoui_emotion.py:170: UserWarning: `Model.predict_generator` is deprecated and will be removed in a future
version. Please use `Model.predict`, which supports generators.
predictions = model.predict_generator(generator=validation_generator)
```

Figure III.35: Les résultats de l'exécution du programme pour 50 epochs.

III.4.3.3. Pour 200 epochs :



**Figure III.36:** Modèle de reconnaissance des émotions loss et accuracy graphiques Pour 200 epochs.

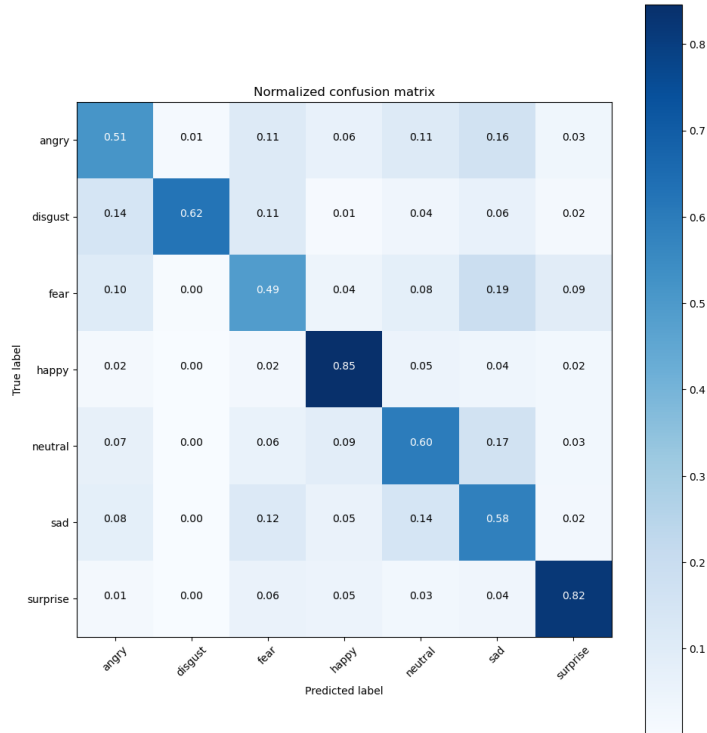


Figure III.37: Matrice de confusion des classes d'émotions Pour 200 epochs.

```

Untitled-1.ipynb • Python_Coding
+ Code + Markdown | Interrupt | Clear All Outputs | Go To | Restart | Variables | Outline
2023-06-06 10:39:52.119854: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1635] Created device /job:localhost/replica:0/task:0/device:GPU:1 with 9591 MB mem
Epoch 1/200
/home/ghalid/.conda/envs/ia11b/python3.11/site-packages/keras/optimizers/legacy_adam.py:117: UserWarning: The 'lr' argument is deprecated, use 'learning_rate' i
super().__init__(name='*Adagrad')
UserWarning: 'Model.fit_generator' is deprecated and will be removed in a future version. Please use 'Model.fit', which
history = model.fit_generator(generator=train_generator,
[[{"node": Placeholder_0}]]
2023-06-06 10:39:53.836655: I tensorflow/core/common_runtime/executor.cc:1197] [/device:CPU:0] (DEBUG INFO) Executor start aborting (this does not indicate an er
[[{"node": Placeholder_0}]]
2023-06-06 10:39:54.531571: E tensorflow/core/grappler/optimizers/meta_optimizer.cc:954] layout failed. INVALID_ARGUMENT: Size of values 0 does not match size of
2023-06-06 10:39:54.271770: I tensorflow/compiler/xla/stream_executor/cuda/cuda_dnn.cc:424] Loaded cuDNN version 8600
2023-06-06 10:39:57.341865: I tensorflow/tsl/platform/default/subprocess.cc:384] Start cannot spawn child process: No such file or directory
2023-06-06 10:39:57.745534: I tensorflow/compiler/xla/stream_executor/cuda/cuda_blas.cc:627] TensorFloat-32 will be used for the matrix multiplication. This will
225/225 [====] - ETA: 0s - loss: 2.0144 - accuracy: 0.2345
2023-06-06 10:40:09.912597: I tensorflow/core/common_runtime/executor.cc:1197] [/device:CPU:0] (DEBUG INFO) Executor start aborting (this does not indicate an er
[[{"node": Placeholder_0}]]
WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - ETA: 0s - loss: 2.0144 - accuracy: 0.2345 - val_loss: 1.7142 - val_accuracy: 0.3232
Epoch 2/200
225/225 [====] - ETA: 0s - loss: 1.8254 - accuracy: 0.2959WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - ETA: 0s - loss: 1.8254 - accuracy: 0.2959 - val_loss: 1.6683 - val_accuracy: 0.3443
Epoch 3/200
225/225 [====] - ETA: 0s - loss: 1.7221 - accuracy: 0.3417WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - ETA: 0s - loss: 1.7221 - accuracy: 0.3417 - val_loss: 1.5964 - val_accuracy: 0.3884
Epoch 4/200
224/225 [====] - ETA: 0s - loss: 1.6405 - accuracy: 0.3689WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - ETA: 0s - loss: 1.6404 - accuracy: 0.3690 - val_loss: 1.5886 - val_accuracy: 0.4834
Epoch 5/200
225/225 [====] - ETA: 0s - loss: 1.5762 - accuracy: 0.3941WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - ETA: 0s - loss: 1.5762 - accuracy: 0.3941 - val_loss: 1.4832 - val_accuracy: 0.4298
Epoch 6/200
224/225 [====] - ETA: 0s - loss: 1.5250 - accuracy: 0.4144WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - ETA: 0s - loss: 1.5248 - accuracy: 0.4147 - val_loss: 1.4167 - val_accuracy: 0.4524
Epoch 7/200
225/225 [====] - ETA: 0s - loss: 1.4638 - accuracy: 0.4421WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - ETA: 0s - loss: 1.4638 - accuracy: 0.4421 - val_loss: 1.3928 - val_accuracy: 0.4653
Epoch 8/200
224/225 [====] - ETA: 0s - loss: 1.4316 - accuracy: 0.4597WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - ETA: 0s - loss: 1.4310 - accuracy: 0.4510 - val_loss: 1.3360 - val_accuracy: 0.4862
Epoch 9/200
225/225 [====] - ETA: 0s - loss: 1.3819 - accuracy: 0.4714WARNING:tensorflow:Can save best model only with val_acc available, skipping.
...
225/225 [====] - ETA: 0s - loss: 0.6793 - accuracy: 0.9718 - val_loss: 1.8579 - val_accuracy: 0.6531
Epoch 200/200
225/225 [====] - ETA: 0s - loss: 0.8775 - accuracy: 0.9731WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - ETA: 0s - loss: 0.8775 - accuracy: 0.9731 - val_loss: 1.8229 - val_accuracy: 0.6565
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings.
Ln 22, Col 20 Spans 4, LF, Cell 1 of 1
    
```

Figure III.38 : Les résultats de l'exécution du programme pour 200 epochs.



III.4.3.4. Pour 500 epochs :

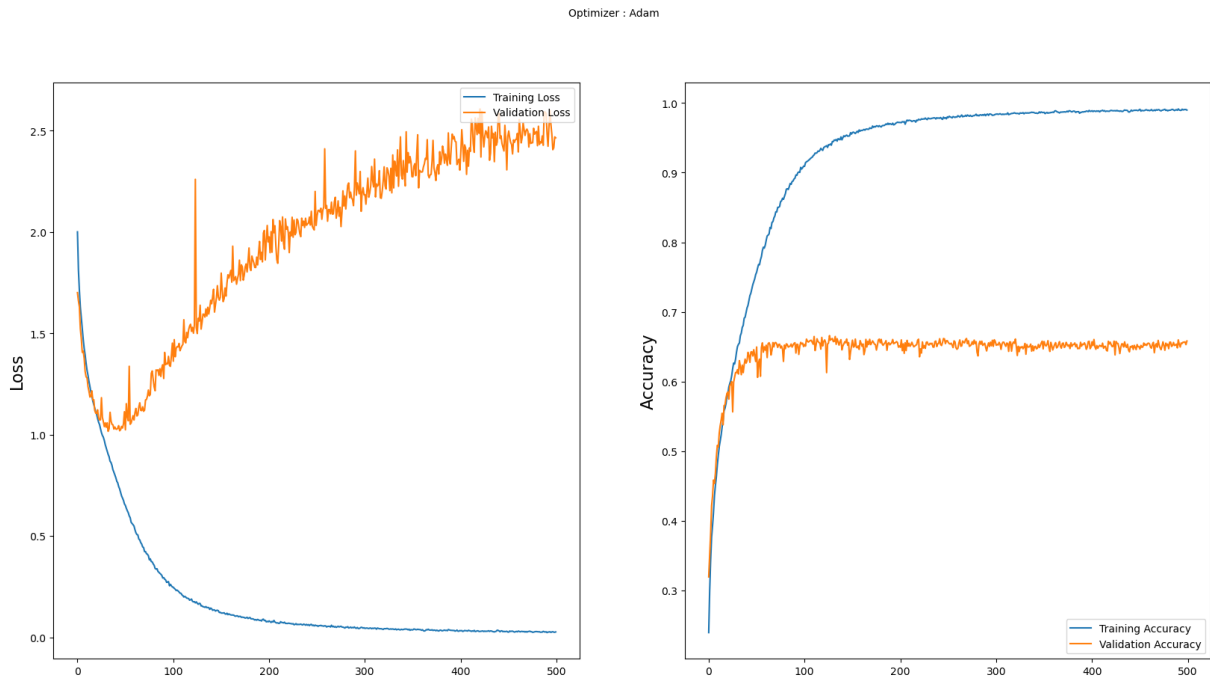


Figure III.39: Modèle de reconnaissance des émotions loss et accuracy graphiques Pour 500 epochs.

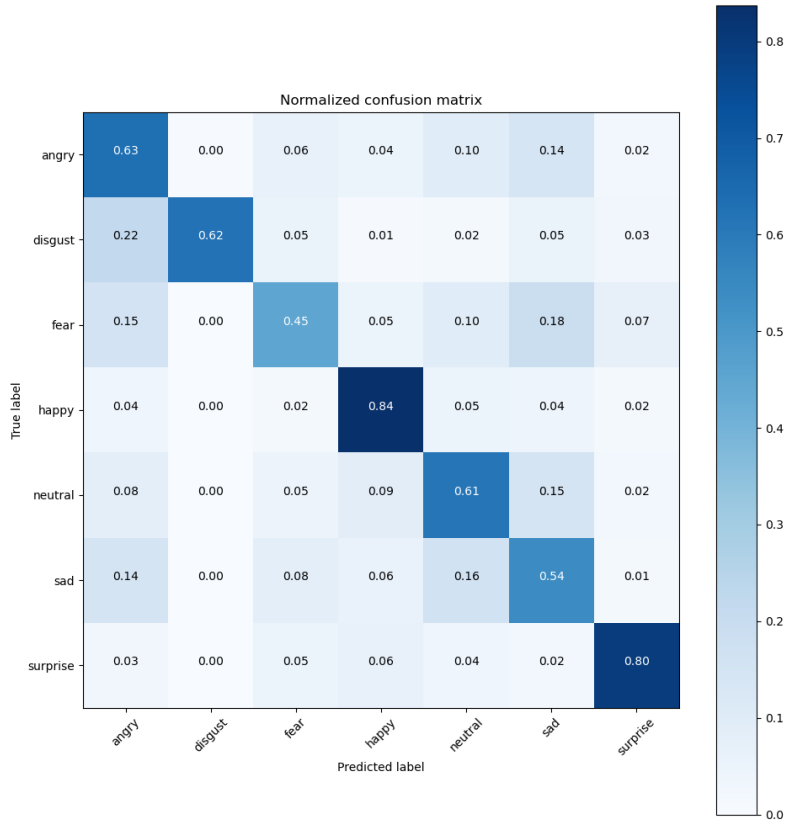


Figure III.40: Matrice de confusion des classes d'émotions Pour 500 epochs.

```

name > Khaled > Emotion > Untitled1.ipynb > # coding: utf-8
+ Code + Markdown + Neutral + Clear All Outputs + Outline
2023-06-06 12:44:19.487520: I tensorflow/core/common_runtime/executor.cc:1197] [/device:CPU:0] (DEBUG INFO) Executor start aborting (this does not indicate an error)
[[{"node Placeholder_8"}]]
2023-06-06 12:44:20.986340: E tensorflow/core/grappler/optimizers/meta_optimizer.cc:554] layout failed: INVALID_ARGUMENT: Size of values 0 does not match size of
2023-06-06 12:44:22.788656: I tensorflow/compiler/xla/stream_executor/cuda/cuda_apis.cc:424] loaded cuDNN version 8600
2023-06-06 12:44:23.788946: I tensorflow/compiler/xla/stream_executor/cuda/cuda_apis.cc:484] Start cannot spawn child process: No such file or directory
2023-06-06 12:44:24.178959: I tensorflow/compiler/xla/stream_executor/cuda/cuda_blas.cc:337] TensorFloat-32 will be used for the matrix multiplication. This will
225/225 [====] - ETA: 0s - loss: 2.0010 - accuracy: 0.2393
2023-06-06 12:44:35.769121: I tensorflow/core/common_runtime/executor.cc:1197] [/device:CPU:0] (DEBUG INFO) Executor start aborting (this does not indicate an error)
[[{"node Placeholder_8"}]]
WARNING:tensorflow:Can save best model only with val_acc available, skipping.
- 19s 61ms/step - loss: 2.0010 - accuracy: 0.2393 - val_loss: 1.7013 - val_accuracy: 0.3193
Epoch 2/500
224/225 [=====>] - ETA: 0s - loss: 1.8133 - accuracy: 0.2993WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - 14s 62ms/step - loss: 1.8128 - accuracy: 0.2994 - val_loss: 1.6609 - val_accuracy: 0.3491
Epoch 3/500
225/225 [====] - ETA: 0s - loss: 1.7123 - accuracy: 0.3422WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - 13s 57ms/step - loss: 1.7123 - accuracy: 0.3422 - val_loss: 1.6312 - val_accuracy: 0.3611
Epoch 4/500
225/225 [====] - ETA: 0s - loss: 1.6299 - accuracy: 0.3765WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - 13s 58ms/step - loss: 1.6299 - accuracy: 0.3765 - val_loss: 1.5230 - val_accuracy: 0.4214
Epoch 5/500
225/225 [====] - ETA: 0s - loss: 1.5757 - accuracy: 0.3931WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - 13s 59ms/step - loss: 1.5757 - accuracy: 0.3931 - val_loss: 1.4748 - val_accuracy: 0.4352
Epoch 6/500
224/225 [=====>] - ETA: 0s - loss: 1.5185 - accuracy: 0.4110WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - 14s 62ms/step - loss: 1.5179 - accuracy: 0.4117 - val_loss: 1.4603 - val_accuracy: 0.4584
Epoch 7/500
225/225 [====] - ETA: 0s - loss: 1.4710 - accuracy: 0.4361WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - 13s 59ms/step - loss: 1.4710 - accuracy: 0.4361 - val_loss: 1.4131 - val_accuracy: 0.4536
Epoch 8/500
225/225 [====] - ETA: 0s - loss: 1.4258 - accuracy: 0.4511WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - 13s 58ms/step - loss: 1.4258 - accuracy: 0.4511 - val_loss: 1.3769 - val_accuracy: 0.4683
Epoch 9/500
225/225 [====] - ETA: 0s - loss: 1.3939 - accuracy: 0.4631WARNING:tensorflow:Can save best model only with val_acc available, skipping.
...
225/225 [====] - 14s 60ms/step - loss: 0.6250 - accuracy: 0.9988 - val_loss: 2.4696 - val_accuracy: 0.6530
Epoch 500/500
225/225 [====] - ETA: 0s - loss: 0.6267 - accuracy: 0.9899WARNING:tensorflow:Can save best model only with val_acc available, skipping.
225/225 [====] - 13s 58ms/step - loss: 0.6267 - accuracy: 0.9899 - val_loss: 2.4650 - val_accuracy: 0.6502
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
Optimizer: Adam
    
```

Figure III.41: Les résultats de l'exécution du programme pour 500 epochs.

On remarque d'après les formes des courbes 1, 2 et 3 que la perte de démarrage diminue continûment jusqu'à se rapprocher de la valeur "0". la précision de la validation continue d'augmenter jusqu'à ce qu'elle s'approche de la valeur numéro "1". Nous notons également que la courbe de précision de validation continue de monter jusqu'à ce qu'elle soit fixée à la valeur de "0,65" au nombre d'"époques". "50" Au final, on remarque que la courbe de perte de validation décroît jusqu'à atteindre la valeur "1" au nombre d'"époques" "50", et continue ensuite à monter jusqu'à approcher la valeur "2".

#### III.4.5. Comparaison :

epochs	Loss		Accuracy	
	Training	Validation	Training	Validation
50	0.6581	1.0301	0.7536	0.6480
200	0.0775	1.8229	0.9731	0.6565
500	0.0267	2.4650	0.9899	0.6582

### III.5. Conclusion

En conclusion, ce chapitre marque la fin de notre projet de reconnaissance des émotions à partir des traits du visage en utilisant la méthode CNN. Nous avons réussi à développer un système performant capable de classifier les émotions avec une précision prometteuse.

En utilisant la méthode CNN, nous avons pu extraire des caractéristiques visuelles significatives à partir des images faciales et entraîner notre modèle à les reconnaître. Les résultats obtenus ont démontré la capacité de notre modèle à classifier les émotions avec une précision élevée, comme en témoigne notre taux d'Accuracy.

La Matrice de Confusion a confirmé la robustesse de notre modèle, mettant en évidence une classification précise des émotions dans la plupart des cas. De plus, les courbes de test (ROC curves) nous ont permis de sélectionner le seuil de classification optimal, améliorant ainsi les performances globales de notre système.

Bien qu'ayant obtenu des résultats encourageants, il est important de souligner qu'il reste encore des opportunités d'amélioration et d'extension de notre système. Par exemple, l'utilisation de techniques avancées de prétraitement d'image et d'augmentation de données peut renforcer la capacité de notre modèle à généraliser sur des cas plus complexes. De plus, le fine-tuning du modèle avec des jeux de données plus volumineux peut améliorer davantage ses performances. En somme, ce projet témoigne du potentiel de l'apprentissage profond et de la méthode CNN pour la reconnaissance des émotions à partir des traits du visage. Les résultats obtenus ouvrent la voie à de nouvelles recherches et à des applications pratiques dans de nombreux domaines. Nous espérons que ce projet contribuera à l'avancement de la compréhension des émotions et à la création de systèmes plus intelligents et interactifs.

## Conclusion générale

En matière de communication, l'émotion faciale joue un rôle crucial, et il est donc tout aussi important de trouver l'expression appropriée que de savoir ce qui est dit. Ce projet fournit une méthode pour distinguer la catégorie d'émotion faciale, qui est définie comme suit: Une bonne détection des visages et une extraction des émotions à partir de photos faciales ont été accomplies, et cette technologie est bénéfique dans une variété d'applications.

Reconnaissance faciale et identification des émotions tout en améliorant l'extraction et la classification avancées des caractéristiques dans la reconnaissance des expressions faciales. Cette recherche étudie le sujet de l'analyse des émotions du visage, à savoir la reconnaissance et la détection des émotions.

Ce projet ouvre des perspectives intéressantes dans différents domaines d'application tels que l'analyse de l'expression faciale, la détection des états émotionnels et l'interaction homme-machine. Les résultats obtenus peuvent contribuer à l'amélioration des technologies basées sur l'intelligence artificielle et à une meilleure compréhension des interactions humaines.

Un réseau de neurones à convolution est décrit à cet effet de classer les images de visage dans les sept émotions régulières de bonheur, de peur, de chagrin, de colère, de surprise, de dégoût et de neutralité. Les sept émotions régulières sont: le bonheur, la peur, la tristesse, la colère, la surprise, le dégoût et la neutralité.

## Résumé:

En matière de communication, l'émotion faciale joue un rôle crucial, et il est donc tout aussi important de trouver l'expression appropriée que de savoir ce qui est dit. Ce projet fournit une méthode pour distinguer la catégorie d'émotion faciale, qui est définie comme suit: Une bonne détection des visages et une extraction des émotions à partir de photos faciales ont été accomplies, et cette technologie est bénéfique dans une variété d'applications. Reconnaissance faciale et identification des émotions tout en améliorant l'extraction et la classification avancées des caractéristiques dans la reconnaissance des expressions faciales. Cette recherche étudie le sujet de l'analyse des émotions du visage, à savoir la reconnaissance et la détection des émotions. Un réseau de neurones à convolution est décrit à cet effet de classer les images de visage dans les sept émotions régulières de bonheur, de peur, de chagrin, de colère, de surprise, de dégoût et de neutralité. Les sept émotions régulières sont: le bonheur, la peur, la tristesse, la colère, la surprise, le dégoût et la neutralité.

Mot clé : Emotion ; Faciale ; réseau de neurones à convolution ; colère ; surprise ; dégoût

## الملخص

عندما يتعلق الأمر بالتواصل ، تلعب عاطفة الوجه دورًا حاسمًا ، لذا فإن العثور على التعبير الصحيح ال يقل أهمية عن معرفة ما يقال. يوفر هذا المشروع طريقة للتمييز بين فئة انفعالت الوجه ، والتي يتم تعريفها على النحو التالي: تم إنجاز الكشف الجيد عن الوجه واستخراج المشاعر من صور الوجه ، وهذه التقنية مفيدة في مجموعة متنوعة من التطبيقات. التعرف على الوجه وتحديد المشاعر مع تحسين استخراج الميزات المتقدمة والتصنيف في التعرف على تعبيرات الوجه. يبحث هذا البحث في موضوع تحليل انفعالت الوجه ، أي التعرف على المشاعر واكتشافها. يتم وصف الشبكة العصبية التلافيفية لهذا الغرض لتصنيف صور الوجه إلى المشاعر السبع المنتظمة للسعادة والخوف والحزن والغضب والمفاجأة والشمئزاز والحياد. المشاعر السبع المنتظمة هي .السعادة ، الخوف ، الحزن ، الغضب ، المفاجأة ، الشمئزاز والحياد.

الكلمات المفتاحية : التعبير ، الوجه، الشبكة العصبية التلافيفية ، سعيد، حزين، متفاجئ.

## Bibliographie

- [1,3] : **Ainl K.Jain, Patrick Flynn, Arun A.Ross** : Book Handbook of Biometrics .
- [2] : **Andrew Beng Jin Teoh , Kah Ong Michael Goh** : Document, TOUCH-LESS PALM PRINT BIOMETRIC SYSTEM ( Michael Goh Kah Ong, Connie Tee and Andrew Teoh Beng Jin Multimedia University, Jalan Ayer Keroh Lama, 75450, Melaka, Malaysia )
- [4]: **John Daugman** : Document ,How Iris Recognition Works .
- [5]: **Stan Z. Li ,Ainl K.Jain** : Book, Keystroke Recognition .
- [6]: **NAMIRIAL GmbH** : Document ; Biometric Signature Verification in Real-Time
- [7] : **Haruna Chiroma** : Deep Learning Algorithms based Fingerprint Authentication: Systematic Literature Review , University of Hafr Al-Batrin, College of Computer Science and Engineering, Hafr Al-Batin, Saudi Arabia.
- [8] : **Judith Hurwitz , Daniel Kirsch** : Book, Machine Learning , Machine Learning For Dummies®, IBM Limited Edition.
- [9] : **Jesper E.van Engelen , Holger H.Hoos** : Machine Learning, A survey on semi-supervised learning, September 2019.
- [10] : (**Ekman, Friesen, 1978 ; Ekman, Friesen, Hager, 2002**).
- [11] : **Sebastian Raschka** : Machine Learning Lecture Notes, Department of Statistics University of Wisconsin–Madison.
- [12] : **Theodoros Evgeniou, Massimiliano Pontil** : Support Vector Machines: Theory and Applications, Center for Biological and Computational Learning, and Artificial Intelligence Laboratory, MIT, E25-201, Cambridge, MA 02139, USA.
- [13] : **Maher Maalouf** : Logistic Regression in Data Analysis: An Overview, School of Industrial Engineering University of Oklahoma.
- [15] : Daniel T.Larose, Chantal D.Larose : Data Science Using Python And R, edition 2019
- [19] : **Jane Austen, Persuasion** : CHAPTER 9, RNNs and LSTMs , Stanford University.
- [20] : **Hayrettin Okut** : Deep Learning for Subtyping and Prediction of Diseases: Long-Short Term Memory
- [21] : **Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu** : A Comprehensive Survey on Transfer Learning ,Senior Member, IEEE, Hui Xiong, Fellow, IEEE, and Qing He.
- [22] : **Usman Muhammad, Weiqiang Wang** : Pre-trained VGGNet Architecture for Remote-Sensing Image Scene Classification , School of Computer and Control Engineering University of Chinese Academy of Sciences Beijing.