

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



Université Abderhmane Mira - Bejaia -
Faculté des Sciences Exactes
Département d'Informatique

MÉMOIRE DE MASTER RECHERCHE

En Informatique

Option : Intelligence artificielle

Thème

Reconnaissance des expressions faciales

Présenté par : M^{lle}Imene Allouache

Soutenance le 25 octobre 2020 devant le jury composé de :

Président	Dr ALOUI Soraya.	Maître Conférence A.	U. A/Mira Bejaia, Algerie.
Rapporteur	Dr AMROUN Kamal.	Maître Conférence A.	U. A/Mira Bejaia, Algerie.
Examinatrice	M SAKAAN Nadjim.	Doctorant.	U. A/Mira Bejaia, Algerie.

Promotion 2019-2020.

- *Remerciement* -

Je tiens à remercier ma famille qui m'a toujours fait confiance dans mon parcours d'étude aussi bien dans les hauts, que dans les bas. Mes amis pour leur soutien et encouragement, ainsi que toutes les personnes proches ou éloignées que j'ai pu rencontrer durant mes années d'étude.

Enfin merci à toutes les personnes qui ont contribué de près ou de loin à la réalisation de ce travail.

- *Dedicace* -

Je dédie ce travail à ma mère et ma grande mère

Table des matières

Table des matières	i
Tables des figures	iii
Liste des tableaux	iv
Introduction générale	1
Résumé	1
ABSTRACT	1
1 Généralités sur la reconnaissance des expressions faciales	2
1.1 Introduction	2
1.2 Historique	2
1.3 Qu'est-ce qu'une expression faciale?	3
1.4 Qu'est-ce qu'une émotion?	3
1.4.1 Types d'émotions	4
1.4.1.1 Émotion primaire	4
1.4.1.2 Émotion secondaire	4
1.4.1.3 Émotion social	4
1.5 Apprentissage automatique	4
1.5.1 Type d'apprentissage	5
1.5.1.1 Apprentissage Supervisé	5
1.5.1.2 Apprentissage non-supervisé	6
1.5.2 Différence entre l'apprentissage supervisé et non supervisé	6
1.5.3 Quelques algorithmes d'apprentissage	6
1.5.3.1 Machine à vecteur du support (SVM)	6
1.5.3.2 Réseau de neurone	7
1.5.3.3 K plus proches voisins (KNN)	7
1.5.3.4 Arbre de décision	8
1.6 Conclusion	8

2	Etat de l'art sur la reconnaissance des expressions faciales	9
2.1	Introduction	9
2.2	Processus de reconnaissance d'expression faciale	9
2.2.1	Détecter un visage	9
2.2.1.1	Méthode pour la détection du visage	9
2.2.2	Extraction des caractéristiques faciales	10
2.2.2.1	Méthode d'extraction des caractéristiques faciales	11
2.2.3	Reconnaissance d'émotion	12
2.2.3.1	Méthode pour la reconnaissance d'émotion	12
2.3	comparaison des performances des méthodes pour la reconnaissance des expressions faciales :	15
2.4	Conclusion	15
3	Etude comparative des deux méthodes de reconnaissance des expressions faciales sur les datasets CK+ et JAFFE	16
3.1	Introduction	16
3.2	Datasets	16
3.2.1	Datasets CK+	16
3.2.2	Dataset JAFFE	16
3.3	Méthodes de reconnaissance d'expressions faciales	17
3.3.1	Réseau de neurone	17
3.3.2	Machine à vecteur du support (SVM)	18
3.4	Critères d'évaluation des solutions existantes	18
3.5	Résultat de la comparaison	19
3.5.1	Par rapport à CK+	19
3.5.2	Par rapport à JAFEE	19
3.6	Conclusion	20
4	Mise en œuvre d'une méthode de reconnaissance d'expressions faciales	21
4.1	Introduction	21
4.2	Méthode proposée	21
4.2.1	Fonctionnement du réseau de neurone	22
4.2.1.1	Couche de convolution	22
4.2.1.2	Convolution matricielle	23
4.2.1.3	Fonction ReLU	23
4.2.1.4	Couche de POOLING	23
4.2.1.5	Fonction d'activation softmax	24
4.3	Exemple illustratif	25
4.4	Conclusion	25
	Conclusion générale	1

Bibliographie

vi

Table des figures

1.1	Librairie Scikit-Learn [11]	5
1.2	Exemple de classification KNN	7
1.3	Exemple d'un arbre de décision	8
2.1	Les 70 points d'intérêt du visage de FaceSDK [17]	11
3.1	Dataset CK+ [33]	17
3.2	Dataset JAFFE [33]	17
4.1	Architecture du réseau de neurone proposé	21
4.2	Effet de la couche de convolution	22
4.3	Illustration de l'opération de pooling.	24

Liste des tableaux

2.1	Comparaison des performances des méthodes pour la reconnaissance d'émotion . . .	15
3.1	Comparaison des performances des méthodes de la reconnaissance d'émotion pour le dataset CK+	19
3.2	Comparaison des performances des méthodes de la reconnaissance d'émotion pour le dataset JAFEE	19

Introduction Générale

Les expressions faciales du visage jouent un rôle important et fondamental dans la communication interpersonnelle et social et révèlent une grande variété d'informations telles que l'émotion. La reconnaissance automatique des expressions faciales est devenue un domaine de recherche actif au cours des dernières décennies, avec un certain nombre d'applications importantes telles que l'interaction homme machine et la surveillance visuelle. Bien que les humains détectent et analysent les visages et les expressions faciales facilement, le développement d'un système automatique pour accomplir cette tâche présente de nombreuses difficultés.

Le problème de reconnaissance d'expressions faciales a été abordé depuis bien de longues années, la question est de savoir comment les émotions sont reconnues, de nombreux chercheurs ont beaucoup travaillé sur la reconnaissance des émotions à partir d'expressions faciales en utilisant les techniques de traitement d'image et de vision par ordinateur.

Dans le premier chapitre nous avons présenté quelques notions relatives aux émotions telles que leurs historiques, leurs définitions ainsi que leurs différents types. Dans le deuxième chapitre on a fait une étude comparative entre les différentes méthodes de la reconnaissance d'émotions.

Dans le troisième chapitre nous avons effectué une étude comparative entre les deux méthodes de reconnaissance faciales, les réseaux de neurones à convolution (CNN) et les machines à vecteur du support (SVM) pour la reconnaissance d'expressions faciales sur les datasets Cohn-knake (CK+) et Japanese Female Facial Expressions (JAFFE).

Dans le quatrième chapitre nous avons proposé un réseau de neurone qu'on va appeler RNREF (Réseau de neurone pour la reconnaissance d'expressions faciales) composé de six couches pour la reconnaissance des sept émotions suivantes (colère, dégoût, peur, bonheur, neutre, tristesse, surprise).

Résumé

La reconnaissance des expressions faciales devient un sujet brûlant en raison de ses vastes applications dans les domaines de la recherche en vision par ordinateur. Les méthodes traditionnelles adoptent des fonctionnalités artisanales combinées à des classificateurs pour atteindre l'objectif de reconnaissance. Cependant, la précision de ces méthodes dépend souvent fortement des caractéristiques extraites et des paramètres du classificateur, et ne peut donc pas obtenir de bons résultats avec des données invisibles. Récemment, l'apprentissage en profondeur, qui simule le mécanisme du cerveau humain pour interpréter les données, a montré des résultats remarquables dans la reconnaissance visuelle d'objets. Dans cet article, nous présentons un nouveau réseau de neurones convolutifs, composé de six couches pour reconnaître les sept expressions faciales suivantes (colère, dégoût, peur, bonheur, neutre, tristesse, surprise).

ABSTRACT

Facial expression recognition is becoming a hot topic due to its wide applications in computer vision research fields. Traditional methods adopt handcrafted features combined with classifiers to achieve the recognition goal. However, the accuracy of these methods often relies heavily on the extracted features and the classifier's parameters, and thus cannot get good result with unseen data. Recently, deep learning, which simulates the mechanism of human brain to interpret data, has shown remarkable results in visual object recognition. In this paper, we present a novel convolutional neural network composed of six layers to recognize the following seven facial expressions (anger, disgust, fear, happiness, neutral, sadness, surprise).

Chapitre 1

Généralités sur la reconnaissance des expressions faciales

1.1 Introduction

Au cours de la dernière décennie, la communauté de la recherche en vision par ordinateur a montré beaucoup d'intérêt pour l'analyse et la reconnaissance automatique des expressions faciales. Cette communauté de la vision par ordinateur et de la recherche scientifique envisageait de développer des systèmes capable de reconnaître les expressions faciales dans des vidéo ou de images, La plus plupart de ces systèmes d'analyse des expressions faciales tentent de classer les expressions en quelques grandes catégories émotionnelles, telles que la joie, la tristesse ; la colère, la surprise, la peur et le dégoût.

Dans ce chapitre, nous allons présenter quelques notions relatives aux émotions telles que leurs historiques, leurs définition ainsi que leurs différents types.

1.2 Historique

Le premier système semi-automatisé a été développé en 1960, il consiste à détecter les oreilles, les yeux, le nez et la bouche sur une image et de calculer les distance à un point de référence commun ensuite comparer ces derniers aux données de référence, tout cela a été fait manuellement par un administrateur [1]

Dans les années 1970, Goldstein et al [2] ont utilisé l'épaisseur des lèvres et la couleur des cheveux pour automatiser la reconnaissance d'un visage dans une image, cette méthode est comme la précédente, elle est fait manuellement.

En 1988, Kirby et Sirovich [3] ont appliqué au problème de la reconnaissance des visages, l'analyse des composantes principales qui est une technique standard d'algèbre linéaire, le codage se

fait sur une image correctement alignée et normalisée.

En 1991, Turk et Pentland [4] ont découvert que grâce à la technique Eigen-faces, l'apprentissage en temps réel de nouveaux visages dans une image de manière non supervisée est possible.

En janvier 2001, lors de la finale du championnat de football américain la technologie de la reconnaissance faciale automatique a été mise en essai en comparant les images capturées de la caméra de surveillance à une base de donnée numérique [5].

Ces méthodes vues avant les années 2000 consistent uniquement à détecter un visage dans une image afin de reconnaître une personne, quant aux méthodes suivantes, la notion d'émotion est introduite.

En 2018, Boughida et al [6] ont utilisé l'algorithme de classification K-Nearest Neighbors(KNN) qui signifie les k plus proches voisins ; ils ont sélectionné 18 points d'intérêt d'un visage afin de caractériser les sept expressions faciales universelles (Joie, Peur, Dégoût, Surprise, Tristesse, Colère, Neutre) définies par Ekman [7].

Aujourd'hui, la technologie de la reconnaissance faciale est utilisée dans les endroits publics pour soutenir les forces de l'ordre en identifiant des personnes qui pourront nuire, lutter contre la fraude des passeports, identifier les personnes disparues, etc...

1.3 Qu'est-ce qu'une expression faciale ?

Le visage porte des informations sur l'identité d'une personne telle que la couleur des yeux, la forme de la bouche, la couleur des cheveux ainsi que la forme des oreilles, etc. Comme il peut exprimer aussi des expressions de communication et d'émotion qu'on peut reconnaître en prêtant attention aux expressions faciales.

L'expression faciale est avant tout un ensemble des signes du visage, perceptible visuellement, dû à l'activation (volontaire ou non) à l'un ou plusieurs des 44 muscles qui composent le visage, qui traduisent un sentiment de changement dans le visage [5]

1.4 Qu'est-ce qu'une émotion ?

L'émotion est l'état mental d'une personne qui dépend de son humeur, elle peut changer à cause d'un événement qui se produit ou d'une situation particulière.

Une émotion peut rester intérieure à une personne et rester limitée à son expérience intime, mais même dans ce cas l'envie de passer à l'action est toujours présente, que ce soit par un ressenti ou à

travers l'imagination [8]

1.4.1 Types d'émotions

Les émotions sont des séquences courante qui interviennent tout au long de nos journées, on les reconnaît grâce à des expressions faciales, à la tonalité de la voix ou même a des gestes corporelle, il existe trois types d'émotions : émotion primaire, émotion secondaire et émotion sociale.

1.4.1.1 Émotion primaire

Les émotions primaires (joie, tristesse, colère, peur, dégoût, surprise) sont engendrées par des évènements quotidiens et elles sont à la base des réactions humaines qui cause des comportements spécifiques [9].

Les deux types d'émotions suivants sont appris et constitué à partir des émotions primaires.

1.4.1.2 Émotion secondaire

Les émotions secondaires sont les plus difficiles à identifier, car elles sont souvent le résultat d'une combinaison de deux émotions primaires, elles sont à l'origine d'un souvenir évoqué et arrivent à maturation à l'âge adulte [9].

1.4.1.3 Émotion social

Les émotions sociales (culpabilité, honte, jalousie, timidité, humiliation, etc.) sont acquises en fonction de l'éducation et la culture d'une personne, elles nous permettent de nous adapter aux autres afin de pouvoir vivre en société [9].

1.5 Apprentissage automatique

Le Machine Learning ou apprentissage automatique en français est un concept qui fait de plus en plus parler de lui dans le monde de l'informatique et qui se rapporte au domaine de l'intelligence artificielle, il s'agit d'une sorte de programme permettant à l'ordinateur ou à une machine un apprentissage automatisé, en lui donnant la capacité d'apprendre plutôt de le programmer d'une façon explicite pour pouvoir réaliser un certain nombre d'opérations très complexe, L'objectif visé et de rendre la machine ou l'ordinateur capable d'apporter des solutions à des problèmes compliqués, par le traitement astronomique d'information. Cela offre ainsi une possibilité d'analyser et de mettre en évidence les corrélations qui existent entre deux ou plusieurs situations données et de prédire leurs différentes implications. [10]

1.5.1 Type d'apprentissage

Le machine Learning implique trois systèmes d'apprentissage qui définissent ses différents modes de fonctionnement, il s'agit d'apprentissage supervisé, apprentissage non-supervisé et apprentissage par renforcement, nous on va se focalisé sur l'apprentissage supervisé et non-supervisé puisque c'est les plus utilisé.

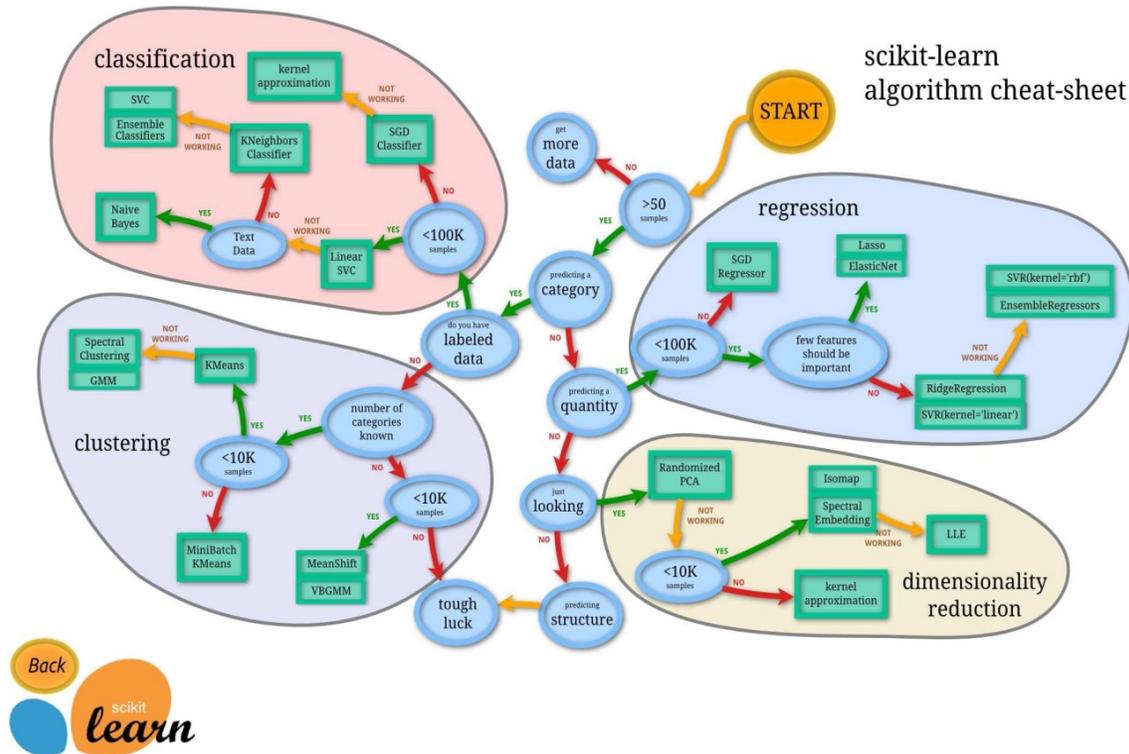


FIGURE 1.1: Librairie Scikit-Learn [11]

1.5.1.1 Apprentissage Supervisé

Dans le mode apprentissage supervisé, la machine s'appuie sur des classes prédéterminées et sur un certain nombre de paradigme connus pour mettre en place un système de classement à partir de modèles déjà catalogué. Dans ce cas, deux étapes sont nécessaires pour compléter le processus, à commencer par le stade d'apprentissage qui consiste à la modélisation des données catalogués, Ensuite, il s'agira au second stade de se basé sur les données ainsi définies pour attribuer des classes aux nouveaux modèles introduits dans le système, afin de les cataloguer eux aussi [12]. Ce problème peut être soit :

- **Classification** : Les échantillons appartiennent à deux classes ou plus et nous voulons apprendre à partir de données déjà étiquetées comment prédire la classe des données non étiquetées. Un exemple de problèmes de classification serait l'exemple de reconnaissance de chiffres manuscrits, dans lequel le but est d'attribuer chaque vecteur d'entrée à l'un des nombre finis de catégories discrètes. Une autre façon de penser à la classification est une

forme discrète (par opposition à continue) d'apprentissage supervisé ou l'on a un nombre limité de catégories et pour chaque des n échantillons fournis, on essaie de les étiqueter avec la catégorie ou la classe correcte.

- **Régression** : Si la sortie désirée consiste en une ou plusieurs variables continues, alors la tâche est appelée régression, par exemple prédire le prix d'un appartement.

1.5.1.2 Apprentissage non-supervisé

Dans ce mode de fonctionnement du Machine Learning, il n'est pas question de s'appuyer sur de éléments prédéfinis, la tâche revient à la machine de procéder toute seule à la catégorisation des données. Pour ce faire, le système va croiser les informations qui lui sont soumises de manière à pouvoir assembler dans une même classe les éléments présentent certaines similitudes. Ainsi en fonction du but recherché, il reviendra à l'opérateur ou au chercheur de les analyser afin d'en déduire les différentes hypothèses.[12]

1.5.2 Différence entre l'apprentissage supervisé et non supervisé

La différence entre ces deux principes de fonctionnement réside dans le fait que l'apprentissage supervisé peut être influencé par l'étiquetage des données, ce n'est pas le cas de l'apprentissage non-supervisé qui se révèle ainsi beaucoup plus fiable dans la mesure où les réponses obtenus vont plus loin que la compréhension humaine des faits. Par ailleurs, il faut noter que le machine Learning peut également faire intervenir un mode de fonctionnement mixte qui utilise les deux type d'apprentissage pour arriver à des résultats plus précis.

1.5.3 Quelques algorithmes d'apprentissage

1.5.3.1 Machine à vecteur du support (SVM)

L'algorithme Machine à vecteurs du support (SVM) est un ensemble de méthodes d'apprentissage supervisé utilisées pour la classification, la régression et la détection des valeurs aberrantes, il est capable d'effectuer une classification multi-classes sur un ensemble de données qui peut être utilisé pour la classification, la régression et d'autres tâches. Intuitivement, une bonne séparation est obtenu par l'hyper-plan qui a la plus grande distance des points de formation les plus proches de n'importe quelle classe, car plus la marge est grande, plus l'erreur de généralisation du classificateur est faible

●Avantage des machines vectorielles du support

- Efficace dans les espaces de grande dimension.
- Toujours efficace dans le cas où le nombre de dimensions est supérieur au nombre d'échantillons.
- Utilise un sous ensemble de points d'apprentissage dans la fonction de décision, donc il est également efficace en mémoire.

● Inconvénients des machines vectorielles du support

- Les SVM ne fournissent pas directement d'estimation de probabilité, celles-ci sont calculées à l'aide d'une validation croisée.

1.5.3.2 Réseau de neurone

Les réseaux de neurones [13] sont des modèles de calcul qui date des années 40, c'est une technique inspiré des réseaux de neurones biologiques pour exécuter des tâches calculatoires, elle a la particularité de s'adapter, d'apprendre, de généraliser pour classer des données en entrée, un réseau de neurone s'inspire du fonctionnement des neurones biologiques et prend corps dans un ordinateur sous forme d'un algorithme, il peut se modifier lui-même en fonction des résultat de ses actions, ce qui permet l'apprentissage et la résolution de problèmes sans algorithme, donc sans programmation classique.

1.5.3.3 K plus proches voisins (KNN)

L'algorithme K-Nearest Neighbors (KNN) [14] est l'un des algorithmes de classification les plus simple et l'un des algorithmes d'apprentissage les plus utilisés, KNN est un algorithme d'apprentissage paresseux. Son but est d'utiliser une base de données dans laquelle les points de données sont séparés en plusieurs classe pour prédire la classification d'un nouveau point d'échantillonnage Une technique est non paramétrique, cela signifie qu'elle ne fait aucune hypothèse sur la distribution de données, en d'autre termes la structure du modèle est déterminé à partir des données. L'algorithme KNN est basé sur la similarité des caractéristiques, la mesure dans laquelle les caractéristiques hors échantillon ressemblent à l'ensemble d'apprentissage détermine la manière dont nous classons un point de données donné.

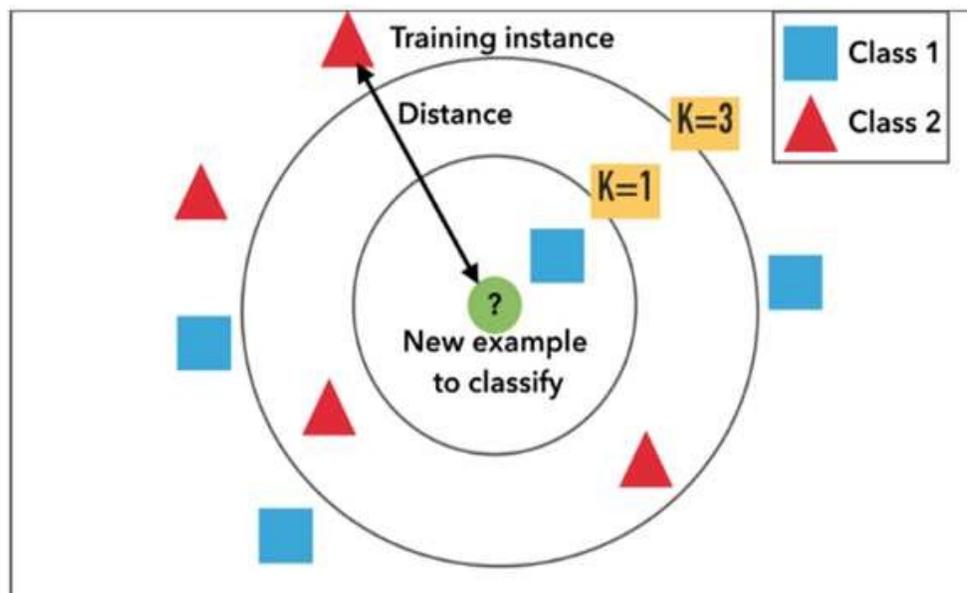


FIGURE 1.2: Exemple de classification KNN

1.5.3.4 Arbre de décision

C'est un algorithme de Machine Learning qui permet d'effectuer automatiquement les tâches de classification et de régression, il peut s'adapter à des données complexes pour résoudre des problèmes linéaire et non-linéaire, C'est une catégorie d'arbre utilisé dans l'exploration de données en informatique décisionnelle, ils emploient une représentation hiérarchique de la structure des données sous forme des séquences de décisions en vue de la prédiction d'un résultat ou d'une classe. Chaque individu(ou observation) qui doit être attribué à une classe est décrite par un ensemble de variables qui sont testées dans les nœuds des arbres, les tests s'effectuent dans les nœuds internes et les décisions sont prise dans les nœuds feuille.

Exemple d'arbre de décision, Accorder ou non un prêt bancaire. Chaque individu est évalué sur un ensemble de variables testées dans les nœuds internes et les décisions sont prises dans les feuilles

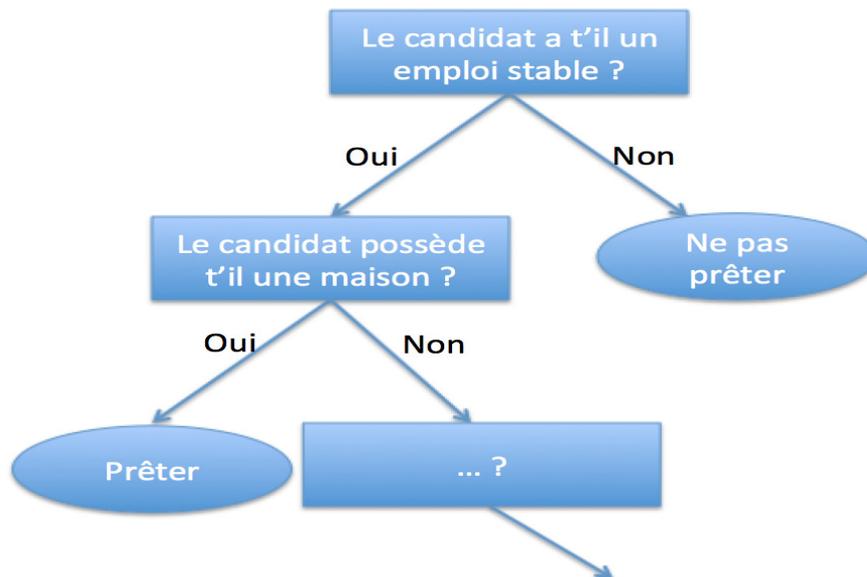


FIGURE 1.3: Exemple d'un arbre de décision

1.6 Conclusion

Dans ce chapitre nous avons d'abord présenté un aperçu général du développement dans ce domaine, ensuite nous avons évoqué le contexte d'expression faciale, émotions et apprentissage automatique, nous avons aussi décrit brièvement quelques algorithmes d'apprentissage. Dans l'étude qui suit nous donnons une comparaison synthétique des systèmes existants de reconnaissance d'expressions faciales.

Chapitre 2

Etat de l'art sur la reconnaissance des expressions faciales

2.1 Introduction

Ce chapitre est consacré à l'état de l'art sur les travaux déjà fait concernant la reconnaissance des émotions dans le domaine de l'intelligence artificielle. Enfin nous terminons le chapitre par une étude comparative entre les différentes méthodes de la reconnaissance.

2.2 Processus de reconnaissance d'expression faciale

Le processus de reconnaissance d'expression faciale d'un visage est constitué de trois étapes principales :

2.2.1 Détecter un visage

La première étape d'un système d'analyse d'expressions faciales entièrement automatique consiste à localiser la région du visage et ses limites. Le but de la détection du visage est de déterminer si un visage est présent ou non sur l'image et de localiser son emplacement.

2.2.1.1 Méthode pour la détection du visage

Kar et al [?] ont utilisés l'algorithme de Viola et Jones pour la détection du visage, d'abord les images sont converties en une image en niveau de gris pour extraire ensuite les caractéristiques des images.

Farfadi et al [24] ont proposé une méthode appelée Deep Dense Face Detector (DDFD) qui consiste à détecter un visage dans différentes orientations on utilise un modèle unique basé sur les réseaux de neurone convolutionnels profonds, similaires à AlexNet, qui sont composés de 8 couches

ou les 5 première couches sont convolutives et les 3 derniers couches sont entièrement connectées entre elles.

2.2.2 Extraction des caractéristiques faciales

Une fois que l'enregistrement du visage effectué, l'étape suivante consiste à extraire les changements faciaux causés par une expression faciale. Les expressions faciales sont définies principalement par la contraction des muscles faciaux qui produisent des changements dans l'apparence et la forme du visage. Cette étape est importante pour la reconnaissance d'une expression d'un visage, On peut distinguer deux opérations d'extraction différentes :

La première est souvent utilisée avec une approche globale de reconnaissance de visage, elle se repose sur une extraction entière des régions du visage [16]

La deuxième est implémente avec une approche locale, elle consiste à extraire quelques points des régions particulières du visage tels que la bouche, les coins des yeux et le nez. [16]

Dans l'outil FaceSDK [17], il existe 70 caractéristiques du visage que plusieurs auteurs ont utilisé dans leurs travaux, la figure suivante illustre les 70 caractéristiques du visage.

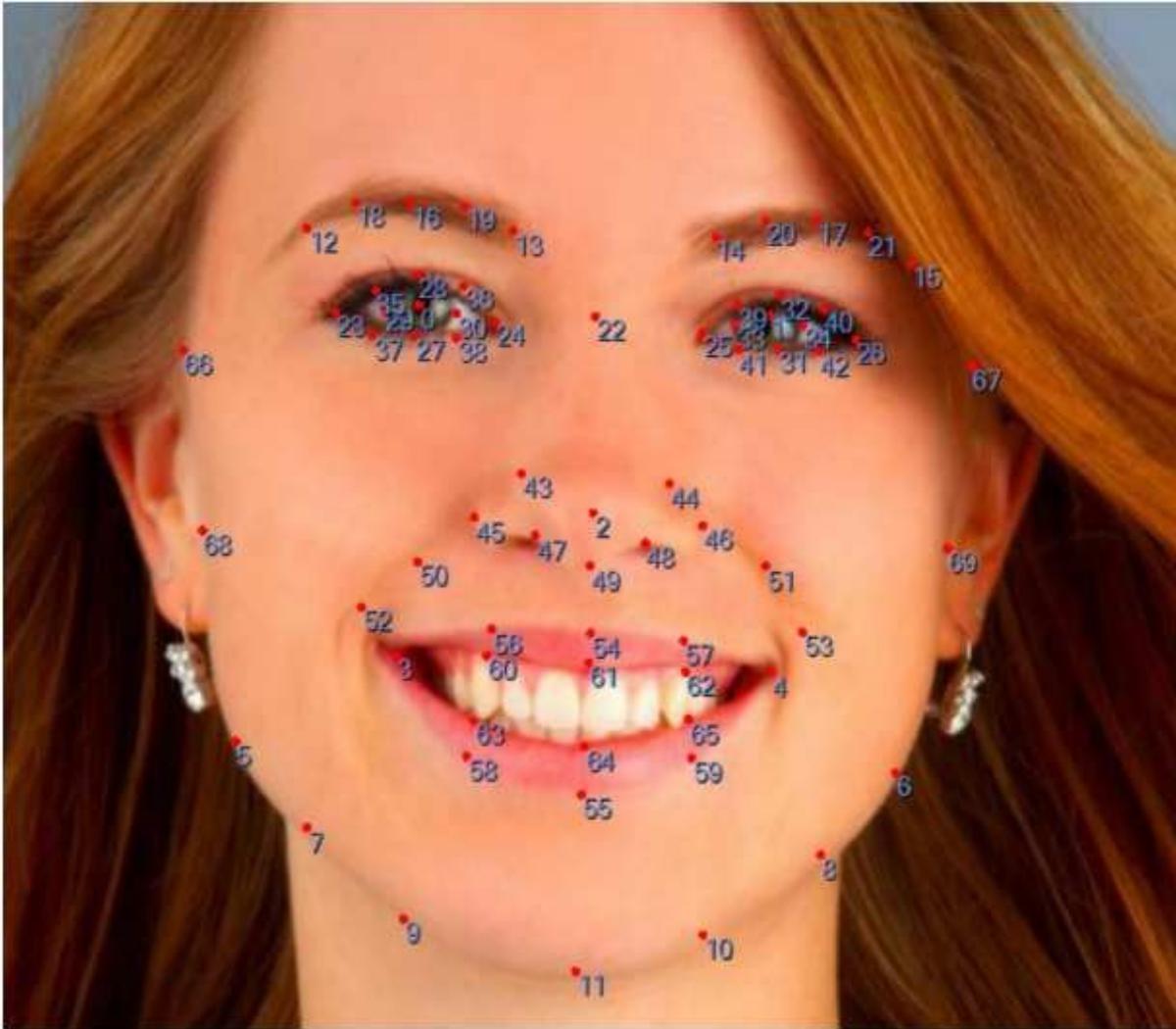


FIGURE 2.1: Les 70 points d'intérêt du visage de FaceSDK [17]

2.2.2.1 Méthode d'extraction des caractéristiques faciales

Farajzadeh et al [?] ont proposés d'utiliser la méthode de l'histogramme de gradients orientés (HOG) pour extraire les caractéristiques les plus informative des images d'expression faciales. cette méthode utilise des dégradé locaux pour décrire la forme d'un objet, ensuite les gradients horizontaux et verticaux d'une image donnée sont calculés afin de capturer l'information spatiale des gradients orientés et de rendre la méthode robuste contre les bruits, l'image est divisé en petites régions, appelées cellules. pour chaque cellule, un histogramme des orientations de gradients pondéré par la magnitude du gradient de ses pixels est calculée, comme les vecteurs des histogrammes sont basées sur les amplitudes des gradients, toute variation de l'éclairage affectera directement les histogramme. par conséquent, une autre étape appelée normalisation des blocs est également appliquée aux histogramme pour rendre caractéristiques invariantes aux variations d'éclairage. pour former le vecteur caractéristique final, les histogramme normalisées sont concaténés.

Hammal et al [18] ont proposé des modèles paramétriques spécifique associés à une méthode de segmentation afin d'extraire automatiquement les caractéristiques des trois régions du visage tel que les yeux, la bouche et le nez. Pour vérifier l'efficacité de cette méthode, ils ont effectué des tests sur une base de séquences vidéo acquises au laboratoire, ils ont ensuite utilisé les caractéristiques acquises pour une reconnaissance de haut niveau des émotions.

Kar et al [?] ont utilisés la transformation en ondulation de type2 (Ripplet2) pour extraire les caractéristiques faciales, la transformation Ripplet2 est basée sur la transformation généralisée du radon (GRT), le GRT convertit les courbes en points ce qui crée des pics situés sur la courbe correspondante, après il la transforme en ondelette (WT) pour régler les singularités ponctuelle en un domaine aléatoire généralisé.

2.2.3 Reconnaissance d'émotion

La dernière étape d'un système automatique d'analyse d'expression est la reconnaissance d'expression faciale en fonction des caractéristiques extraites.

2.2.3.1 Méthode pour la reconnaissance d'émotion

Li et al [21] ont proposés un nouveau réseau de neurone à convolution profonde inspiré par l'architecture Inception-ResNet qui possède une fonction double entrée pour reconnaître automatiquement sept expressions faciales, ils ont choisi les couche Inception-v4 et ils alimenté les images d'entrée avec la taille 299*299*1 dans le module «Stem»..... Dans le réseau neuronal proposé, ils ont incorporé les fonctionnalités des modèles binaires locaux (LBP) pendant l'étape de formation, ce qui fournit des informations sur la texture des images et reflètent donc de légers changements sur le visage, ce qui aide le réseau à prêter attention à la caractéristique faciale afin d'améliorer la précision de reconnaissance.

Au-delà ils ont utilisé la normalisation par lots, ce qui permet au réseau de se concentrer sur des images plus réelles en les déformant moins, et pour garantir aussi la distribution stable des donné d'entrée à chaque couche, enfin ils ont utilisé l'augmentation de données pour éviter le sur-ajustement. Ils ont testés leur méthode sur trois ensemble de données CK+, JAFFE et FER2013 .

Jain et al [19] ont présentés leur approche qui consiste à classer chaque image dans l'une des six classes d'émotions faciales. Le modèle proposé est basé sur des réseaux de neurones à convolution profonde unique (DNN), qui contient six couches de convolution et deux blocs d'apprentissage résiduel profond, Les couches de convolution extraient des caractéristiques hiérarchiquement, et la couche entièrement connectée et la couche softmax sont utilisés pour indiquer les 6 classes d'expression. L'entrée du réseau est 128*96*k pour tous les patches, ou k=3 pour les patches de couleur et k=1 pour les tâches grises, la sortie finale est l'une des six expressions des classes, juste après chaque couche de convolution il y a une couche de regroupement maxi, Les blocs résiduels profond

mis en place après les 2eme et 4eme couche de convolution, il existe également 2 couches entièrement connectée chacune avec une fonction d'activation ReLU et un décrochage pour la formation, enfin ils ont utilisé Softmax pour la classification. Chaque bloc résiduel a quatre couches de convolution deux connexion courtes et une connexion par saut, la première couche de convolution en résiduel a une taille de $1*1*64$, le second a une taille $3*3*64$, la troisième a une taille de $3*3*128$ et le dernier a une taille de $1*1*256$. Leur modèle a été formé sur deux ensembles de données Extended Cohn – Kanade (CK +) et ensemble de données sur l'expression du visage des femmes japonaises (JAFFE).

Neha Jain et al [20] ont utilisé la méthode des réseaux de neurones pour la connaissance des émotions, pour cela ils ont proposés un modèle hybride entre un réseau de neurone à convolution (CNN) et un réseau de neurone récurrent (RNN), le modèle CNN est utilisé pour l'extraction d'entités pour fixer tous ses paramètres et éliminer la couche de régression et le réseau de neurone récurrent (RNN) classe l'émotion faciale, Lorsque l'image est passé dans le réseau, 200 vecteurs dimensionnels seront extraits des couches entièrement connectées, pour un temps supposé t, ils prennent P trames du passé, chaque image passe ensuite de l'instant t-P a t au CNN et extraire complètement P vecteurs pour chaque image, chaque vecteur passe par un nœud du modèle RNN, ensuite chaque nœud du RNN renvoie certains résultat de l'étiquette de valence leur modèle a été formé sur deux ensembles de données (MMI) + et ensemble de données sur l'expression du visage des femmes japonaises (JAFFE).

Farajzadeh et al [22] ont appliqués la méthode de machine à vecteur de support (SVM) pour la reconnaissance de 7 émotions (colère, dégoût, peur, bonheur, neutre, tristesse, surprise), ils ont utilisés la méthode d'histogramme de gradient orienté (HOG) pour extraire les caractéristiques les plus informatives des images d'expression faciale. Dans leur méthode proposée, ils ont utilisés l'algorithme SVM pour construire les modèles, ils ont formé des classificateurs séparé pour chaque émotion perçue et les recueillir, une fois qu'un nouvel échantillon est obtenu, le score de chaque modèle est calculé, et enfin celui avec le score le plus élevé est choisi pour être l'émotion liée à l'expression. Dans un premier temps ils ont collecté des échantillons de différentes émotions (T) et après les avoir prétraités, une phase d'extraction des caractéristique est effectuée pour extraire la fonction souhaité (F), plus tard ils ont utilisé chaque exemplaire comme échantillon positif (N+) et certains comme échantillons négatifs (N-), ils ont construit un ensemble de modèles simple (un pool d'exemplaire P), En fin de compte une procédure d'étalonnage est appliqué pour rendre la sortie des modèles compatible. A des fins du test, une image d'expression faciale donnée est prétraitée et ses caractéristiques sont extraites en premier, Ensuite les exemplaire (P) sont alimentés avec la fonction (F), enfin l'exemple avec le score le plus élevé set trouvé et son étiquette correspondante est affecté à l'image d'expression faciale donnée. Ils ont effectués plusieurs tests sur 5 jeux de données, en savoir Cohen-Kanade(CK), Cohen – Kanade (CK+), JAFFE , TFEID et MMI.

Kar et al [23] ont utilisés la méthode des moindres carrés de la machine à vecteur du support (LS-SVM) comme classifieur dans leur travail, dans (LS-SVM) les solutions sont sorties en résolvant un ensemble d'équations linéaires au lieu d'équation quadratiques comme dans les SVM classiques. Dans un premier temps, les images de visage sont converties en une image en niveau de gris, ensuite le contraste de l'image a été ajustée de sorte que 1% des informations soient immergées à des intensités faibles et élevées, par la suite le visage est détecté en utilisant l'algorithme de détection de visage Viola et Jones populaire [15]. La région du visage détectée est recadrée à partir de l'image du visage d'origine et le visage recadré est remodelé à la taille 128*128. Ils ont utilisés la transformation Ripplet2 pour extraire les caractéristiques des images, le nombre total de fonctionnalités extraite par le coefficient d'ondulation est de $128 \times 128 = 16384$ ce qui est assez important pour tout classificateur existant, c'est pour ça ils ont appliqué l'approche par analyse discriminante linéaire (ADL) + le principe de l'analyse par composante principale (ACP) pour réduire les fonctionnalités a seulement 6 pour les deux ensembles de données Cohn-Kanade étendue (CK+) et l'expression faciale féminine japonaise (JAFFE).

2.3 comparaison des performances des méthodes pour la reconnaissance des expressions faciales :

Source	Année	Base de données	Méthode	Taux de reconnaissance
Deepak et al	2019	JAFFE et CK+	Réseaux de neurone à convolution profonde (DNN)	97.91 % (JAFFE) 94.98 % (CK+)
Lee et al	2019	JAFFE, CK+ et FER2013	Réseau de neurone à convolution (CNN)	98.92% (JAFFE) 97.34 % (CK+) 67.71 % (FER2013)
Farajzadeh et al	2018	JAFFE, CK+, TFEID et MMI	La méthode de machine à vecteur du support (SVM)	94.64 % (JAFFE) 98.50 % (CK+) 98.90 % (TFEID) 83.12 % (MMI)
Kar et al	2019	JAFFE et CK+	La méthode des moindres carrés de la machine à vecteur du support (LS-SVM)	98.97 % (CK+) 99.46 % (JAFFE)
Jain et al	2018	JAFFE et MMI	Modèle hybride CNN et RNN	94.91 % (JAFFE) 92.07 % (MMI)

TABLE 2.1: Comparaison des performances des méthodes pour la reconnaissance d'émotion

2.4 Conclusion

Dans ce chapitre nous avons établi un état de l'art sur les travaux déjà fait concernant la reconnaissance des émotions faciales dans le domaine d'intelligence artificielle. Pour ce fait, nous avons étudié quelques méthodes de la détection des expressions faciales, extraction des caractéristiques ainsi que la reconnaissance des émotions, finalement on a abouti à une comparaison entre les différentes méthodes pour la reconnaissance des émotions.

Dans l'étude qui suit, nous allons présenter une étude comparative des deux méthodes de reconnaissance des expressions faciales (Réseaux de neurone a convolution , machine à vecteur du support(svm)) sur les dataset CK+ et GAFFE.

Chapitre 3

Etude comparative des deux méthodes de reconnaissance des expressions faciales sur les datasets CK+ et JAFFE

3.1 Introduction

Dans ce chapitre nous allons effectuer une étude comparative entre les deux méthodes de reconnaissance faciale, les réseaux de neurone à convolution (CNN) et les machines à vecteur du support (SVM) pour la reconnaissance d'expression faciale qui utilisent les datasets Cohn-Kanade (CK+) et Japanese Female Facial Expressions (JAFEE). On s'intéresse au taux de reconnaissance et à la méthode elle-même pour pouvoir en déduire quelle méthode est la plus efficace avec les deux datasets séparément.

3.2 Datasets

3.2.1 Datasets CK+

C'est une version étendue de la base de données Cohn-Kanade (CK), introduite en 2010, se compose de 593 séquences d'expressions faciales de 123 sujets mais les étiquettes des émotions n'ont été attribuées qu'à 327 séquences qui ont été jugées répondre aux critères de l'une des sept émotions primaires (colère, dégoût, méprise, peur, tristesse, bonheur et surprise). Chacune des séquences contient des images du début, qui représentent l'image neutre, à l'expression maximale qui est la dernière image.

3.2.2 Dataset JAFEE

C'est une base de données qui contient 213 images de sept expressions faciales (6 expressions faciales de base + 1 neutre) posées par 10 modèles féminines japonais, dans ce dataset chaque sujet possède au moins 3 à 4 images d'expression de chaque étiquette d'émotion.



FIGURE 3.1: Dataset CK+ [33]



FIGURE 3.2: Dataset JAFFE [33]

3.3 Méthodes de reconnaissance d'expressions faciales

3.3.1 Réseau de neurone

Les réseaux de neurones [13] sont des modèles de calcul qui date des années 40, c'est une technique inspiré des réseaux de neurones biologiques pour exécuter des tâches calculatoires, elle a la particularité de s'adapter, d'apprendre, de généraliser pour classer des données en entrée, un

réseau de neurone s'inspire du fonctionnement des neurones biologiques et prend corps dans un ordinateur sous forme d'un algorithme, il peut se modifier lui-même en fonction des résultats de ses actions, ce qui permet l'apprentissage et la résolution de problèmes sans algorithme, donc sans programmation classique.

3.3.2 Machine à vecteur du support (SVM)

L'algorithme Machine à vecteurs du support (SVM) est un ensemble de méthodes d'apprentissage supervisé utilisées pour la classification, la régression et la détection des valeurs aberrantes, il est capable d'effectuer une classification multi-classes sur un ensemble de données qui peut être utilisé pour la classification, la régression et d'autres tâches. Intuitivement, une bonne séparation est obtenue par l'hyper-plan qui a la plus grande distance des points de formation les plus proches de n'importe quelle classe, car plus la marge est grande, plus l'erreur de généralisation du classificateur est faible.

●Avantage des machines vectorielles du support

- Efficace dans les espaces de grande dimension.
- Toujours efficace dans le cas où le nombre de dimensions est supérieur au nombre d'échantillons.
- Utilise un sous ensemble de points d'apprentissage dans la fonction de décision, donc il est également efficace en mémoire.

●Inconvénients des machines vectorielles du support

- Les SVM ne fournissent pas directement d'estimation de probabilité, celles-ci sont calculées à l'aide d'une validation croisée.

3.4 Critères d'évaluation des solutions existantes

Afin de bien évaluer les travaux que nous avons déjà lu, nous avons établi une liste de critères d'évaluation qui se compose de datasets, complexité et taux de reconnaissance.

- **Datasets** : Le dataset JAFFE est utilisé dans cette étude comparative, il contient 213 images de 7 expressions faciales (6 expressions faciales de base + 1 neutre). Ainsi que le dataset CK+ qui se compose de 327 séquences d'expressions faciales de 123 sujets.
- **Taux de reconnaissance** : Dans tous les traitements, les tests de reconnaissance des expressions faciales permettent de reconnaître les émotions des personnes (joie, tristesse, colère, peur, dégoût, surprise), pour être efficace le processus de reconnaissance devrait être assez simple et il est recommandé d'utiliser des méthodes moins complexes.

3.5 Résultat de la comparaison

Dans les tableaux 3.1 et 3.2 suivants nous allons effectuer une étude comparative de quelques travaux qui ont utilisé les deux méthodes pour la reconnaissance d'expressions faciales.

3.5.1 Par rapport à CK+

Dans ce qui suit nous avons effectué une étude comparative entre les différents travaux qui ont effectué leurs test sur le dataset CK+

Auteurs	Année	Méthode	Taux
Lee et al	2019	Réseau de neurone à convolution (CNN)	94.98%
Faradjzadeh et al	2019	La méthode de machine à vecteur du support (SVM)	98.50%

TABLE 3.1: Comparaison des performances des méthodes de la reconnaissance d'émotion pour le dataset CK+

En utilisant le dataset CK+, le meilleur taux de reconnaissance est obtenu en utilisant la méthode de machine à vecteur du support (SVM) par les auteurs Faradjzadeh et al, une reconnaissance de 98.50

3.5.2 Par rapport à JAFEE

Dans ce qui suit nous avons effectué une étude comparative entre les travaux qui ont effectué leurs test sur le dataset JAFEE

Auteurs	Année	Méthode	Taux
Lee et al	2019	Réseau de neurone à convolution (CNN)	97.91%
Faradjzadeh et al	2019	La méthode de machine à vecteur du support (SVM)	94.64%

TABLE 3.2: Comparaison des performances des méthodes de la reconnaissance d'émotion pour le dataset JAFEE

Le taux de reconnaissance le plus élevé est 97.91%, il a été obtenu en utilisant la méthode des réseaux de neurones à convolution (CNN) avec le dataset JAFEE par Lee et al.

3.6 Conclusion

Après avoir effectué une étude comparative entre les deux méthodes de la reconnaissance d'émotion, il s'est avéré que les deux méthodes ont un taux très élevé, la méthode de machine à vecteur du support avec le dataset CK+ et la méthode des réseaux de neurone avec le dataset JAFFE.

Dans l'étude qui suit, nous allons présenter notre solution qui consiste à reconnaître des émotions on utilisant les réseaux de neurones à convolution.

Chapitre 4

Mise en œuvre d'une méthode de reconnaissance d'expressions faciales

4.1 Introduction

Ce chapitre est consacré à la mise en œuvre de notre méthode proposée pour la reconnaissance d'expression faciale qu'ont appelée RNREF (Réseau de neurone pour la reconnaissance d'expressions faciales) ainsi qu'aux outils utilisés pour l'implémentation et enfin nous allons présenter les résultats obtenus.

4.2 Méthode proposée

L'architecture du réseau de neurone est comme suit :

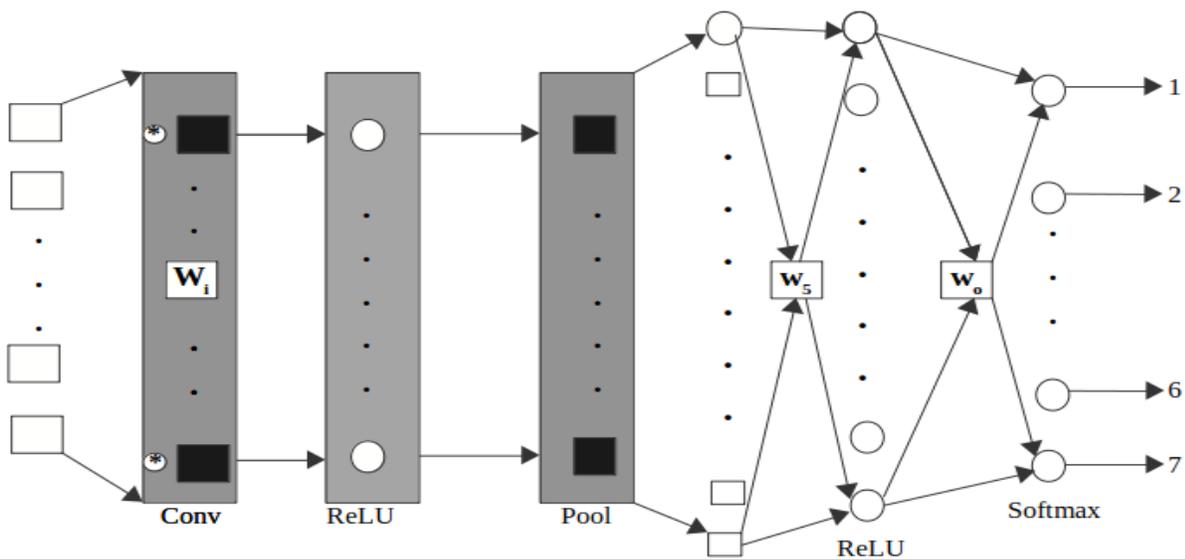


FIGURE 4.1: Architecture du réseau de neurone proposé

4.2.1 Fonctionnement du réseau de neurone

Dans notre approche, nous avons considéré une classe de réseaux avec 6 couches, une couche d'entrée convolutionnelle, deux autres avec une fonction d'activation ReLu, une avec Pooling, une autre avec Reshape et une couche de sortie avec la fonction d'activation Softmax De plus, nous avons effectué une régularisation pour chaque matrice des poids W .

4.2.1.1 Couche de convolution

Cette section explique comment fonctionne la couche de convolution, qui sert à l'extraction d'entités à partir d'images. La couche de convolution génère de nouvelles images appelées cartes d'entités. La carte des entités accentue les caractéristiques uniques de l'image d'origine. La couche de convolution fonctionne de manière très différente par rapport aux autres couches du réseau neuronal. Cette couche n'utilise pas une somme pondérée des connexions. Au lieu de cela, elle contient des filtres qui convertissent les images. Nous appellerons ces filtres des filtres convolutifs. Le processus d'entrée de l'image à travers les filtres de convolution donne la carte des caractéristiques (des entités).

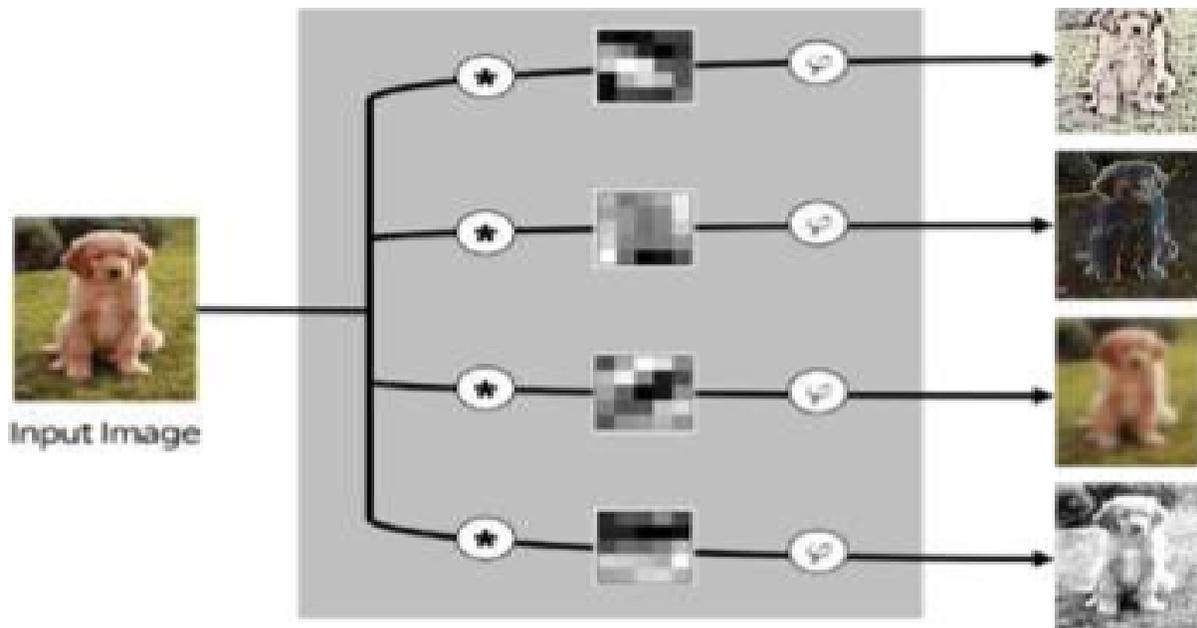


FIGURE 4.2: Effet de la couche de convolution

La figure précédente montre le processus effectué par la couche de convolution, où la marque encadrée * indique l'opération de convolution et la marque σ est la fonction d'activation. Les icônes carrées en niveaux de gris entre ces opérateurs indiquent les filtres de convolution. La couche de convolution génère le même nombre de cartes d'entités que les filtres de convolution. Par conséquent, par exemple, si la couche de convolution contient quatre filtres, elle générera quatre cartes d'entités.

4.2.1.2 Convolution matricielle

Une image est représentée généralement sur ordinateur comme une matrice. En noir est blanc, le cas le plus simple, serait de considérer chaque pixel en binaire (noir =1, blanc=0). Le pixel peut contenir aussi le niveau en gris (méthode utilisée dans les écrans noir et blanc). Les images en couleurs sont codées en RGB, c'est-à-dire trois matrices (R pour le niveau de la couleur rouge, G pour le niveau de la couleur verte, B pour le niveau de la couleur bleue), ceci donne trois valeurs pour chaque pixel.

convolution suivante entre une image de 44 pixels et un filtre de 22.

$$\begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 3 \\ \hline 4 & 6 & 4 & 8 \\ \hline 30 & 0 & 1 & 5 \\ \hline 0 & 2 & 2 & 4 \\ \hline \end{array} \circledast \begin{array}{|c|c|} \hline 1 & 0 \\ \hline 0 & 1 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 7 & 5 & 9 \\ \hline 4 & 7 & 9 \\ \hline 32 & 2 & 5 \\ \hline \end{array}$$

L'opération de convolution est la somme des produits des éléments qui se trouvent aux mêmes positions des deux matrices. Le résultat de 7 dans le calcul précédent est calculé comme suit :

$$(11) + (10) + (40) + (61) = 7$$

4.2.1.3 Fonction ReLU

La fonction ReLU est ne fonction d'activation qui sert à réduire certaines connexions du réseau profond. Sa formule mathématique est la suivante :

$$(v) = \max(V, 0) = v \text{ si } v > 0 \\ 0 \text{ ailleurs}$$

4.2.1.4 Couche de POOLING

La couche de regroupement (POOLING) réduit la taille de l'image, car elle combine les pixels voisins d'une certaine zone de l'image en une seule valeur représentative. Afin de mener les opérations dans la couche de POOLING, nous devons déterminer comment sélectionner les pixels de

POOLING de l'image et comment définir la valeur représentative. Les pixels voisins sont généralement sélectionnés dans la matrice carrée et le nombre de pixels combinés diffère d'un problème à l'autre. La valeur représentative est généralement définie comme la moyenne (mean pooling) ou le maximum (max pooling) des pixels sélectionnés. Le fonctionnement de la couche de mise en commun est étonnamment simple. Comme il s'agit d'une opération bidimensionnelle et qu'une explication dans le texte peut créer plus de confusion, passons en exemple. Considérons l'image d'entrée de 4 4 pixels, qui est exprimée par la matrice suivante :

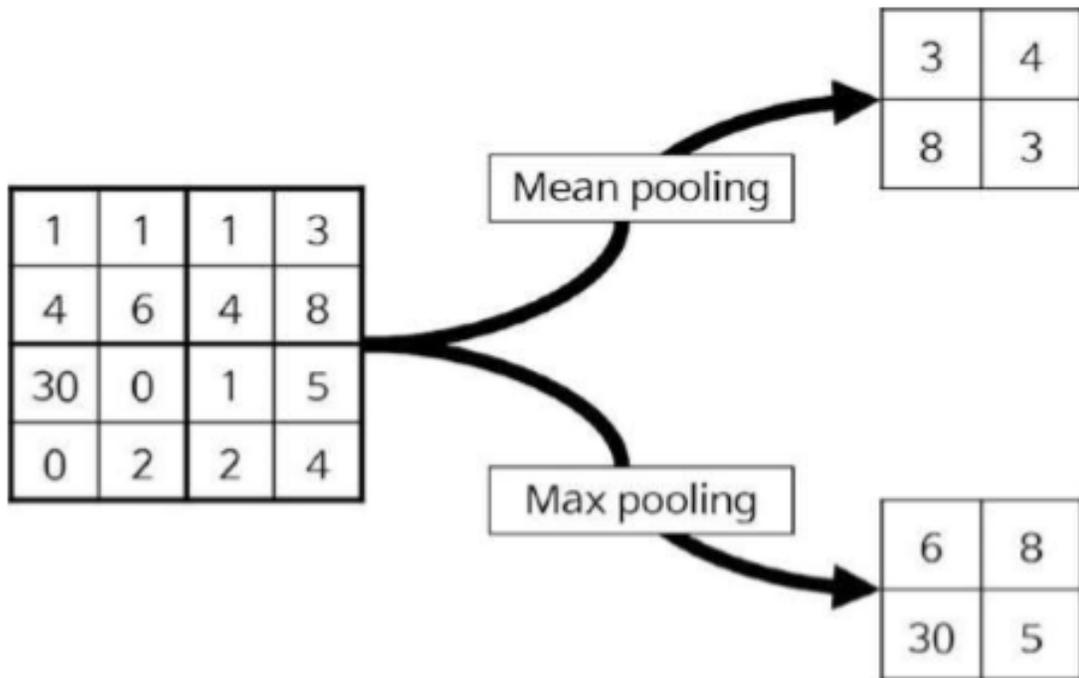


FIGURE 4.3: Illustration de l'opération de pooling.

En fait, dans un sens mathématique, le processus de regroupement (pooling) est un type d'opération de convolution. La différence avec la couche de convolution est que le filtre de convolution est stationnaire et que les zones de convolution ne se chevauchent pas. De plus, comme le processus POOLING réduit la taille de l'image, il est très bénéfique pour soulager la charge de calcul et empêcher le sur-ajustement.

4.2.1.5 Fonction d'activation softmax

Très utilisée pour la classification, la formule mathématique est donnée par :

La majorité des algorithmes du Deep learning utilisent cette fonction d'activation pour la couche de sortie. La fonction softmax maintient la somme des valeurs de sortie à un et limite également les sorties individuelles à des valeurs de 0 à 1. Comme elle tient compte des grandeurs relatives de toutes les sorties, la fonction softmax est un choix approprié pour les réseaux de neurones de classification multiclassés.

$$\Psi(v_i) = \frac{e^{v_i}}{\sum_{i=1}^M e^{v_i}}$$

4.3 Exemple illustratif

En premier lieu le réseau de neurone prendra les images de taille 128*128 comme données d'entrée, ensuite il appliquera les 10 filtre de taille 11*11 sur les images dans la couche de convolution on aura donc $Y1 = \text{Conv}(X, W1) = 118*118*10$, deuxièmement les données passeront par la fonction d'activation Relu, le résultat de ce dernier est de la même taille que ce qui lui est passé en entrée, avec simplement toutes les valeurs négatives éliminées $Y2 = \text{Relu}(Y1)$, $\text{Dim}(Y2) = 118*118*10$, troisièmement les données passeront par la couche de regroupement POOLING qui permet de prendre une large image et d'en réduire la taille tout en préservant les informations les plus importantes qu'elle contient $Y3 = \text{Pool}(Y2)$ dont la dimension est égale à : $\text{Dim}(Y3) = 59*59*10$, quatrièmement, les données passeront par la couche RESHAPE qui transforme les données sous forme de matrice en données linéaire $Y4 = \text{Reshape}(Y3, [], 1) = [(59*59*10)*1]$, cinquièmement le réseau de neurone multiplie le résultat obtenu avec la matrice des poids $W5$ et on aura $V5 = W5*Y4$, sixièmement les données passeront par la fonction d'activation Relu, $Y5 = \text{Relu}(V5)$, septièmement le réseau de neurone multiplie le résultat obtenu avec la matrice des poids $W5$ et on aura : $V = W_o * Y5$, Et au final nous utilisons la fonction d'activation Softmax pour la classification $Y = \text{Softmax}(V)$

4.4 Conclusion

Dans ce chapitre nous avons proposé une méthode pour la reconnaissance des expressions faciales qu'on a appelée RNREF (Réseau de neurone pour la reconnaissance d'expressions faciales) pour la reconnaissance de sept émotions suivantes (anger, disgust, fear, happiness, neutral, sadness, surprise).

Conclusion générale

Les expressions faciales sont la méthode la plus visuelle pour transmettre des émotions et elles diffèrent d'une personne à une autre. La reconnaissance automatique des expressions faciales est importante pour les interfaces naturelles homme-machine. Bien que les humains reconnaissent les expressions faciales sans délai, la réorganisation des expressions par ordinateur reste un défi.

Dans ce travail on s'est intéressé à la reconnaissance des émotions faciales, pour cela nous avons appliqué un algorithme de type deep learning CNN. Notre méthode proposée pour cette reconnaissance est appelée RNREF (Réseau de neurone pour la reconnaissance d'expressions faciales), elle consiste à reconnaître les émotions des personnes, on les classe en suite dans les sept classe prédéfinie (colère, dégoût, peur, bonheur, neutre, tristesse, surprise).

Après avoir étudié plusieurs travaux qui existe déjà dans la littérature, on a décidé d'utiliser les réseaux de neurone pour le problème de la reconnaissance d'émotion, et on a constaté que nos résultats théoriques obtenus sont vraiment encourageante.

Bien que de nombreux progrès ont été réalisés dans ce mémoire nous envisageant d'implémenter et simuler notre approche.

Bibliographie

- [1] K. Bouchra. Mise au point d'une application de reconnaissance faciale. Master's thesis, université de Tlemcen, Algerie, 2013.
- [2] A. J. Goldstein, L. D. Harmon, and A. B. Leski. Identification of humman faces. *Proc. IEEE*, 59(5) :748–760, May 1971.
- [3] L. Sirovich and M. Kirby. A low 47 -dimensional procedure for the characteri- zation of human faces. *Josa a*, 4(3) :559–524, 1987.
- [4] M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. In *Pro- ceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 586–591. IEEE, 1991.
- [5] National Science and Technology Concil (NSTC). Face recongnition. *Comittee on Technology*, page 10, 7 Aout 2006.
- [6] A. Boughida, M. N. Kouahla, and F. Z. Bouhlaci et al. Emad : Un système d'apprentissage humain adaptatif à base d'émotions. *ISKO Maghreb*, pages 25– 26, (2018).
- [7] P. Ekman. Facial expression and emotion. *American psychologist*, 48(4) :384, 1993.
- [8] F. Khalfi. Reconnaissance automatique des émotions par données multimodales : expressions faciales et des signaux physiologiques. PhD thesis, Université Paul Verlaine de Metz, France, 2010.
- [9] E. Couzon and F. Dorn. *Les émotions : développer son intelligence émotionnelle*. ESF editeur, 2011.
- [10] M. Meila. Comparing clusterings : an axiomatic view. In *Proceedings of the 22nd international conference on Machine learning*, pages 577–584. ACM, 2005.
- [11] scikit-learn. <https://scikit-learn.org/stable/>, (Consulte le 20 janvier 2019).
- [12] O. Chapelle, B. Scholkopf, and A. Zien. Semi-supervised learning. *IEEE Tran- sactions on Neural Networks*, 20(3) :542–542, 2009.
- [13] Réseau neuronal. [https://www.futura-sciences.com/tech/definitions/ informatique-reseau-neuronal-601/](https://www.futura-sciences.com/tech/definitions/informatique-reseau-neuronal-601/), (Consulte le 14 juin 2020).
- [14] K-nearest neighbors implementation using scikit-learn. https://etav.github.io/algorithms/knn_implementation_scikit-learn.html, (Consulte le 10 juin 2020) Reference pour les SVM et les arbre de decision.

- [15] National Science and Technology Concil (NSTC). Face recongnition. Comittee on Technology, page 10, 7 Aout 2006.
- [16] S.Guerfi. Authentification d'individus par reconnaissance de caractéristique bio- metriques liees aux visages 2D/3D. PhD thesis, universite d'Evry-val d'Essonne, France, 2008.
- [17] L. FaceDSK. 1.7 : Face detection and recognition library developer's guide. Online Document, 2008.
- [18] Z. HAMMAL, N.EVENO, A.CAPLIER, and PY.COULON. Extraction des traits caractéristiques du visage a l'aide de modèles paramétriques adaptes. I- Revues, 2005.
- [19] Deepak Jain, Pourya Shamsolmoali, and Paramjit Sehdev. Extended deep neural network for facial emotion recognition. Pattern Recognition Letters, 120, 04 2019.
- [20] Neha Jain, Shishir Kumar, Amit Kumar, Pourya Shamsolmoali, and Masoumeh Zareapoor. Hybriddeepneuralnetworksforfaceemotionrecognition. Pattern Recognition Letters, 115 :101 106, 2018.
- [21] Jing Li, Yang Mi, Jiahui Yu, and Zhaojie Ju. A novel convolutional neural network for facial expression recognition. In International Conference on Cognitive Systems and Signal Processing, pages 310320. Springer, 2018.
- [22] Nacer. Farajzadeh, Mahdi. Hashemzadeh. Exemplar-based Facial expression recognition, 2018
- [23] Nikunja Bihari Kar, Korra Sathya Babu¹, Arun Kumar Sangaiah et Sambit Bakshi¹, Face expression recognition system based on ripplelet transform type II and least square SVM, Springer Science+Business Media, LLC, part of Springer Nature 2017