

République Algérienne Démocratique et Populaire
Ministère de l'enseignement Supérieur et de la Recherche Scientifique
Université de Mohamed El Bachir El Ibrahimi de Bordj Bou Arreridj
Faculté des Mathématiques et d'Informatique
Département d'informatique



MEMOIRE

Présenté en vue de l'obtention du diplôme

Master en informatique

Spécialité : Réseau et multimédia

THEME

Systeme de détection de visage, d'émotion, de sexe et
classification d'âge en Deep Learning.

Présenté par :

Aggoune Manel.

Soutenu publiquement le : 13/09/2023.

Devant le jury composé de:

Président : Ms Charikhi Mourad.

Examineur : Mme Belalta Ramla.

Encadrant : Mme Saad Saoud Manel.

2022/2023

Remerciements :

En préambule à ce mémoire je remercie **ALLAH** qui m'a aidé et m'a donné la patience et le courage durant ces longues années d'études et d'accomplir ce modeste travail.

Je tiendrai d'abord à remercier très chaleureusement :

- **Mme Saad Saoud Manel** qui m'a permis de bénéficier de son encadrement, ses conseils et son orientation.
- **Les membres du jury** qui me fera l'honneur d'estimer et de juger ce travail.

Ensuite, J'adresse mes remerciements les plus sincères à mes **très chers PARENTS** qui ont apportés leurs aides durant ces longues années.

Merci à **tous** ceux qui m'ont apportée d'aide de près ou de loin.

Dédicace :

Premièrement et avant tout « الحمد لله على تمام النعمة »

Je tiens à dédier ce modeste travail à tous ceux qui m'ont encouragée durant toute la période de réalisation de ce travail.

En particulièrement :

- A mes chers parents qui se sacrifient pour me voir réussir.
- A mon très cher mari **Kennedy Bilal** qui a toujours été à mes côtés.
- A mes chers frères **Akram, Mohamed** et **Youcef**.
- A toute ma famille paternelle et maternelle (**Aggoune** et **Aissaoui**).
- A ma chère belle-famille **Kosse** un par un.
- A ma chère sœur du cœur **LADA**.
- A ma deuxième maman **Bouchra**.
- A ma chère copine **Sonia**.
- A Mes collègues : **Idir Sonia , Boussoualim Wiam**.

✚ **Aggoune Manel.**

Résumé :

La reconnaissance de formes et la classification automatique sont des domaines de recherche très actifs, dont l'objectif principal est de développer des systèmes intelligents capables d'apprendre et de reconnaître des modèles objet. Une partie importante de ces applications concerne la biométrie, souvent utilisée à des fins de sécurité. La reconnaissance faciale est devenue dominante en tant que technologie biométrique de base de plus en plus importante dans le domaine de la recherche.

L'objectif de ce travail est de développer un système de détection de visage, d'expressions faciale et de classification d'âge et de sexe en Deep Learning basé sur les réseaux de neurones convolutifs (CNN) qui s'appuie sur une image, une vidéo ou en temps réel.

Mots clés : Reconnaissance faciale; détection de visage; expression faciale ; classification d'âge et de sexe ; Deep Learning; réseaux de neurones convolutifs (CNN).

Abstract :

Pattern recognition and automatic classification are very active areas of research, the main objective of which is to develop intelligent systems capable of learning and recognizing object models. An important part of these applications concerns biometrics, often used for security purposes. Facial recognition has become dominant as an increasingly important core biometric technology in research.

The objective of this work is to develop a system for face detection, facial expressions and age and gender classification in Deep Learning based on convolutional neural networks (CNN) which is based on an image, video or in real time.

Keywords: Facial recognition; face detection; facial expression; age and sex classification; Deep Learning; convolutional neural networks (CNN).

المخلص :

أنظمة تطوير هو منها الرئيسي والهدف، للغاية النشطة البحث مجالات من التلقائي والتصنيف الأشكال على التعرف إن غالبًا والتي، الحيوية بالقياسات التطبيقات هذه من مهم جزء ويتعلق. الكائنات نماذج على والتعرف التعلم على قدرة ذكية مجال في متزايدة أهمية ذات أساسية حيوية كتنقية الوجه على التعرف طريقة أصبحت وقد . أمنية لأغراض تستخدم ما ..البحث

استنادًا العميق التعلم في والجنس العمر وتصنيف الوجه وتعبيرات الوجه عن للكشف نظام تطوير هو العمل هذا من الهدف. الفعلي الوقت في أو فيديو مقطع أو صورة على تعتمد التي التلافيفية العصبية الشبكات إلى

الشبكات عميق؛ تعلم والجنس؛ العمر تصنيف الوجه؛ تعابير الوجه؛ عن الكشف الوجه؛ على التعرف: **المفتاحية الكلمات**
التلافيفية العصبية

Liste de figures :

Figure 1.1 Détection de visage	14
Figure 1.2 Analyse du visage.....	14
Figure 1.3 Reconnaissance faciale.	15
Figure 1.4 Exemple de variation d'éclairage.	16
Figure 1.5 Exemple de variation de poses	16
Figure 1.6 Exemple de variation d'expressions faciales.	17
Figure 1.7 Exemple de présence ou absence des composants structurels.....	17
Figure 2.1 La relation entre IA, ML et DL.....	22
Figure 2.2 Neurone biologique et neurone artificiel	23
Figure 2.3 Structure d'un réseau de neurones artificiels.....	24
Figure 2.4 Fonctionnement d'un réseau de neurones artificiels	25
Figure 2.5 Couche de convolution	27
Figure 2.6 La forme de la fonction ReLU.....	29
Figure 2.7 La couche fully-connected.	29
Figure 2.8 Exemple d'architecture d'un CNN.	30
Figure 3.1 Le détecteur HOG.....	36
Figure 3.2 Détection de visage sur une image statique.....	36
Figure 3.3 Détection de visage sur une vidéo pré-enregistrée.	37
Figure 3.4 Détection de visage en temps réel	37
Figure 3.5 Détection d'expression faciale dans une image statique	38
Figure 3.6 Détection d'expression faciale dans une vidéo pré-enregistrée.....	39
Figure 3.7 Détection d'expression faciale en temps réel	39
Figure 3.8 Classification d'âge et de sexe dans une image statique.....	41
Figure 3.9 Classification d'âge et de sexe dans une vidéo pré-enregistrée.....	41
Figure 3.10 Classification d'âge et de sexe en temps réel	42
Figure 3.11 L'interface de notre système	42
Figure 3.12 Choisir une image.....	43
Figure 3.13 Le résultat sur une image statique	43
Figure 3.14 Le résultat sur une vidéo pré-enregistrée.....	44
Figure 3.15 Le résultat en temps réel	44

Liste des tableaux :

Tableau 2.1 Les avantages et les inconvénients des réseaux de neurones artificiels.....	26
Tableau 2.2 Les types de pooling	28
Tableau 3.1 Les métriques de performance	45

Liste des abréviations :

IA Intelligence Artificielle.

ML Machine Learning.

DL Deep Learning.

ANN Artificial Neural Network.

CNN Convolution Neural Network.

DNN Deep Neural Network.

FC Fully-Connected.

ReLU Rectified Linear Units.

HOG Histogram of Oriented Gradients.

Table de matières :

: الملخص.....	5
Liste de figures :	6
Liste des tableaux :	7
Liste des abréviations :	8
Introduction Générale.....	10
Chapitre 1 : La reconnaissance faciale	13
1- Introduction	13
2- Qu'est-ce que la reconnaissance faciale ?	13
3- Le fonctionnement de la reconnaissance faciale	13
3-1 Détection du visage.....	13
3-2 Analyse du visage	14
3-3 Conversion de l'image	14
3-4 Reconnaissance faciale	15
3-	15
4- Principales difficultés de la reconnaissance faciale	15
4-1 Variation d'éclairage.....	15
4-2 Variation de poses.....	16
4-3 Variation d'expressions faciales	16
4-4 Présence ou absence des composants structurels	17
4-5 Occultations partielles.....	17
5- Les applications de la reconnaissance faciale	18
6- Les avantages et les inconvénients de la reconnaissance Faciale	18
6-1 Les avantages.....	18
6-2 Les inconvénients.....	19
7- Conclusion.....	19
Chapitre 2 : L'apprentissage profond (Deep Learning).....	21
1- Introduction	21
2- Concepts généraux	21
2-1 L'intelligence artificielle (IA).....	21
2-2 L'apprentissage automatique	21
3- L'apprentissage profond (Deep Learning)	21
3-1 Définition.....	21
3-2 Le fonctionnement de l'apprentissage profond	22

3-3 Les réseaux de neurones artificiels.....	22
3-3-1 Définition	22
3-3-2 La structure d'un réseau de neurones artificiels	23
3-3-3 Le fonctionnement d'un réseau de neurones artificiels	25
3-3-4 Les avantages et les inconvénients des réseaux de neurones artificiels.....	26
3-4 Les algorithmes d'apprentissage profond	26
3-5 Réseaux de neurones convolutifs (CNN)	26
3-5-1 Définition	26
3-5-2 Le fonctionnement des différents types de couches d'un CNN	27
3-5-3 L'architecture d'un CNN.....	29
3-5-4 Les CNNs pré-entraînés.....	30
3-5-5 Les principales applications d'un CNN.....	31
3-6 Les applications d'apprentissage profond (DL).....	31
4- L'apprentissage profond et la reconnaissance faciale.....	31
5- Conclusion.....	32
Chapitre 3 : Analyse et Résultats	34
1- Introduction	34
2- Outils de développement.....	34
3- Les étapes du développement	35
3-1 Détection de visage.....	35
3-1-1 Histogram of oriented Gradients (HOG).....	35
3-1-2 Implémentation	35
3-2 Détection d'expressions faciales	37
3-2-1 The Kaggle facial expression recognition challenge dataset	37
3-2-2 Implémentation	37
3-3 Classification d'âge et de sexe	39
3-3-1 Adience collection of unfiltered faces for gender and age classification	39
3-3-2 Implémentation	40
4- L'interface de développement.....	42
5- Résultat et interprétation :.....	44
5-1 Précision	44
5-2 Rappel	45
5-3 Score F1	45
5-4 Exactitude	45
5-5 Taux d'erreur.....	45
6- Conclusion.....	45

Conclusion Générale	46
Bibliographies	47

Introduction Générale

- **Contexte**

La biométrie est une science qui implique l'analyse de caractéristiques physiques ou comportementales spécifiques de chaque personne, permettant de vérifier son identité de manière fiable. Elle présente des avantages non négligeables par rapport aux technologies d'authentification traditionnelles (ex : mots de passe, cartes à puce, etc.), d'autant qu'une partie de la biométrie est inaccessibles, unique à chaque individu et ne peut être ni perdue ni volée. Le choix de la solution biométrique dépend de l'acceptation par l'utilisateur, du niveau de sécurité, du coût de précision et du délai de mise en œuvre.

Au cours de la dernière décennie, la reconnaissance faciale a été l'un des domaines de recherche les plus actifs en matière de vision par ordinateur et de reconnaissance de formes, et de nombreuses applications s'appuient sur cette technologie car le visage humain est considéré comme une source d'informations très importante, si riche que certains psychologues le considèrent être la « fenêtre sur l'âme ». En particulier, l'expression faciale, le sexe et l'âge sont des caractéristiques faciales très utiles pour diverses applications, telles que l'utilisation de systèmes automatisés de prédiction du sexe, de l'âge et d'émotion pour analyser les clients intéressés par des produits ou des publicités ciblées.

Détecter l'expression faciale, prédire le sexe et estimer l'âge sont des tâches que les humains accomplissent naturellement et sans effort dans la vie quotidienne ; doter les ordinateurs de telles capacités de raisonnement reste un défi de taille.

Afin de donner aux ordinateurs une spécificité pour la reconnaissance faciale et la classification automatique des objets, plusieurs méthodes ont été proposées ces dernières années. Dans notre travail, nous avons choisi les réseaux de neurones convolutifs (CNN) qui appartiennent à la catégorie des réseaux de neurones feedforward basés sur l'apprentissage profond.

- **Objectifs**

L'objectif de notre travail consiste à implémenter un système qu'on a nommé KE-MA FACE RECOGNITION pour détecter le visage, l'expression faciale et prédire l'âge et le sexe dans les médias visuels (Image, vidéo et en temps réel).

Ce mémoire s'articule autour de trois chapitres qui nous permettront de présenter les différents aspects de notre travail.

Chapitre 1 : Ce chapitre est consacré à des généralités sur la reconnaissance faciale et son mode de fonctionnement.

Chapitre 2 : Ce chapitre est consacré à la présentation de l'apprentissage profond, où nous détaillons plus les réseaux de neurone convolutifs.

Chapitre 3 : Dans ce dernier chapitre, nous allons présenter les résultats expérimentaux obtenus par notre système de reconnaissance de visage, détection d'expression et classification de genre et d'âge suivi d'une discussion avec interprétation des résultats.

CHAPITRE 1 :

La reconnaissance Faciale.

Chapitre 1 : La reconnaissance faciale

1- Introduction

La reconnaissance faciale est récemment devenue un domaine de recherche très actif, en raison de l'intérêt accru pour les systèmes de sécurité biométriques en général. Le début de ce chapitre présente la reconnaissance faciale, ensuite les différentes étapes de son fonctionnement. La seconde moitié du chapitre se penchera sur les principales difficultés qui entravent la technologie de reconnaissance faciale, ainsi que sur les nombreuses applications de cette technologie. Enfin, les avantages et les inconvénients de cette technologie seront discutés.

2- Qu'est-ce que la reconnaissance faciale ?

La reconnaissance faciale est une parmi tant d'autres méthodes d'identification biométrique. La biométrie implique l'identification d'un individu grâce à ses caractéristiques physiques, biologiques et parfois comportementales. D'autres exemples de ce type de technologie incluent la reconnaissance des empreintes digitales, la reconnaissance de l'iris ou de la rétine et la reconnaissance vocale. Lorsqu'il s'agit de reconnaissance faciale, il s'agit d'une technologie qui permet d'identifier ou de vérifier une personne à partir d'une image numérique, d'une séquence vidéo ou en temps réel. Son principal cas d'utilisation est souvent celui de mécanisme de contrôle d'accès dans les systèmes de sécurité. [1]

3- Le fonctionnement de la reconnaissance faciale

La plupart des gens connaissent la technologie de reconnaissance faciale, utilisée dans FaceID pour déverrouiller les smartphones et les iPhones. De manière générale, la technologie de reconnaissance faciale ne dépend pas d'une vaste base de données d'images pour reconnaître l'identité d'une personne. Sa fonction principale est de reconnaître et de valider le propriétaire de l'appareil pour empêcher tout accès non autorisé. [2]

Maintenant, lorsqu'il ne s'agit pas d'un appareil personnel, la reconnaissance faciale fonctionne différemment. L'une des principales différences est que le système scanne un visage et essaie de le faire correspondre à d'autres images stockées dans une base de données de visages. Essentiellement, il ne s'agit pas de se souvenir d'un individu en particulier, mais d'identifier si cet individu existe au sein du système.

La reconnaissance faciale implique essentiellement quatre étapes, sont :

3-1 Détection du visage

Est le processus de localisation automatique des visages humains dans les médias visuels (images numériques ou vidéo) (Figure 1.1). Pour trouver un visage, l'image peut être recherchée pour des repères tels que les yeux et le nez.

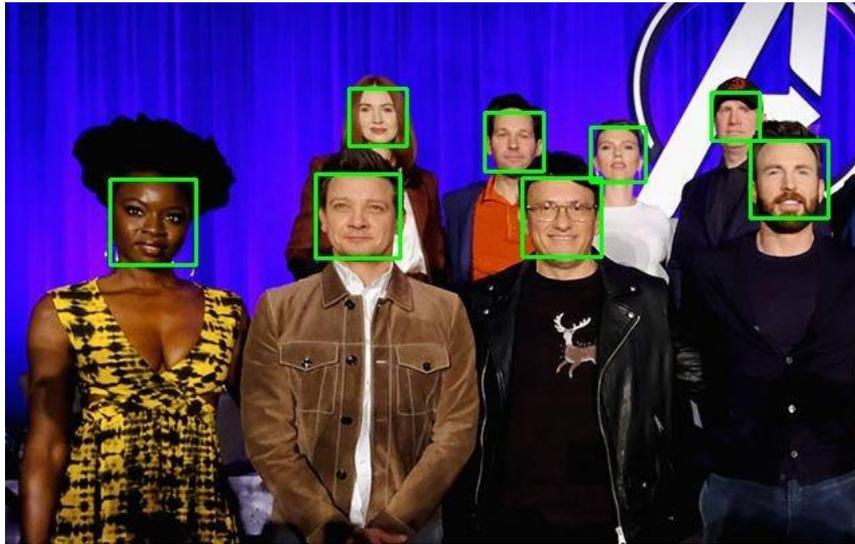


FIGURE 1.1 DÉTECTION DE VISAGE. [1]

3-2 Analyse du visage

Cette étape consiste à relever les traits distinctifs en se basant sur la géométrie du visage. L'analyse du visage implique d'abord celle des yeux, notamment leur distance et la profondeur des orbites (Figure 1.2). En outre, le système évalue la forme du menton ainsi que sa distance par rapport au front. Il étudie également la forme des pommettes, des lèvres et des oreilles. L'objectif est de créer des points de repère qui permettent d'identifier le visage. [3]

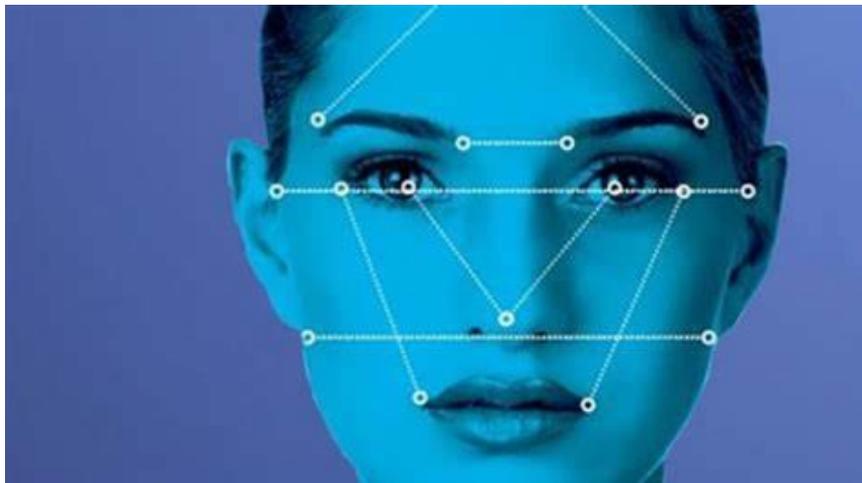


FIGURE 1.2 ANALYSE DU VISAGE. [3]

3-3 Conversion de l'image

Cette étape consiste à impliquer la reconnaissance de caractéristiques faciales uniques basées sur sa géométrie. L'examen du visage commence par les yeux, plus précisément leur

proximité les uns par rapport aux autres et la profondeur des orbites. De plus, le système évalue le contour de la mâchoire et sa distance par rapport à la racine des cheveux. L'étude englobe également la forme des pommettes, des lèvres et des oreilles. L'objectif ultime est d'établir des repères identifiables qui distinguent le visage de l'individu.

3-4 Reconnaissance faciale

La dernière phase est celle qui répond le plus à la définition de la reconnaissance faciale. Au cours de cette étape, l'empreinte faciale est méticuleusement comparée à une vaste base de données de visages pour découvrir une correspondance. Cette compilation de données d'images peut provenir d'une multitude d'emplacements, notamment des caméras publiques, des plateformes de médias sociaux et des bases de données privées. [2]



FIGURE 1.3 RECONNAISSANCE FACIALE. [3]

4- Principales difficultés de la reconnaissance faciale

4-1 Variation d'éclairage

Un Certains facteurs tels que l'éclairage (répartition de la source de lumière, intensité, spectre) et les caractéristiques de la caméra affectent l'apparence d'un visage dans l'image acquise. [4]



FIGURE 1.4 EXEMPLE DE VARIATION D'ÉCLAIRAGE. [4]

4-2 Variation de poses

Lorsqu'il y a des changements de pose dans l'image, le taux de reconnaissance faciale diminue considérablement. Les tests d'évaluation développés sur les bases de données FERET et FRVT ont démontré cette difficulté [Bla01 ; Φ03]. La variation de pose a été identifiée comme un problème majeur pour les systèmes de reconnaissance faciale. Lorsqu'un visage est de côté (orientation $< 30^\circ$) dans le plan de l'image, il peut être normalisé en détectant au moins deux traits du visage (au niveau des yeux). Cependant, lorsque la rotation est supérieure à 30° , la normalisation géométrique n'est plus possible. [4]



FIGURE 1.5 EXEMPLE DE VARIATION DE POSES. [4]

4-3 Variation d'expressions faciales

Les déformations du visage causées par les expressions faciales touchent principalement la partie inférieure du visage. Les informations faciales trouvées sur le dessus du visage restent presque inchangées, ce qui est généralement suffisant pour le processus de reconnaissance. Cependant, comme les expressions faciales modifient l'apparence du visage (Figure 6), le taux de reconnaissance va inévitablement baisser. La reconnaissance faciale des expressions faciales est un problème toujours d'actualité et non résolu. [5]

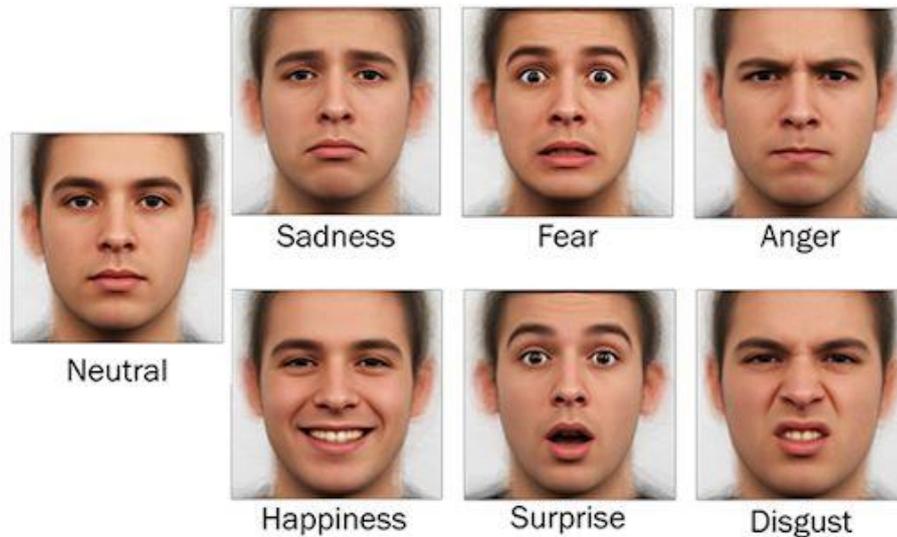


FIGURE 1.6 EXEMPLE DE VARIATION D'EXPRESSIONS FACIALES. [5]

4-4 Présence ou absence des composants structurels

La présence des composants structurels comme la barbe, la moustache ou les lunettes peuvent altérer grandement les caractéristiques du visage comme la forme, la couleur ou la taille du visage. De plus, ces composants peuvent masquer des caractéristiques faciales essentielles, provoquant l'échec des systèmes de reconnaissance. Par exemple, une moustache ou une barbe peuvent modifier la forme du visage (Figure 1.7). [5]



FIGURE 1.7 EXEMPLE DE PRÉSENCE OU ABSENCE DES COMPOSANTS STRUCTURELS. [5]

4-5 Occultations partielles

Le visage peut être partiellement masqué par d'autres objets recouvrant le visage. En effet, par exemple, dans une image contenant un groupe de personnes, les visages peuvent en masquer partiellement d'autres.

5- Les applications de la reconnaissance faciale

Les systèmes de reconnaissance faciale sont plus applicables à la vie quotidienne. Des exemples d'applications de reconnaissance faciale [3], comprennent :

- **Authentification** : La reconnaissance faciale est utilisée pour authentifier les utilisateurs dans diverses applications telles que les smartphones, les ordinateurs portables, les systèmes bancaires en ligne, etc.
- **Surveillance de la sécurité** : La reconnaissance faciale est utilisée pour la surveillance de la sécurité dans les lieux publics tels que les aéroports, les gares, les banques, les centres commerciaux, les hôtels, etc.
- **Marketing** : Les entreprises utilisent la reconnaissance faciale pour identifier les émotions des clients lorsqu'ils visitent les magasins, les centres commerciaux ou les restaurants. Cela permet aux entreprises de personnaliser leurs offres et d'améliorer l'expérience client.
- **Santé** : La reconnaissance faciale est utilisée pour diagnostiquer les troubles médicaux tels que le syndrome de Down, la progeria, le syndrome de Williams, etc.
- **Gestion des ressources humaines** : Les entreprises utilisent la reconnaissance faciale pour surveiller la présence des employés au travail, pour faciliter l'identification des visiteurs ou pour la reconnaissance des employés sur des bases de données.
- **Sécurité des réseaux sociaux** : Les plateformes de réseaux sociaux utilisent la reconnaissance faciale pour identifier les personnes dans les photos et les vidéos et pour prévenir les cyberharcèlements et les fraudes.
- **Loisirs** : La reconnaissance faciale est utilisée pour le divertissement, notamment pour faire des selfies, des filtres de photos pour les réseaux sociaux, les jeux vidéo et les applications de réalité augmentée. [3]

6- Les avantages et les inconvénients de la reconnaissance Faciale

6-1 Les avantages

- **Rapidité et efficacité** : la reconnaissance faciale permet de reconnaître rapidement les visages, ce qui est particulièrement utile dans les lieux de grande affluence tels que les aéroports ou les stades.
- **Sécurité** : la reconnaissance faciale peut être utilisée pour identifier les criminels ou pour protéger des zones sensibles, telles que des bases militaires.
- **Expérience utilisateur améliorée** : la technologie de reconnaissance faciale peut être utilisée pour faciliter l'accès aux appareils mobiles, tels que les smartphones, en remplaçant les mots de passe.

- **Amélioration de la qualité de vie** : la reconnaissance faciale peut être utilisée pour aider les personnes atteintes de troubles cognitifs, telles que la maladie d'Alzheimer, en leur permettant de reconnaître plus facilement les personnes de leur entourage.

6-2 Les inconvénients

- **Vie privée** : la reconnaissance faciale peut être considérée comme une violation de la vie privée, car elle permet de suivre et d'identifier les individus sans leur consentement.
- **Erreurs de reconnaissance** : la technologie de reconnaissance faciale n'est pas infaillible et peut conduire à des erreurs qui peuvent avoir des conséquences graves, telles que l'identification erronée d'un suspect.
- **Discrimination** : la reconnaissance faciale peut introduire des biais et des discriminations en fonction de la race, de l'âge ou du sexe, ce qui peut conduire à des conséquences négatives pour les personnes concernées.
- **Coût** : la mise en place de systèmes de reconnaissance faciale peut être coûteuse, ce qui peut limiter la portée de cette technologie dans certains contextes.

7- Conclusion

Dans ce chapitre, nous avons défini la technologie de reconnaissance faciale et son histoire. Nous décrivons également les étapes de la reconnaissance faciale. Nous avons également évoqué les différentes difficultés inhérentes à la technologie de reconnaissance faciale et les différentes applications de cette technologie. Enfin, nous avons souligné ses avantages et ses inconvénients.

CHAPITRE 2 :

L'apprentissage Profond
(Deep Learning).

Chapitre 2 : L'apprentissage profond (Deep Learning).

1- Introduction

L'apprentissage profond est devenu un outil essentiel pour la reconnaissance faciale. Dans ce chapitre, nous présentons une étude détaillée sur le Deep Learning et nous aborderons également l'essentiel sur les réseaux de neurones artificiels (ANN). Ce chapitre vise aussi à fournir une présentation au concept de réseaux de neurones convolutionnels (CNNs) en détaillant ses différents types de couche d'un CNN, son architecture et son fonctionnement. etc. Après cela, nous terminerons cette discussion en expliquant l'efficacité de l'apprentissage profond dans la reconnaissance faciale.

2- Concepts généraux

2-1 L'intelligence artificielle (IA)

(ou **(AI)** en anglais pour **Artificial Intelligence**) est un domaine de l'informatique dont l'objectif est de créer des programmes qui exécutent des tâches normalement assignées à l'intelligence humaine et de la simuler. [7]

2-2 L'apprentissage automatique

(ou **(ML)** en anglais pour **Machine Learning**) est une branche de l'intelligence artificielle (Figure 2.1), elle a la capacité d'apprendre à partir de données en utilisant un algorithme d'apprentissage dont l'objectif est d'effectuer des analyses explicatives, prédictives ou préventives. [7]

3- L'apprentissage profond (Deep Learning)

3-1 Définition

Est sous-ensemble de l'apprentissage automatique (Figure 2.1), où les réseaux neuronaux artificiels des algorithmes conçus pour fonctionner comme le cerveau humain émettent à partir de grandes quantités de données. [7]

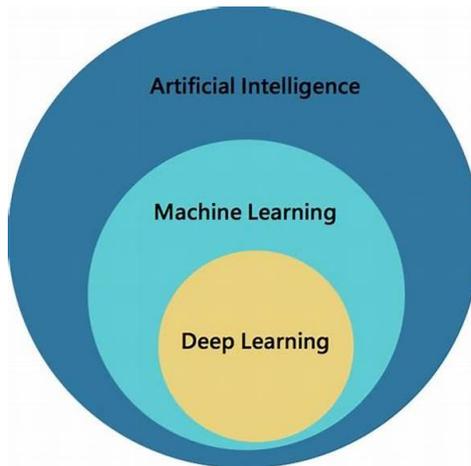


FIGURE 2.1 LA RELATION ENTRE AI, ML ET DL. [7]

3-2 Le fonctionnement de l'apprentissage profond

Le Deep Learning exploite des réseaux de neurones artificiels (appelés aussi « **perceptrons** ») inspirés directement du fonctionnement des neurones humains. Chaque perceptron prend des données et effectue une série d'opérations de base, puis transforme ses résultats via une fonction non linéaire appelée « **fonction d'activation** ». Le réseau est composé de plusieurs couches d'unités neuronales, parfois au nombre de dizaines, voire de centaines de couches. Chaque couche est composée de perceptrons parallèles qui s'envoient et reçoivent des informations entre eux. Les résultats de ces calculs sont ensuite transmis d'une couche à la suivante jusqu'à ce que la couche ultime génère **une prédiction**. [8]

3-3 Les réseaux de neurones artificiels

3-3-1 Définition

Les réseaux de neurones artificiels (**ANN**, (ou Deep Neural network)), également appelés Perceptrons, sont utilisés dans les algorithmes d'apprentissage automatique et constituent un élément clé de l'apprentissage en profondeur. Ces réseaux sont structurés et portent le nom du cerveau humain car ils copient la manière dont les neurones biologiques communiquent entre eux (Figure 2.2). Au minimum, il se présente sous la forme d'au moins deux couches de neurones qui interagissent les unes avec les autres et s'adaptent pour apprendre à effectuer une tâche spécifique en recevant des données d'apprentissage.

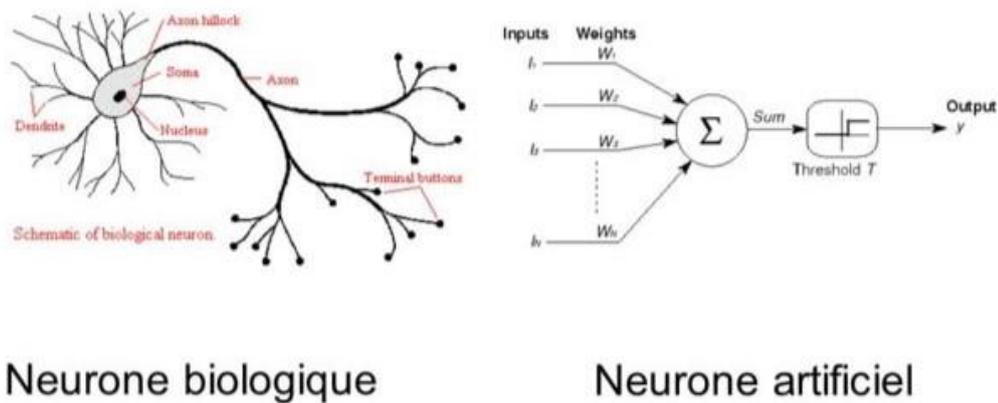


FIGURE 2.2 NEURONE BIOLOGIQUE ET NEURONE ARTIFICIEL. [8]

3-3-2 La structure d'un réseau de neurones artificiels

Un réseau neuronal typique est composé de plusieurs couches disposées en séquence, chaque couche suivante recevant les entrées des sorties de la couche précédente. Chaque couche (i) est constituée de N neurones dont chacun obtient ses entrées des N neurones de la couche précédente. Chaque synapse est associée à un poids synaptique, qui est multiplié par les entrées puis résumé par les neurones à ce niveau. Ce processus équivaut à la multiplication du vecteur d'entrée par une matrice de transformation. Si les différentes couches du réseau neuronal étaient placées les unes à la suite des autres, cela équivaudrait à mettre en cascade plusieurs matrices de transformation en une seule matrice obtenue à partir de leur produit. Cependant, la présence d'une fonction de sortie à chaque couche, qui introduit une non-linéarité à chaque étape, rend cela impossible. Cela souligne l'importance de sélectionner une fonction de sortie appropriée. Un réseau de neurones avec des sorties linéaires ne sert à rien.

Au-delà de cette structure simple (Figure 2.3), Les boucles au sein d'un réseau neuronal peuvent potentiellement modifier radicalement ses capacités et ses subtilités. De la même manière que les boucles peuvent modifier la logique combinatoire et la rendre séquentielle, les boucles dans un réseau neuronal peuvent transformer un système de reconnaissance d'entrée de base en une machine à multiples facettes capable d'un large éventail de comportements.

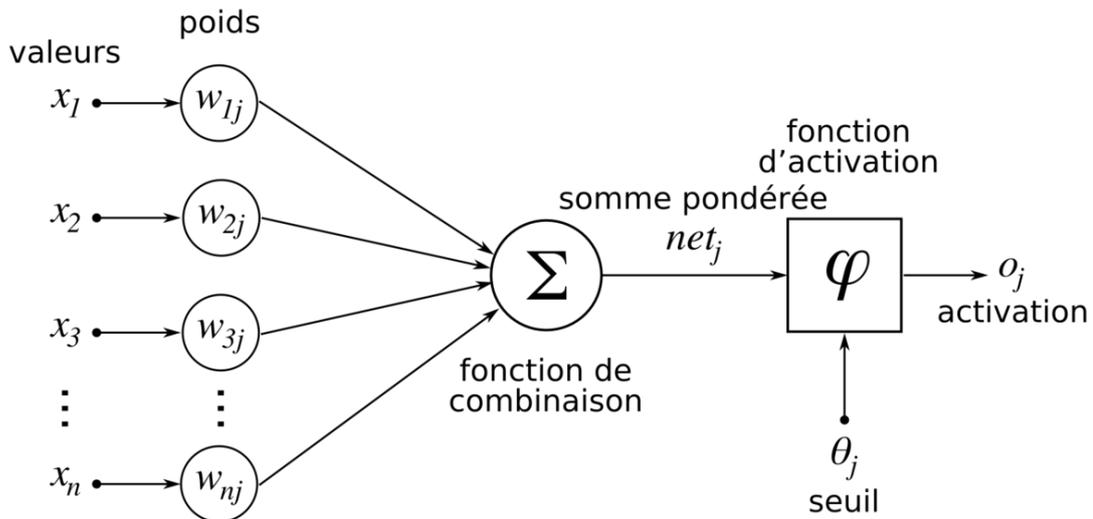


FIGURE 2.3 STRUCTURE D'UN RESEAU DE NEURONES ARTIFICIELS. [9]

Une fois que le neurone reçoit ses entrées \mathbf{x} , il procède au calcul de leur somme, qui est ensuite pondérée par les valeurs \mathbf{w} . La valeur résultante est ensuite soumise à la fonction d'activation ϕ , qui est chargée de générer la sortie \mathbf{o} . [9]

- **Fonction de combinaison** : Considérons un neurone quelconque. Grâce à ses connexions synaptiques, le neurone absorbe un nombre spécifié de valeurs provenant des neurones en amont, puis génère une valeur singulière à l'aide **d'une fonction de combinaison**. Cette fonction peut être formalisée comme une fonction vectorielle à scalaire, notamment :
 - Le type de réseaux **MLP** (Multi-Layer Perceptron) calcule une combinaison linéaire d'entrées. En d'autres termes, la fonction de combinaison produit le produit scalaire du vecteur d'entrée et du vecteur de poids synaptique.
 - Le type de réseaux **RBF** (Radial Basis Function). [9]
- **Fonction d'activation** : également connue sous le nom de fonction de **seuillage** ou fonction de **transfert**, est utilisée pour intégrer la non-linéarité dans le fonctionnement d'un neurone. Les fonctions de seuillage sont généralement caractérisées par trois intervalles :
 1. Lorsque le neurone descend en dessous d'un seuil, il devient inactif (souvent dans ce cas, sa sortie vaut 0 ou -1).
 2. Aux alentours du seuil, une phase de transition.
 3. Lorsque le neurone dépasse le seuil, il devient actif (souvent dans ce cas, sa sortie vaut 1). [9]

Voici quelques exemples de fonction d'activation les plus utilisés en Deep Learning :

- Fonction **ReLU (Rectified Linear Unit)**.
- Fonction **ELU (Exponential Linear Unit)**.
- Fonction **sigmoïde**. [9]

3-3-3 Le fonctionnement d'un réseau de neurones artificiels

Un réseau neuronal se compose d'au moins deux couches, chacune contenant plusieurs neurones ou nœuds. D'une couche à l'autre, les nœuds sont liés entre eux. Chacun est associé à des données d'entrée, des poids, des seuils et des données de sortie (dont la valeur est le résultat de la multiplication de la valeur d'entrée du neurone par son poids). Si les données de sortie d'un nœud dépassent un seuil spécifié, ce neurone est activé et envoie ses données aux neurones de la couche suivante, et ainsi de suite. (Figure 2.4).

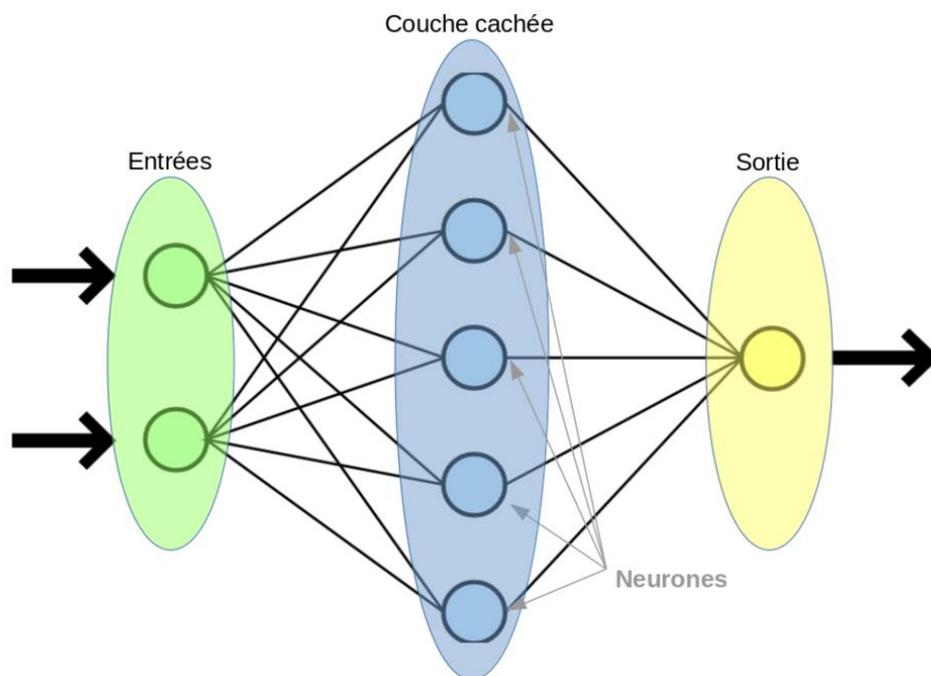


FIGURE 2.4 FONCTIONNEMENT D'UN RESEAU DE NEURONES ARTIFICIELS. [10]

3-3-4 Les avantages et les inconvénients des réseaux de neurones artificiels

Avantages	Inconvénients
<ul style="list-style-type: none">– Taux d’erreur généralement bon.– Robustesse (bruit, données manquantes).– Combinaison avec d’autres méthodes.– Classification rapide.– Outil disponible dans les environnement datamining.	<ul style="list-style-type: none">– Apprentissage très long.– Pouvoir explicatif faible (boite noire).– Plusieurs paramètres (architecture, poids de connections, pas de gradients).

TABLEAU 2.1 LES AVANTAGES ET LES INCONVÉNIENTS DES RÉSEAUX DE NEURONES ARTIFICIELS.

3-4 Les algorithmes d’apprentissage profond

En apprentissage profond, il existe différents algorithmes, chacun avec sa propre spécificité et son application. Voici une liste des 10 algorithmes d'apprentissage en profondeur les plus populaires :

- Réseaux de neurones convolutifs (CNN).
- Réseaux de neurones récurrents (RNN).
- Réseaux de fonction de base radiale (RBFN).
- Réseaux de mémoire à long et court terme (LSTM).
- Réseaux adversariaux génératifs (GAN).
- Machines de Boltzmann restreintes (RBM).
- Les perceptrons multicouches (MLP).
- Les réseaux de croyance profonds (DBN).
- Les cartes auto-organisées (SOM).
- Les auto-encodeurs.

Ce qui nous intéresse dans notre projet sont **Les réseaux de neurones convolutifs (CNN)**.

3-5 Réseaux de neurones convolutifs (CNN)

3-5-1 Définition

Un CNN, également connu sous le nom de **ConvNet**, est un type d'algorithme de deep learning, est souvent utilisé pour analyser et comprendre les caractéristiques visuelles grâce au traitement de quantités massives de données. Bien qu'il soit principalement appliqué aux tâches d'intelligence artificielle basées sur l'image, le CNN peut également être utilisé pour d'autres projets liés à l'IA, tels que le traitement du langage et le développement de moteurs de recommandation.

3-5-2 Le fonctionnement des différents types de couches d'un CNN

Lorsqu'il s'agit d'un réseau neuronal convolutif, il est composé de plusieurs couches qui sont responsables du traitement et de l'extraction des caractéristiques des données et qui transforment progressivement une image en entrée en un vecteur de probabilités en sortie. Ces couches sont généralement de quatre types différents :

- a) **La couche de convolution** : Est la première couche d'un CNN, Pour effectuer la convolution, il prend une image pixelisée carré par carré et la transforme à l'aide d'un calque connue sous le nom **feature detector** (filtre) et l'image transformée est appelée **feature map**. Il existe des millions de feature detector différents de différentes tailles, le choix se fait en fonction des best practice des data scientist. Une fois que l'image a complètement traversé le feature detector, nous avons une feature map. Nous répétons ce processus avec d'autres feature detector pour créer plusieurs feature map de l'image. Cela créera une sorte de couches composée de nombreuses feature map différentes (Figure 2.5). C'est ce qu'on appelle une **couche convolutive**. [10]

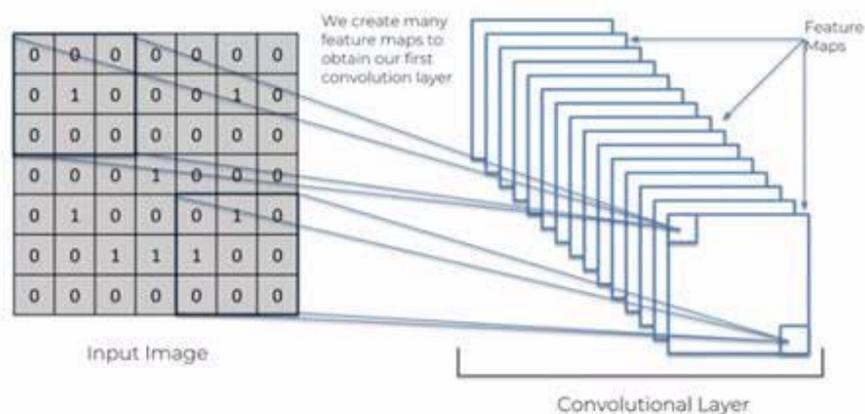


FIGURE 2.5 COUCHE DE CONVOLUTION. [10]

- b) **La couche de pooling** : Est fréquemment insérée entre deux couches de convolution. Le processus de pooling consiste à réduire la taille des images tout en conservant leurs caractéristiques essentielles. Deux types de pooling couramment utilisés sont max et average (Tableau 2.2). [11]

Type	Max Pooling	Average Pooling																																								
Objectif	Lors de chaque opération de pooling, valeur la plus élevée est sélectionnée.	Lors de chaque opération de pooling, la valeur moyenne est prise.																																								
Exemple	<p style="text-align: center;">Max Pooling</p> <table border="1" style="margin: auto;"> <tr><td>29</td><td>15</td><td>28</td><td>184</td></tr> <tr><td>0</td><td>100</td><td>70</td><td>38</td></tr> <tr><td>12</td><td>12</td><td>7</td><td>2</td></tr> <tr><td>12</td><td>12</td><td>45</td><td>6</td></tr> </table> <p style="text-align: center;">↓ 2 x 2 pool size</p> <table border="1" style="margin: auto;"> <tr><td>100</td><td>184</td></tr> <tr><td>12</td><td>45</td></tr> </table>	29	15	28	184	0	100	70	38	12	12	7	2	12	12	45	6	100	184	12	45	<p style="text-align: center;">Average Pooling</p> <table border="1" style="margin: auto;"> <tr><td>31</td><td>15</td><td>28</td><td>184</td></tr> <tr><td>0</td><td>100</td><td>70</td><td>38</td></tr> <tr><td>12</td><td>12</td><td>7</td><td>2</td></tr> <tr><td>12</td><td>12</td><td>45</td><td>6</td></tr> </table> <p style="text-align: center;">↓ 2 x 2 pool size</p> <table border="1" style="margin: auto;"> <tr><td>36</td><td>80</td></tr> <tr><td>12</td><td>15</td></tr> </table>	31	15	28	184	0	100	70	38	12	12	7	2	12	12	45	6	36	80	12	15
29	15	28	184																																							
0	100	70	38																																							
12	12	7	2																																							
12	12	45	6																																							
100	184																																									
12	45																																									
31	15	28	184																																							
0	100	70	38																																							
12	12	7	2																																							
12	12	45	6																																							
36	80																																									
12	15																																									
Propos	<ul style="list-style-type: none"> - Conserver les caractéristiques détectées. - Le plus utilisé. 	<ul style="list-style-type: none"> - Sous-échantillon la feature map. 																																								

TABLEAU 2.2 LES TYPES DE POOLING.

c) **La couche de correction ReLU (Rectified Linear Units)** : Désigne la fonction réelle non-linéaire définie par : $\text{ReLU}(x)=\max(0,x)$. La fonction de cette couche implique le remplacement de toutes les valeurs d'entrée négatives par des zéros (Figure 2.6). Il agit comme une fonction d'activation. [11]

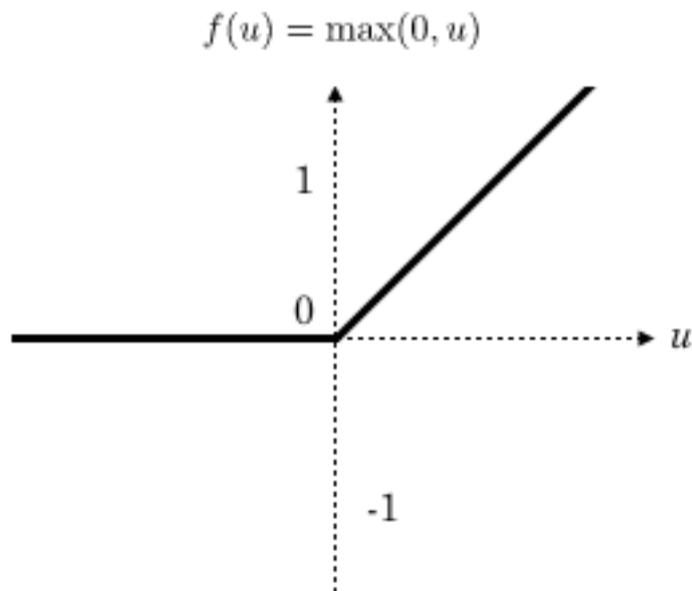


FIGURE 2.6 La forme de la fonction ReLU. [11]

- d) La couche fully-connected (FC) :** Dans cette couche, chaque nœud de la couche de sortie est directement lié au nœud de la couche précédente (Figure 2.7). Cette couche est responsable de la classification d'une image en fonction des caractéristiques qui ont été extraites via les couches précédentes avec leurs filtres correspondants. [10]

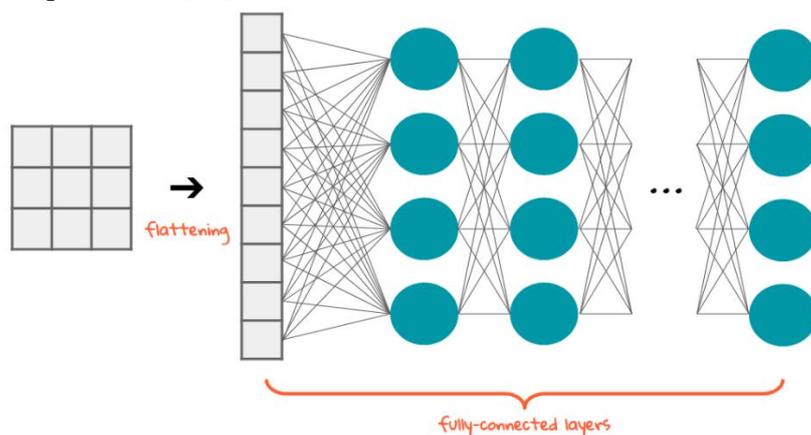


FIGURE 2.7 LA COUCHE DE FULLY-CONNECTED. [10]

3-5-3 L'architecture d'un CNN

Après avoir expliqué le fonctionnement des différents types de couches dans un réseau de neurones convolutifs, nous allons voir son architecture. Un CNN est simplement composé de plusieurs couches convolutives, de couches de pooling, de couches de correction ReLU et de couches fully-connected. Ainsi, chaque image entrée subira divers filtres, réductions et corrections, pour finalement former un vecteur. Dans les problèmes de classification, ce vecteur contient la probabilité d'appartenir à une classe. La couche initiale

de tout réseau neuronal convolutif est la couche de convolution et la couche finale est la couche de fully-connected (Figure 2.8).

Les couches intermédiaires ont une flexibilité dans leur disposition, tant que la sortie d'une couche conserve la même structure que l'entrée de la couche suivante. Exemple, une couche de pooling, qui nécessite une matrice 3D en entrée, ne peut pas précéder une couche entièrement connectée, qui produit toujours un vecteur.

En règle générale, un CNN empile plusieurs couches de convolution et de correction ReLU, puis ajoute (facultatif) des couches de pooling et répète ce modèle plusieurs fois; puis, il empile des couches fully-connected.[12]

Remarque : Plus il y a de couches, plus le CNN devient profond (nous sommes en plein Deep Learning). [13]

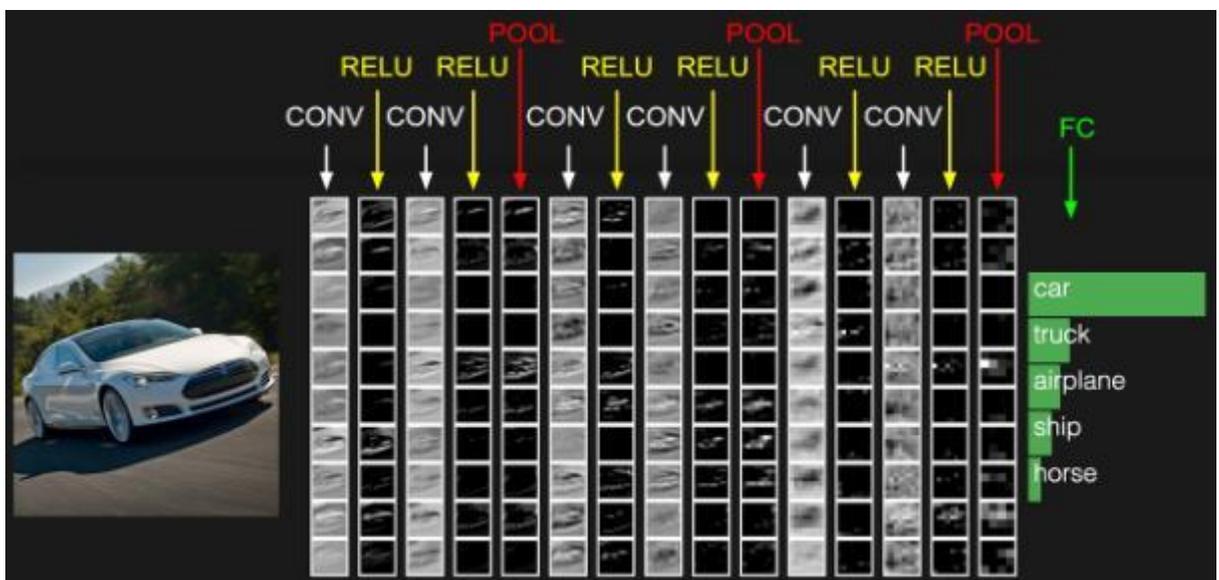


FIGURE 2.8 EXEMPLE D'ARCHITECTURE D'UN CNN. [12]

3-5-4 Les CNNs pré-entraînés

Sont des modèles qui ont été formés sur de grands ensembles de données (grande base de données), tels qu'ImageNet, et peuvent être réutilisés pour diverses tâches, telles que la classification d'images, la détection d'objets ou la reconnaissance faciale. Les modèles CNN pré-entraînés sont disponibles dans différents frameworks de Deep Learning tels que TensorFlow, PyTorch, Keras, etc. [14]

L'entraînement d'un CNN implique de déterminer et de calculer empiriquement la valeur de chaque poids. Voici comment cela fonctionne : un CNN traite une image (à partir d'une base de données d'entraînement) et, en sortie, effectue une prédiction, par exemple à quelle classe il pense que l'image appartient. Sachant que l'on connaît à l'avance la classe de chaque image d'entraînement, on peut vérifier que ce résultat est correct. Sur la base de

l'exactitude de ce résultat, nous mettons à jour tous les poids du CNN selon un algorithme appelé rétropropagation du gradient d'erreur. Pendant la phase de l'entraînement du modèle, le processus ci-dessus est répété plusieurs fois pour toutes les images de la base de données d'entraînement. L'objectif du modèle est de classer au mieux ces données. Une fois que le modèle a fini de mettre à jour ses poids, on évalue le modèle en l'alimentant dans la base de données de validation. Il classe toutes ces images (images que le modèle n'a jamais vues auparavant) puis on calcule son bon taux de classification, que l'on appelle la précision du modèle. [14]

3-5-5 Les principales applications d'un CNN

- **La reconnaissance d'images** est l'une des principales applications des CNNs. Ces réseaux sont connus pour apprendre à un rythme accéléré et posséder un taux d'erreur supérieur. De plus, ils sont utilisés dans une certaine mesure pour **analyser le contenu vidéo**.
- Les modèles CNN sont particulièrement compétents lorsqu'il s'agit de tâches telles que **l'analyse sémantique, la modélisation de phrases, la classification ou la traduction**. [15]
- Les CNNs ont également été utilisés dans le domaine de **la découverte de médicaments**, leur fonction permet d'identifier des traitements prometteurs en prévoyant les interactions pouvant survenir entre protéines et molécules biologiques. [15]
- Les CNNs sont reconnus pour leurs performances exceptionnelles dans le **développement de logiciels des jeux** (précisément les jeux de Go ou d'échecs). Etc.

3-6 Les applications d'apprentissage profond (DL)

L'apprentissage profond est aujourd'hui utilisé dans un certain nombre de domaines, exemples :

- La reconnaissance faciale.
- Assistance Vocale.
- Rédaction de textes.
- Diagnostics médicaux.
- Transaction financière.
- Cyber sécurité.
- Services de traduction automatique.
- Conduite autonome.

4- L'apprentissage profond et la reconnaissance faciale

Au cours des dernières années, l'utilisation de la technologie de reconnaissance faciale a explosé et ses applications se sont étendues à de nombreux secteurs, notamment les médias

sociaux, la surveillance, la sécurité et les médias. En utilisant le concept d'apprentissage profond, le système est capable de collecter automatiquement des connaissances à partir des données fournies.

L'efficacité de l'apprentissage profond dans la reconnaissance faciale ne peut être niée, car il a le potentiel de traiter de grandes quantités de données et d'atteindre des taux de précision remarquables. L'un des principaux avantages du Deep Learning est sa capacité à extraire automatiquement des informations à partir d'images, de vidéos ou d'en temps réel, notamment les traits du visage, les teintes et les contrastes, de manière entièrement automatisée. Cette capacité conduit à un système de reconnaissance fiable et précis, car le réseau neuronal est capable d'apprendre les propriétés cruciales de l'identification. Essentiellement, l'apprentissage profond est devenu un outil essentiel pour la reconnaissance faciale en raison de sa capacité à apprendre et à traiter les informations de manière efficace et précise.

5- Conclusion

Dans cette partie, nous avons présenté les contextes théorique de l'apprentissage profond, son fonctionnement, ses algorithmes et les domaines de son application., nous avons également présenté les fondements essentiels des réseaux de neurones artificiels. Ensuite nous avons abordé le concept de CNN, et nous avons défini ses différentes couches ainsi que son architecture et son fonctionnement pour notre besoin de reconnaissance de visage. Enfin, nous avons défini l'efficacité de l'apprentissage profond dans la reconnaissance faciale.

CHAPITRE 3 :

Développement & Implémentation.

Chapitre 3 : Analyse et Résultats

1- Introduction

Dans ce chapitre, je décris le processus de création du système de reconnaissance faciale « KE-MA Face Recognition ». Ceci est réalisé en précisant les outils de développement et ses différentes étapes, et un aperçu sur l'interface de ce système. Et enfin on donne un résultat et une interprétation.

2- Outils de développement

Dans cette section, nous présentons l'environnement matériel et logiciel utilisé pour implémenter notre projet. Les caractéristiques sont les suivantes :

- **Le langage Python** : Pour atteindre notre objectif, nous avons utilisé la version 3.10 du langage de programmation Python. Python est relativement facile à apprendre, open source, gratuit, interprété et le plus utilisé par les informaticiens de nos jours. Python a été créé au début des années 1990 par Guido van Rossum aux Pays-Bas pour succéder au langage appelé ABC.
- Afin de se concentrer sur ce projet et de profiter de la puissance du langage Python, j'ai utilisée les outils suivants :
 - **Anaconda** : est une distribution libre et open source des langages de programmation comme Python, data science et Machine Learning. J'ai installée dedans le reste de dépendance et des bibliothèques y compris : [17]
 - **OpenCv** : Une bibliothèque open source, considérée comme l'outil standard pour la Computer Vision (vision par ordinateur) et le traitement d'images et de vidéos.
 - **Dlib** : Une bibliothèque d'apprentissage profond, qui est utilisée pour la détection des visages, face mapping et la reconnaissance des visages.
 - **CMake** : Un système de construction logicielle multiplateforme qui est utilisé pour compiler et installer Dlib.
 - **Face_recognition** : Une bibliothèque qu'on l'a utilisée principalement dans notre projet, qui est un outil très simple à utiliser. Elle enveloppe autour de la fonctionnalité de reconnaissance faciale de Dlib.
 - **Numpy** : Une bibliothèque fondamentale pour le calcul numérique en Python, est utilisée pour travailler avec des pixels sous forme de tableaux et de matrices.
 - **Keras** : Une bibliothèque open source de Python, elle permet d'interagir avec les algorithmes de réseaux de neurones profonds et d'apprentissage automatique, notamment Tensorflow.
 - **Tensorflow** : Est une bibliothèque open source de Machine Learning, créée par Google, permettant de développer et d'exécuter des applications de Machine Learning et de Deep Learning.
 - **Pillow** : Est une bibliothèque de traitement d'images pour le langage de programmation Python

- **CustomTK** : CustomTkinter est une extension tkinter qui est une bibliothèque python, qui fournit des éléments d'interface utilisateur supplémentaires comme CtkButton,
- **Spyder** : Pour développer ce système nous avons utilisé l'environnement Spyder, version 5.4.1. Spyder est un environnement de développement pour Python. Libre et multiplateforme, il intègre de nombreuses bibliothèques d'usage scientifique comme : Numpy. [17]

3- Les étapes du développement

Notre système « Ke-Ma Face Recognition » est élaboré par trois étapes, sont :

3-1 Détection de visage

Est un processus de localisation automatique des visages humains dans les médias visuels (image statique, vidéo pré- enregistrée ou temps réel). Pour cette étape, nous avons utilisé le détecteur le plus populaire **HOG**.

3-1-1 Histogram of oriented Gradients (HOG)

Est un descripteur de caractéristiques utilisé en Computer Vision, en traitement d'images à des fins de détection d'objets dans des images, en particulier pour des objets ayant des contours bien définis, tels que des visages, des véhicules ou des personnes. Il a été largement utilisé dans des applications telles que la détection de piétons, la reconnaissance de gestes, la surveillance de vidéos, etc. HOG s'appuie sur la distribution du gradient d'intensité ou direction de bords (Figure 3.1), il fonctionne rapide, sous petit occlusion et la taille minimale du visage doit être de 80x80 pixels. [17]



FIGURE 3.1 LE DÉTECTEUR HOG. [17]

3-1-2 Implémentation

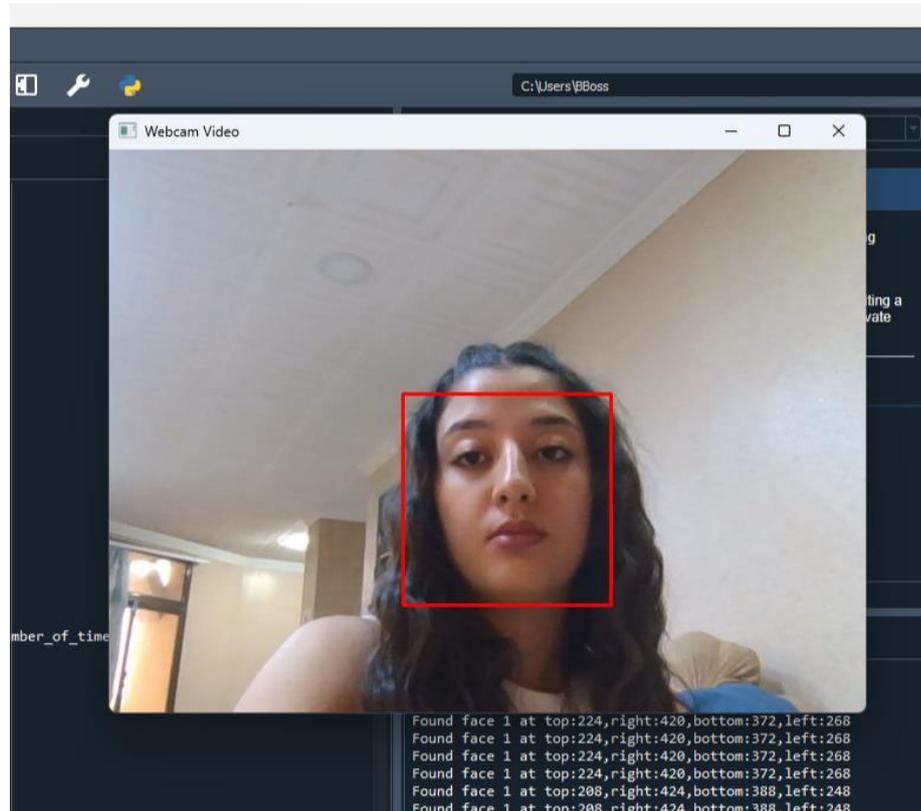


FIGURE 3.4 DÉTECTION DE VISAGE EN TEMPS RÉEL.

3-2 Détection d'expressions faciales

Après avoir détecter la localisation de visage dans l'étape précédente, dans cette étape on essaye de prédire l'emotion en utilisant le modèle pré-entraîné de Kaggle.

3-2-1 The Kaggle facial expression recognition challenge dataset

Est un modèle pré-entraîné basé sur les réseaux de neurones. Il consiste à :

- Des images de visage en niveaux de grayscale de 48x48 pixels.
- Chaque image correspond à une des sept catégories d'expression ('angry', 'disgust', 'fear', 'happy', 'sad', 'surprise', 'neutral').
- L'ensemble des données contient environ 36000 images. [16]

Après avoir l'installer, on va avoir deux fichiers :

- Fichier de structure du modèle :
'facial_expression_model_structure.json'.
- Fichier de poids du modèle :
'facial_expression_model_weight.h5'. [16]

3-2-2 Implémentation



FIGURE 3.5 DÉTECTION D'EXPRESSION FACIALE DANS UNE IMAGE STATIQUE.

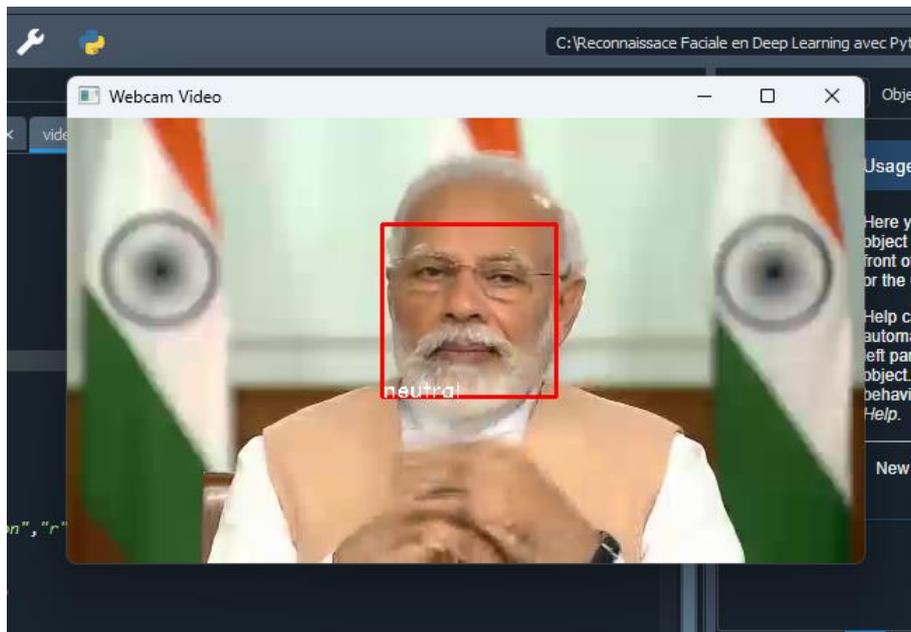


FIGURE 3.6 DÉTECTION D'EXPRESSION FACIALE DANS UNE VIDÉO PRE-ENREGISTRÉE.

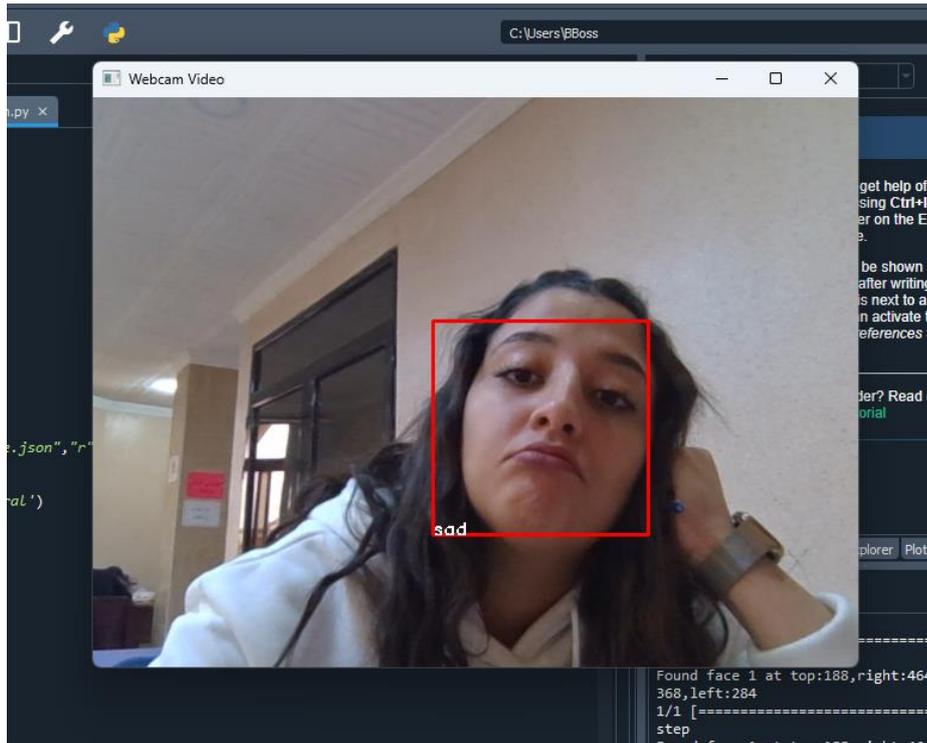


FIGURE 3.7 DÉTECTION D'EXPRESSION FACIALE EN TEMPS RÉEL.

3-3 Classification d'âge et de sexe

Lors de l'étape précédente, nous avons extrait les émotions de chaque visage dans les différents médias visuels. Dans cette étapes, on va essayer de prédire l'âge et le sexe d'un visage en utilisant le modèle pré-entraîné "Adience collection of unfiltered faces for gender and age classification".

3-3-1 Adience collection of unfiltered faces for gender and age classification

Est un modèle pré-entraîné basé sur Deep Learning et les réseaux de neurones convolutifs (CNN), crée et entraîné par **Gil Levi** et **Tal Hassner** à l'aide d'ensemble de données **Adience**. Il a :

- Nombre total de sujets : 2284.
- Nombre total de photos : 26580.
- Nombre de groupes d'âge : 8 [(0-2), (4-6), (8-13), (15-20), (25-32), (38-43), (48-53), (60....) etc].
- Classification : Male / Female. [18]

Après avoir l'installer de GitHub, on va avoir cinq fichiers :

- Age Classification Caffe Model.
- Age Classification protext file.

- Gender Classification Caffe Model.
- Gender Classification prototxt file. [18]
- The Mean Image.

Caffe : est un framework de Deep Learning.

Prototxt : est un fichier de configuration pour indiquer à Caffe comment on souhaite que le réseau soit entraîné.

3-3-2 Implémentation

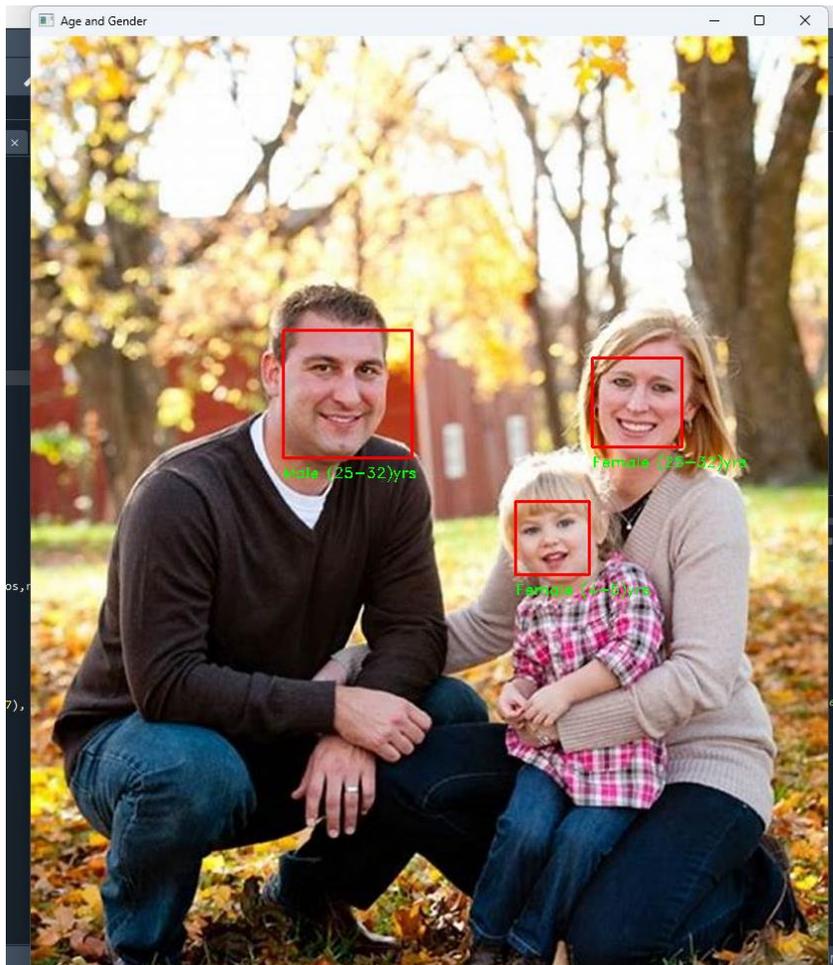


FIGURE 3.8 CLASSIFICATION D'AGE ET DE SEXE SUR UNE IMAGE STATIQUE.

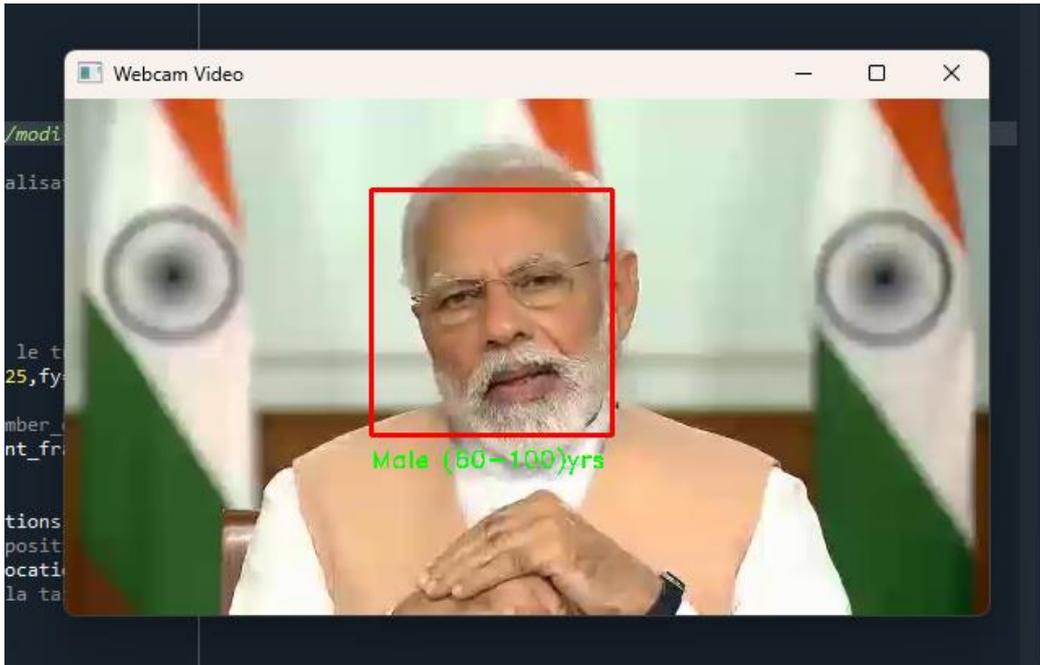


FIGURE 3.9 CLASSIFICATION D'AGE ET DE SEXE SUR UNE VIDÉO PRE-ENREGISTRÉE.

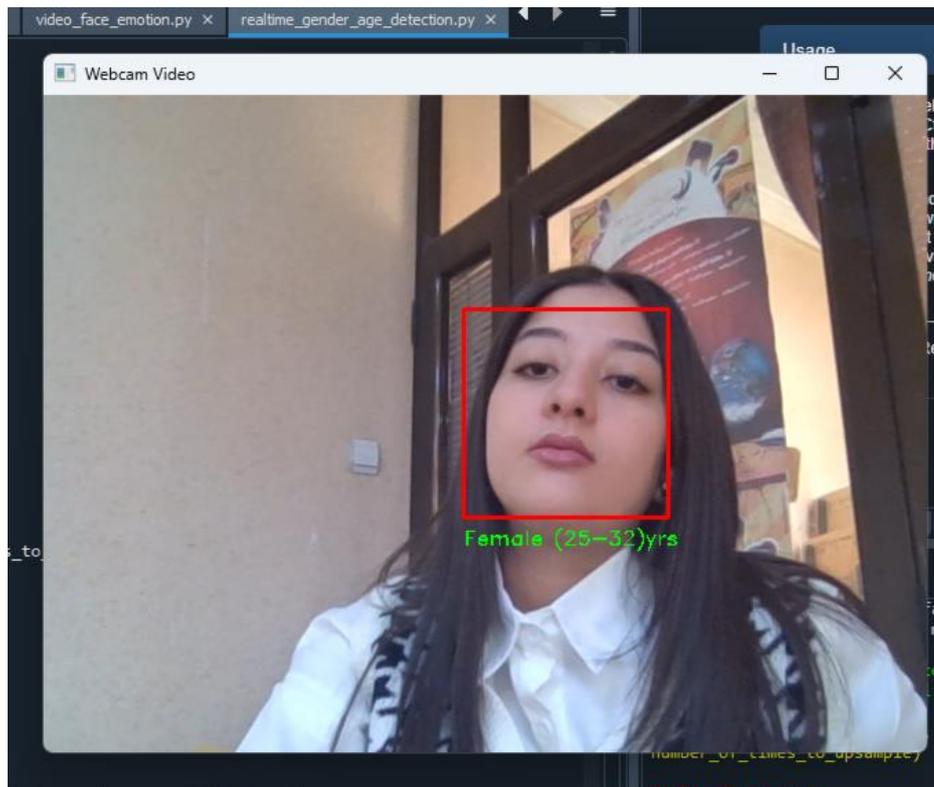


FIGURE 3.10 CLASSIFICATION D'AGE ET DE SEXE EN TEMPS RÉEL.

4- L'interface de développement

Pour rendre visible notre travail, nous avons développé une interface en Spyder pour fournir un environnement de système de reconnaissance faciale. Notre interface contient trois boutons : **Image**, **Video** et **Real-Time** (Figure 3.11).



FIGURE 3.11 L'INTERFACE DE NOTRE SYSTÈME.

- Dès qu'en clique sur le bouton "Image" ça nous donne accès aux images stockées dans le PC (Figure 3.12). Après avoir choisir l'image, ça nous donne le résultat (Figure 3.13).

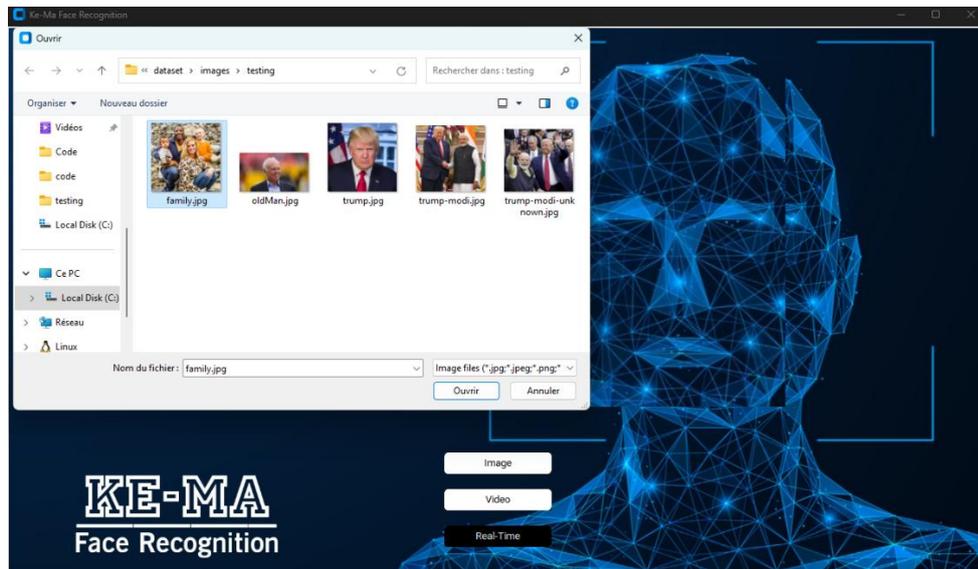


FIGURE 3.12 CHOISIR UNE IMAGE.

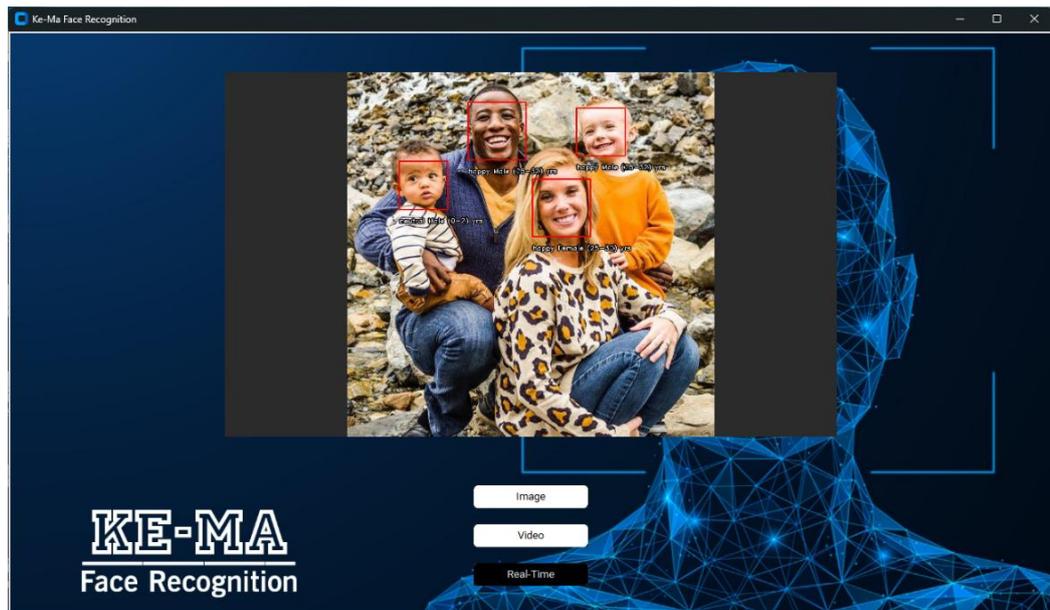


FIGURE 3.13 LE RÉSULTAT SUR UNE IMAGE STATIQUE.

- En cliquant sur le bouton "Video" ça nous donne accès aux vidéos pré-enregistrées dans le PC. Après avoir choisir la vidéo, ça nous donne le résultat (Figure 3.14).

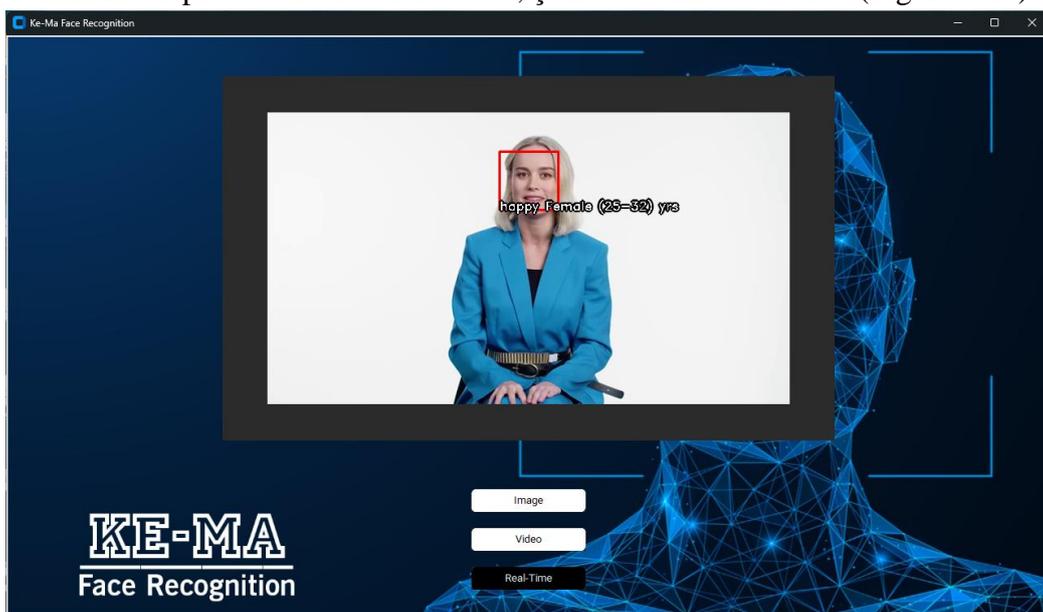


FIGURE 3.14 LE RÉSULTAT SUR UNE VIDÉO PRE-ENREGISTRÉE.

- En cliquant sur le bouton "Real-Time", ça commence directement le diffusion en direct du Web Cam du PC (Figure 3.15).

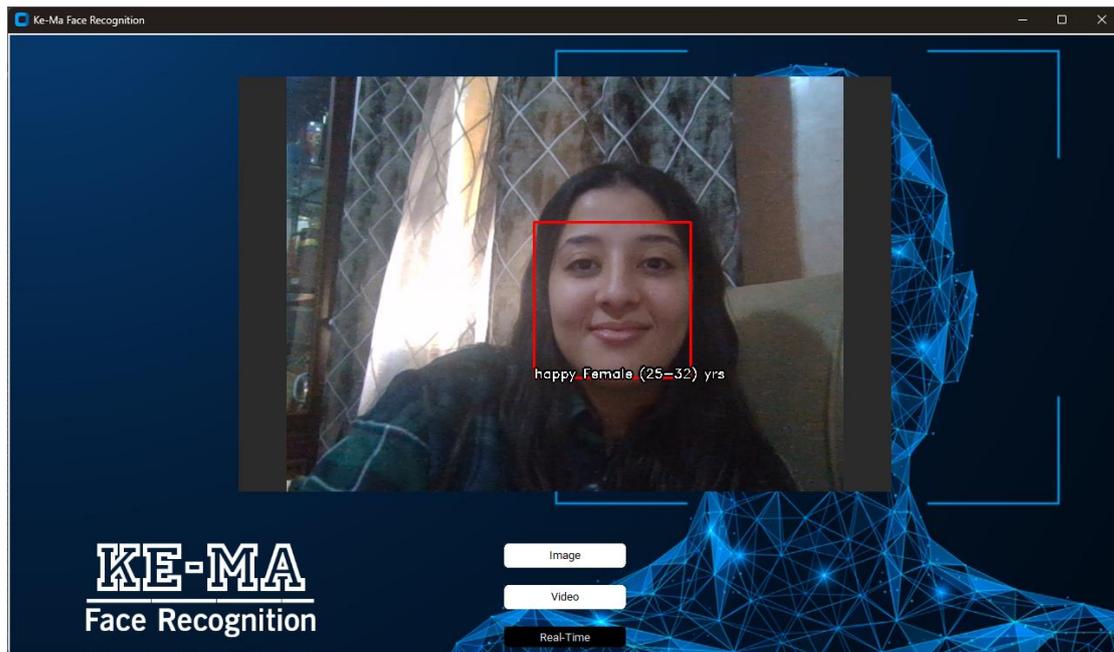


FIGURE 3.15 LE RÉSULTAT EN TEMPS RÉEL.

5- Résultat et interprétation :

Le tableau 3.1 présente les métriques de performance pour les différents classificateurs : Classificateur d'âge, du sexe et d'expressions faciales.

Métrique	Classificateur		
	D'âge	De sexe	D'expressions faciales
Précision	0,71	0,94	0,78
Rappel	0,053	0,44	0,41
Score F1	0,60	0,93	0,58
Exactitude	0,42	0,87	0,56
Taux d'erreur	57%	12%	43%

TABLEAU 3.1 LES MÉTRIQUES DE PERFORMANCE.

5-1 Précision Cela mesure la proportion de résultats positifs correctement prédits par le modèle. On peut observer que le classificateur de sexe a la plus haute précision (0,93), suivi par le classificateur d'expressions faciales (0,78) et le classificateur d'âge (0,71).

- 5-2 Rappel** Cela mesure la proportion de vrais résultats positifs identifiés par le modèle. Le classificateur de sexe a le plus haut rappel (0,44), suivi par le classificateur d'expressions faciales (0,41) et le classificateur d'âge (0,053).
- 5-3 Score F1** C'est une mesure globale qui combine à la fois la précision et le rappel. Le classificateur de sexe obtient le score F1 le plus élevé (0,93), suivi par le classificateur d'âge (0,60) et le classificateur d'expressions faciales (0,58).
- 5-4 Exactitude** Cela mesure la proportion de prédictions correctes par rapport à l'ensemble des prédictions. Le classificateur de sexe a la plus haute exactitude (0,87), suivi par le classificateur d'expressions faciales (0,56). Le classificateur d'âge a la plus faible exactitude (0,42).
- 5-5 Taux d'erreur** Il est calculé en soustrayant l'exactitude de 1. On peut observer que le classificateur de sexe a le taux d'erreur le plus faible (12%), suivi par le classificateur d'expressions faciales (43%). Le classificateur d'âge a le taux d'erreur le plus élevé (57%).

6- Conclusion

Dans ce chapitre, nous avons présenté les différentes étapes suivies pour développer et exploiter notre système de reconnaissance faciale. Nous mettons également à disposition l'interface de ce système, avec interprétations et d'explications des résultats obtenus.

Conclusion Générale

Les systèmes biométriques jouent un rôle de plus en plus important dans la société moderne. La reconnaissance faciale présente un intérêt particulier en tant que technologie biométrique de base, car elle fournit un moyen discret et non intrusif de détection, d'identification et de vérification sans que le sujet n'en ait connaissance ou ait son consentement.

L'objectif de ce travail est de concevoir et d'implémenter un système de reconnaissance faciale, capable de détecter les visages et de prédire le sexe et l'âge d'une personne dans les médias visuels (image, vidéo ou en temps réel) qui peut être utilisé dans les réseaux sociaux et les panneaux publicitaires en répondant à des exigences de rapidité et robustesse. Etc.

Ce travail a relativement bien réussi à développer un système de détection d'expression faciale, de sexe et de classification d'âge. Malgré les problèmes de poses et d'éclairage, notre système nous a donné des résultats acceptables.

L'exploitation de ce travail nous a permis de découvrir plus profondément plusieurs aspects du développement d'un système complexe. Il nous a fallu d'abord nous renseigner sur le domaine de la reconnaissance faciale, et plus généralement de la vision par ordinateur, qui est un domaine en pleine expansion.

A cet effet, si la biométrie est un enjeu important au niveau économique, alors la recherche, notamment dans le domaine de la reconnaissance faciale, offre encore un champ de recherche très ouvert pour améliorer les performances des systèmes de détection et de détection et élargir la portée de ces applications.

Enfin, en perspective, on peut envisager d'étendre ce travail en créant un système de détection et de reconnaissance de visages basé sur CNN comme extracteur de caractéristiques et une machine à vecteurs de support comme classificateur. Une autre perspective est de tester le nôtre sur d'autres bases de données faciales dotées de puissantes capacités. Changements d'éclairage et de posture.

Bibliographies

- [1] Ben mouffok mohamed Chaker, Sayah Ahmed Ridha. « Système de présence basé sur la reconnaissance faciale utilisant Linux PC ou Raspberry Pi OS ». PhD thesis, Université Mohamed Khider Biskra, 2020.
- [2] AO Kaspersky Lab, (2023). « Qu'est-ce que la reconnaissance faciale ? : Kaspersky » [site web]. Consulté Mars 2023.
- [3] Vonintsoa, (Mars 2022). « Le guide complet de la reconnaissance faciale : Intelligence-artificielle » [site web]. Consulté Avril 2023.
- [4] Souhila Guerfi ABABSA. « Authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visages 2D/3D ». PhD thesis, l'Université Evry Val d'Essonne, Octobre 2008.
- [5] W. Zhao, R. Chellappa, P.J. Phillips et A. Rosenfeld, « Face recognition : A literature survey », ACM Computing Surveys, vol. 35, 2003, p. 399-458.
- [6] D.Boukhlof, « Résolution des problèmes par écosystèmes : Application au traitement d'image ». Université de Biskra, 2005.
- [7] « Deep Learning ou Apprentissage Profond : qu'est-ce que c'est ? » : DataScientest [site web] (28 septembre 2022). <https://datascientest.com/deep-learning-definition>
- [8] Hugo Welte, (22 août 2023). « Deep learning définition, fonctionnement et applications : HubSpot » [site web]. Consulté Avril 2023. <https://blog.hubspot.fr>
- [9] Amin Riad Maouche. « Vers une commande neuronale de robots flexibles ». PhD thesis ,Université M'Hamed Bougara Boumerdes, Novembre 2010.
- [10] Ludo LOUIS. (19 Juin 2018). « Les différents types de réseaux de neurones » : Réseau de convolution [site web]. Consulté Mars 2023. <https://ludo-louis.fr/differents-reseaux-de-neurones-reseau-convolution>
- [11] Afshine Amidi et Shervine Amidi, (2019). « Pense-bête de réseaux de neurones convolutionnels » : Stanford University [site web]. Consulté Mai 2023. <https://stanford.edu>
- [12] « Convolutional Neural Network : Tout ce qu'il y a à savoir : DataScientest » [site web] (25 Juin 2023), Consulté Mai 2023. <https://datascientest.com>
- [13] D.Moualek. « Deep Learning pour la classification d'image ». PhD thesis, Université Abou Baker Belkaid Telmcen, 2017.
- [14] Entraînement d'un réseau de neurone convolutif, Publié le 17 mai 2018, Disponible sur <https://www.natural-solutions.eu>

- [15] Amrita Pathak. « Réseaux de neurones convolutifs (CNN) une introduction : GeekFlare»
[site web]. <https://geekflare.com>
- [16] <https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data>
- [17] www.fr.wikipédia
- [18] <https://github.com>