



Université Mohamed Khider de Biskra
Faculté des Sciences et de la Technologie
Département de Génie Electrique

MÉMOIRE DE MASTER

Sciences et Technologies
Télécommunications
Réseaux et télécommunication

Réf. :

Présenté et soutenu par :
Bennoui Manel

Le : 2020

Reconnaissance de visage 3D par un réseau de neurones à convolution

Jury :

Pr.	Debilou Abderrazzak	Pr	Université de Biskra	Président
Dr.	Medouakh Saadia	MCB	Université de Biskra	Examineur
Dr.	Ouamane Abdelmalik	MCA	Université de Biskra	Rapporteur

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'enseignement Supérieur et de la recherche scientifique



Université Mohamed Khider Biskra
Faculté des Sciences et de la Technologie
Département de Génie Electrique
Filière : Télécommunications
Option : Réseau et Télécommunications

Mémoire de Fin d'Etudes
En vue de l'obtention du diplôme:

MASTER

Thème

Reconnaissance de visage 3D par un réseau de neurones à convolution

Présenté par :

Bennoui Manel

Avis favorable de l'encadreur :

Dr. Ouamane Abdelmalik

Avis favorable du Président du Jury

Pr. Debilou Abderrazzak

Cachet et signature

Remerciement

Au terme de ce mémoire, je tiens à exprimer mes remerciements et ma profonde gratitude avant tout au bon

DIEU de m'avoir donné la force et le courage pour mener à bien ce modeste travail.

*Je saisis de cette occasion pour exprimer mon profond remerciement à mon enseignant encadreur **Dr. Abdelmalik OUAMANE** pour son soutien et son louable effort, durant toute la durée de ce travail.*

*Un grand merci aux membres du jury composé de Pr. **DEBILOU Abderrazzak** et Dr. **MEDOUAKH, Saadia**, qui ont accepté d'évaluer et d'examiner notre travail.*

*Mes remerciements y vont aussi à l'endroit de tous les professeurs du département de Génie Electrique de l'université **Université Mohamed Khider – Biskra***

Sommaire

Dédicace

Remerciement

Résumé

Introduction générale **1**

Chapitre 1 : Concepts de base pour la reconnaissance de visage 3D

1.1 Introduction	3
1.2 Difficultés de reconnaissance de visage 2D	3
1.2.1 Illumination	3
1.2.2 Pose	4
1.2.3 Expressions faciales	5
1.2.4 Occlusions	5
1.3 Opportunités de reconnaissance de visage 3D	6
1.4 Techniques d'acquisition de visage 3D	6
1.5 Mesure de la performance	8
1.6 Etat de l'art en reconnaissance de visages 3D	10
1.7 Conclusion	14

Chapitre 2 : Réseau de neurone à convolution

2.1 Introduction	16
2.2 Les couches CNN	16
2.3 L'étape de convolution	17
2.3.1 Profondeur	18
2.3.2 Foulée Stride	18
2.3.3 Remplissage nul (Zero-padding)	18
2.4 Présentation du non linéarité (ReLU)	19
2.5 L'étape Pooling	20
2.6 Couches totalement connecté (Fully Connected Layer)	20
2.7 Les applications de CNN	21
2.8 Conclusion	23

Chapitre 3: Conception de système et résultats

3.1 Introduction	25
3.2 Notre algorithme d'identification du visage 3D	25
3.2.1 Prétraitement d'image	25
3.3. CNN	27
3.4 Base de données CASIA 3D	29
3.5 Résultats	32
3.5.1 Identification uni-modale de visage 3D	32
3.5.2 Identification multimodales (2D + 3D) de visage	38

3.6 Conclusion

40

Conclusion générale

Bibliographie

Liste des figures

Figure 1.1 : Exemple d'un visage d'une même personne subissant un changement de luminosité	4
Figure 1.2 : Exemples de variation de poses variables.	4
Figure 1.3 : Exemple de quelques variations d'expressions faciales de la même personne .	5
Figure 1.4 : Exemples d'occlusion du visage.	5
Figure 1.5 : Exemple de lumière structurée (à gauche) et sa projection sur un visage.	7
Figure 1.6 : Comparaison de la stéréo photométrique (a) procédé d'éclairage photométrique stéréo et (b) stéréo photométrique multi spectrale.	7
Figure 1.7 : Droite : systèmes de capture, Gauche : modèle de visage capturé en utilisant une configuration de studio de plusieurs caméras.	8
Figure 1.8 : La courbe ROC.	9
Figure 1.9 : Exemple de la courbe CMC.	10
Figure 1.10 : Caractéristiques autour desquelles est réalisée une ACP.	11
Figure 1.11: Comparaisons des courbes radiales et la distance géodésique entre les courbes	13
Figure 2.1 : Convolutional neural networks (CNN).	17
Figure 2.2 : L'opération de convolution.	18
Figure 2.3 : ReLU operation .	19
Figure 2.4 : Pooling avec un filtre 2x2 et un pas de 2	20
Figure 2.5 : Couches totalement connecté.	21
Figure 3.1 : Schéma général de notre système	25
Figure 3.2: Prétraitement de l'image de profondeur: (a) données 3D, (b) image 3D coupée, (c) élimination des pointes aberrantes, (d) image de profondeur après étape de prétraitement .	27
Figure 3.3: Prétraitement de l'image 2D: (a) données 2D, (b) image 2D coupée, (c) image d'intensité, (d) normalisation photométrique par l'approche PS (preprocessing sequence approach)	27

Figure 3.4 : Capture des scènes de la base de données CASIA visages 3D .	30
Figure 3.5 : Variations d'éclairage de la base de données CASIA visage 3D	31
Figure 3.6 : Variations d'expressions de la base de données CASIA visage 3D	31
Figure 3.7 : Variations des poses de la base de données CASIA visage 3D	32

Liste des tableaux

Tableau 3.1: Architecture initiale de notre CNN	29
Tableau 3.2 : Taux de reconnaissance (TR) en utilisant Max Pooling et la taille des filtres dans la phase de convolution varient entre 2 et 6 (5 images et 10 images pour l'apprentissage de CNN).	34
Tableau 3.3 : Taux de reconnaissance (TR) en utilisant Moyen Pooling et la taille des filtres dans la phase de convolution varient entre 2 et 6 (5 images et 10 images pour l'apprentissage de CNN).	37
Tableau 3.4 : Taux de reconnaissance (TR) de visage 3D et 2D+3D.	39

RESUMES (Français et Arabe)

Un système de reconnaissance faciale est une application logicielle visant à reconnaître une personne grâce à son visage de manière automatique. C'est un domaine de la vision par ordinateur consistant à reconnaître automatiquement une personne à partir d'une image de son visage.

Les réseaux de neurones à convolution CNNs sont des réseaux de neurones multicouches qui sont spécialisés dans des tâches de reconnaissance et de classification d'images. Dans notre travail on a utilisé ce type de réseaux pour la vérification de visage 3D. Utiliser la base de données des visages **CASIA 3D**.

نظام التعرف على الوجه هو تطبيق برمجي يهدف إلى التعرف على شخص من خلال وجهه تلقائيًا. إنه مجال رؤية الكمبيوتر يتكون تلقائيًا من التعرف على شخص من صورة وجهه.

الشبكات العصبونية الالتفافية هي شبكات عصبية متعددة الطبقات تتمثل مهامها في التعرف على النمط وتصنيف الصورة. في عملنا هذا قمنا باستخدام هذه الشبكات للتعرف على الوجه ثلاثي الأبعاد . باستعمال قاعدة البيانات **CASIA 3D**.

Introduction
Générale

Introduction générale

Introduction générale

Les systèmes biométriques sont des méthodes automatisées qui sont utilisées pour identifier et vérifier les humains au moyen de caractéristiques physiologiques ou comportementales. La reconnaissance faciale est l'un des composants les plus importants du système biométrique, qui a d'énormes applications dans les applications de sécurité, de contrôle et de divertissement [I]. L'idée de la reconnaissance faciale est de donner un ordinateur la capacité de trouver et de reconnaître les visages rapides et précis en images ou en vidéo. De nombreux algorithmes et techniques ont été développés pour améliorer la performance de la reconnaissance faciale [II].

Plusieurs méthodes ont été développées pour la reconnaissance de visage 2D. Cependant, elles présentent un certain nombre de limitations liées à l'orientation du visage ou à la pose, à l'éclairage, à l'expression faciale, aux occultations, etc. Ces dernières années, on parle de plus en plus des techniques de reconnaissance de visage 3D comme une solution alternative pour résoudre les problèmes cités ci-dessus. En effet, la richesse de l'information fournie par les mesures 3D permet de reconstruire la forme tridimensionnelle du visage. Ce type de représentation du visage est invariant aux changements d'illumination et de pose .[III]

L'apprentissage profond est une classe de machines qui peuvent apprendre des niveaux de représentation et d'abstraction en construisant des caractéristiques de haut niveau . Le réseau neuronal conventionnel est un type de modèle d'apprentissage profond . Les CNN (Convolution neurale Networks) sont construits par superposition de couches de convolution qui appliquent un ensemble de filtres locaux à travers les dimensions de la donnée d'entrée. Ainsi, ces réseaux permettent de détecter des motifs en calculant des corrélations locales grâce à un noyau (kernel) dont les paramètres sont déterminés par le processus d'apprentissage. La sortie d'une telle couche, appelée carte de convolution (feature map) est souvent passée à une couche de pooling qui réduit les dimensions de cette carte en le sous-échantillonnant (en général en prenant le maximum ou la moyenne des valeurs sur un ensemble de pixels), donnant ainsi au résultat une propriété d'invariance à la translation. L'empilement de couches de convolution et de pooling à diverses échelles permet de détecter des motifs plus grands et plus complexes . [IV]

Introduction générale

Donc, le but de notre travail est la création d'un réseau de neurone à convolution pour l'identification du visage 3D.

Ce mémoire se présente sous forme de trois chapitres :

Le chapitre 1 nous présente les concepts de base pour la reconnaissance de visage 3D, difficultés de reconnaissance de visage 2D et aussi état de l'art en reconnaissance de visages 3D.

Le chapitre 2 est consacré à l'étude du modèle proposé à savoir le réseau de neurones à convolution et les quatre étapes principales des CNNs.

Le chapitre 3 illustre l'implémentation et l'expérimentation de notre système.

Finalement, une conclusion permet de conclure le travail effectué.

Chapitre 1

Concepts de base pour la reconnaissance de visage 3D

Chapitre 1 : Concepts de base pour la reconnaissance de visage 3D

1.1 Introduction

Le visage humain est une méthode biométrique très efficace. Il est certainement la caractéristique biométrique que les humains utilisent le plus naturellement pour s'identifier entre eux [1], en plus d'être naturel, non-intrusif et précis, la caractéristique la plus importante du visage humain est qu'il peut être capturé à distance et en toute confidentialité, et que la reconnaissance faciale a de nombreux avantages [2] d'accroître le niveau de sécurité dans la société et simplifier le processus d'authentification sur les appareils électroniques. [3]. Et utilisée l'identification et la vérification des humains le premier répondre à la question suivante : suis-je ce que je prétends être ?, et la deuxième pose la question suivante : qui suis-je [4].

1.2 Difficultés de reconnaissance de visage 2D

Les principaux paramètres qui ont un impact sur les performances d'un système de reconnaissance faciale sont [5] :

1.2.1 Changement d'illumination (éclairage)

Le changement d'apparence d'un visage dû à l'illumination (figure 1.1), se révèle parfois plus critique que la différence physique entre les individus. L'intensité et la direction d'éclairage lors de la prise de vue influent énormément sur l'apparence du visage dans l'image ce qui rend la tâche de la reconnaissance du visage très difficile [5].

Chapitre 1 : Concepts de base pour la reconnaissance de visage 3D



Figure 1.1 : Exemple d'un visage d'une même personne subissant un changement de luminosité [6].

1.2.2 Pose

La variation de pose (figure 1.2) est considérée comme un problème majeur pour les systèmes de reconnaissance faciale. Elle permet de baisser considérablement le taux de reconnaissance [5].



Figure 1.2 : Exemples de variation de poses variables [6].

Chapitre 1 : Concepts de base pour la reconnaissance de visage 3D

1.2.3 Expressions faciales

Un autre type d'effet facial est le changement des expressions (figure 1.3). L'expression faciale de l'émotion, combinée avec la parole, peut produire des changements d'apparence importants des visages. Cette expression entraîne forcément une diminution du taux de reconnaissance [5].

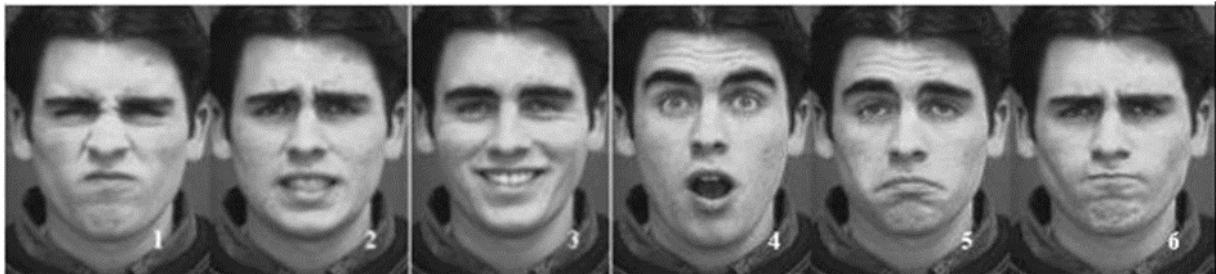


Figure 1.3 : Exemple de quelques variations d'expressions faciales de la même personne (Image recueillie à partir d'Internet).

1.2.4 Occlusions

Les occlusions partielles apparaissent fréquemment dans des applications réelles. La présence des composants structurels (la barbe, la moustache, les lunettes, etc.) peut modifier énormément les caractéristiques à savoir la forme, la couleur, ou la taille du visage. De plus, ces composants peuvent cacher les caractéristiques faciales de bas (figure 1.4) [5].



Figure 1.4 : Exemples d'occlusion du visage (Image recueillie à partir d'Internet).

Chapitre 1 : Concepts de base pour la reconnaissance de visage 3D

1.3 Opportunités de reconnaissance de visage 3D

Il y a tant de difficultés techniques pour la reconnaissance faciale 2D du visage. On peut suggérer d'utiliser les données 3D pour la reconnaissance faciale [5]. Il y a plusieurs raisons à cela :

- L'invention et la vulgarisation des techniques de numérisation 3D : de nos jours, de nombreux types de capteurs 3D sont disponibles pour capturer des données 3D statiques et même dynamiques de surface du visage.
- La disponibilité d'un grand potentiel pour lutter contre les difficultés 2D : les informations sur la forme des approches fondées sur la reconnaissance faciale en 3D sont naturellement moins sensibles aux conditions d'éclairage par comparaison avec les informations de photométrie 2D.
- La surface faciale 3D est plus commode pour effectuer la correction de pose et résoudre le problème de rotation dans la reconnaissance faciale 2D.
- les cosmétiques pour le visage ont tendance à avoir moins d'impact pour les scans du visage 3D [7].
- La forme du visage 3D et la photométrie 2D contiennent des informations comparables et complémentaires

1.4 Techniques d'acquisition de visage 3D

1.4.1 Lumière structurée (structured light)

La lumière structurée a été largement utilisée dans l'acquisition de données de visages 3D. La configuration de base de cette technique comprend une caméra et un projecteur de lumière dans les relations géométriques données. Le projecteur projette un ou plusieurs motifs de lumière codés sur la scène. La forme des signaux est extraite par la mesure de la déformation de motifs lumineux à la surface de l'objet [9]. La figure 1.5 illustre un exemple des motifs de lumière codés et projetés sur un visage.

Chapitre 1 : Concepts de base pour la reconnaissance de visage 3D



Figure 1.5 : Exemple de lumière structurée (à gauche) et sa projection sur un visage.[9]

1.4.2 Photométrie stéréo (Stereo photometric)

Est une technique pour estimer les normales de la surface d'objets en capturant un ensemble d'images de l'objet sous différentes conditions d'éclairage. Cette technique est très sensible à la présence d'ombres, aux faits saillants et à l'éclairage non uniforme [9].

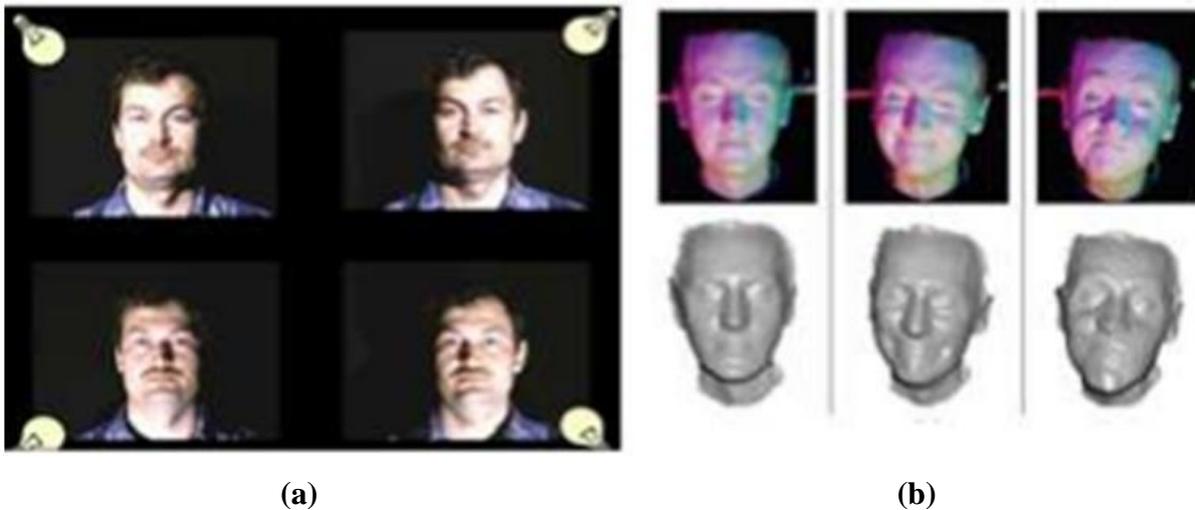


Figure 1.6 : Comparaison de la stéréo photométrie (a) procédé d'éclairage photométrique stéréo et (b) stéréo photométrie multi spectrale (image recueillie à partir d'Internet).

1.4.3 Multi-vues stéréo (Multi-view stereo)

Cette technique utilise plusieurs caméras calibrées placées à différents points de vue du sujet. Sous diverses contraintes, les points correspondants de différentes images de bon sens peuvent être mis en place et ainsi la forme du visage peut être reconstruite. Cette

Chapitre 1 : Concepts de base pour la reconnaissance de visage 3D

technique ne nécessite pas de clignotant, puisque tout l'appareil peut enregistrer le même sens en même temps avec des lumières constantes.



Figure 1.7 : Droite : systèmes de capture, Gauche : modèle de visage capturé en utilisant une configuration de studio de plusieurs caméras (Image recueillie à partir d'Internet).

1.5 Mesure de la performance

Les critères de mesure de la performance utilisés pour le **mode vérification** sont :

- ✓ **Le Taux de Faux Rejet (TFR) ou False Reject Rate (FRR) :** c'est le pourcentage d'individus qui ont été faussement rejetés par le système alors qu'ils sont des clients .

$$TFR = (\text{Nombre de faux rejet})/(\text{Nombre de clients})$$

- ✓ **Le Taux de Fausse Acceptation (TFA) ou False Acceptation Rate (FAR) :** c'est le pourcentage d'individus qui ont été acceptés par le système alors qu'ils sont des imposteurs.

$$TFA = (\text{Nombre de Fausse Acceptation})/(\text{Nombre d'imposteurs})$$

- ✓ **Le Taux d'Égale Erreur (TEE) ou Equal Error Rate (EER)** est calculé pour vérifier la stabilité du système. Ce point de fonctionnement correspond à une valeur où TFR est égal à TFA. Le taux d'erreur TEE est un compromis entre le TFA et le TFR, c'est-à-dire, si un des deux taux est augmenté l'autre sera diminué automatiquement. En d'autres termes, plus l'EER est faible, plus le système est performant. À noter que ce taux d'erreur est le plus couramment utilisé dans la littérature pour illustrer la performance des systèmes biométriques [16].
- ✓ **Le Taux de Vérification (TV) ou Vérification Rate (VR) :** ce taux représente les individus qui sont correctement vérifiés par le système, soit les clients acceptés ou

Chapitre 1 : Concepts de base pour la reconnaissance de visage 3D

bien les imposteurs rejetés. Le taux de vérification permet d'examiner la sensibilité du système de vérification

$$TV = (100 - (TFA + TFR))$$

- ✓ **La courbe ROC (Receiver Operating Characteristic):** une courbe ROC (voir la figure 1.8) présente le taux de faux rejeté en fonction du taux de faux accepté. C'est une courbe strictement décroissante, qui pour un système performant va avoir tendance à épouser le repère.

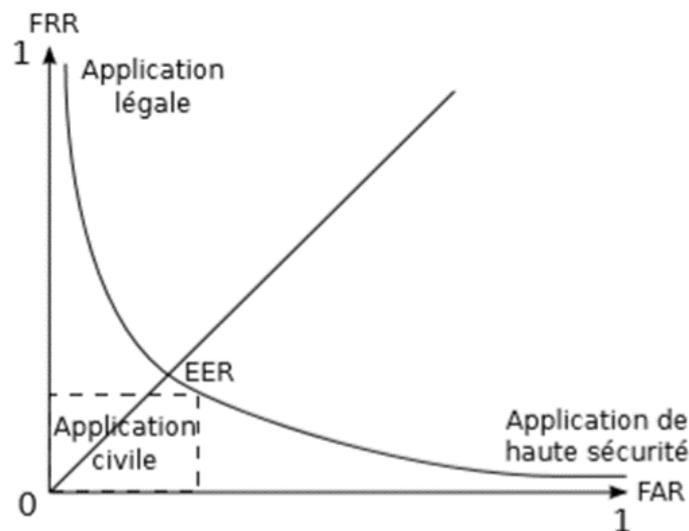


Figure 1.8 : La courbe ROC (Image recueillie à partir d'Internet).

Mesure des performances pour le mode identification :

- ✓ **Rank-1 Recognition Rate (R1RR) :** est le pourcentage de tous les individus pour lesquels le meilleur match dans la galerie appartient au même sujet. Le R1RR représente le pourcentage de toutes les entrées du système qui sont correctement identifiés.
- ✓ **La courbe CMC (Cumulative Match Characteristic) :** la courbe CMC est utilisée pour mesurer la performance d'un système d'identification. Cette courbe donne le pourcentage de personnes reconnues en fonction d'une variable que l'on appelle le rang [8].

Chapitre 1 : Concepts de base pour la reconnaissance de visage 3D

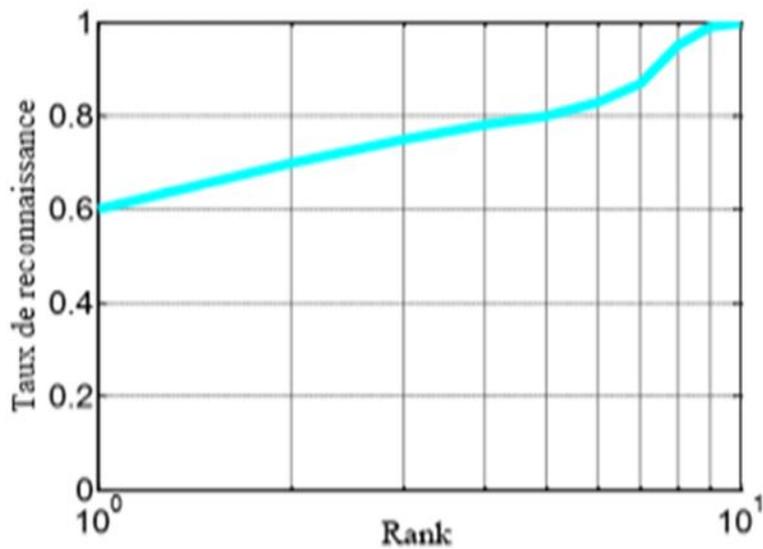


Figure 1.9 : Exemple de la courbe CMC (Image recueillie à partir d'Internet).

1.6 État de l'art en reconnaissance de visages 3D

La reconnaissance faciale 3D est définie comme une technique biométrique qui utilise la forme faciale 3D individuelle pour reconnaître les visages humains à l'aide de modèles de visage 3D.

1.6.1 Approches globales

Les approches globales prennent l'image du visage comme un tout et utilisent des techniques d'analyse statistique bien connues. L'idée est généralement de projeter l'image d'entrée du visage, préalablement vectorisée, dans un espace de plus faible dimension, où la reconnaissance est supposée être plus aisée. La projection est souvent conçue pour ne sélectionner que les caractéristiques importantes et suffisamment discriminantes pour différencier les personnes entre elles. Contiens deux types de techniques : les techniques linéaires et les techniques non linéaires.

Les techniques linéaires

Les techniques linéaires réalisent une projection linéaire des visages (espace dont la dimension est égale à la dimension des images, donc grande) sur un espace de plus faible dimension. La plus connue de ces approches est la technique dite des Eigenfaces présentée par

Chapitre 1 : Concepts de base pour la reconnaissance de visage 3D

Turk et Pentland dans [9]. L'analyse en composantes principales (ACP) est réalisée sur un ensemble d'apprentissage d'images de visages. Les principaux vecteurs propres résultant de l'ACP définissent le nouvel espace. Les images de visages sont ensuite projetées sur cet espace, et les vecteurs obtenus sont utilisés pour la classification. Ainsi, Kirby et al. [10] proposent un critère basé sur l'énergie des valeurs propres associées aux vecteurs propres. Les vecteurs propres correspondant aux plus grandes valeurs propres sont retenus jusqu'à ce que la somme des valeurs propres dépasse un certain seuil de l'énergie totale (90% dans [10]). Une autre approche bien connue présentée par Belhumeur et al. [11] réalisent une Analyse Discriminante Linéaire (LDA), elle est ainsi souvent nommée Fisherfaces. En effet, cette technique consiste à maximiser sur un ensemble d'apprentissage le critère de Fisher, à savoir le quotient de la variance inter-classe par la variance intra-classe [11].

D'autres techniques linéaire sont également été utilisés pour le calcul de vecteurs caractéristiques :

- l'analyse en composantes indépendantes (ICA) dans [12]
- la factorisation de matrices non négatives (NMF) dans [13] ou [14]
- l'analyse discriminante bilinéaire (BDA) dans [15]
- la technique dite de «Vecteurs communs discriminants» (DCV) dans [16]

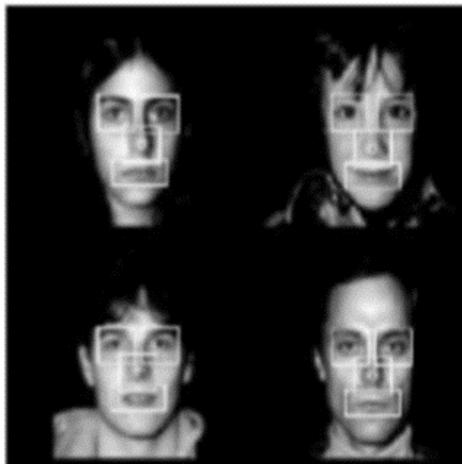


Figure 1.10 : Caractéristiques autour desquelles est réalisée une ACP dans [17].

Techniques non linéaires

Des techniques globales non linéaires ont été développées, souvent à partir des techniques linéaires. Ont également été utilisées dans le contexte de la reconnaissance faciale :

Chapitre 1 : Concepts de base pour la reconnaissance de visage 3D

- le Multi Dimensional Scaling (MDS) dans [18] ou [19]
- l'Isomap dans [20]
- les diffusions maps dans [21]
- le Local Linear Embedding (LLE) dans [22] ou [23]
- les Laplacian eigenmaps dans [24] [25] ou [26]
- le Hessian LLE dans [27],
- le Local Tangent Space Analysis (LTSA) dans [28],
- les approches neuronales dans [29] ou [30] (autoencodeurs), dans [31] (cartes de Kohonen), et dans [32] (réseaux de neurones convolutionnels).

1.6.2 Approches locales

Les approches locales ont été largement étudiées dans la reconnaissance faciale 3D car elles procurent généralement de meilleurs résultats que les approches globales. Ces approches sont classées dans [37] en cinq classes.

Méthodes basées sur le calcul du point

Les méthodes basées sur le calcul du point enlèvent des informations d'un ensemble de points initiaux pour la classification, comme: les coordonnées originales, la distance d'un bord entre deux points, la surface d'une région composée de plusieurs points, l'angle entre les deux bords, exigent des descripteurs plus complexes.

Méthodes basées sur le calcul de courbe

Les méthodes basées sur le calcul de courbe mesurent les courbes correspondantes de différentes surfaces de visage 3D ou calculent les distances en utilisant des fonctionnalités plus puissantes.

Jahanbin et al. [38] décrivent les surfaces 3D par iso-profondeur et courbes iso-géodésique. La dernière étape consiste à vérifier ou rejeter les demandes des utilisateurs à l'aide d'un classificateur de distance euclidienne ou SVM.

Chapitre 1 : Concepts de base pour la reconnaissance de visage 3D

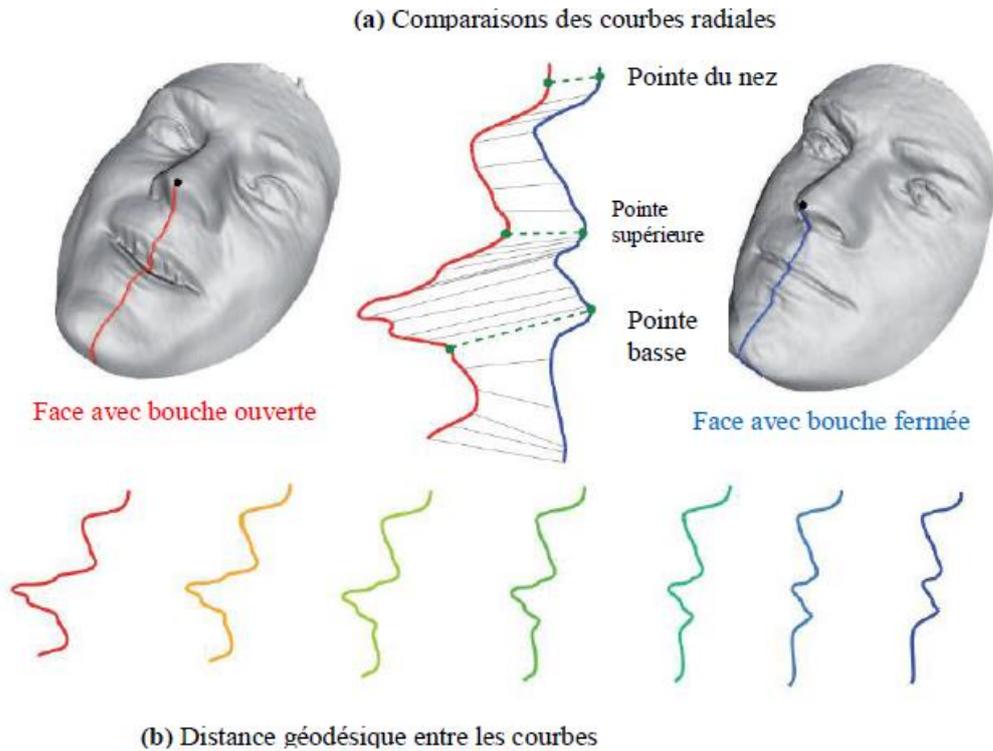


Figure 1.11: Comparaisons des courbes radiales et la distance géodésique entre les courbes [38].

Méthodes basées sur les régions

La totalité du champ facial peut être segmentée relativement en parts rigides et non rigides en donnant un poids plus faible sur les régions du visage non rigides pour diminuer leurs effets. Les régions du visage choisies sont ensuite comparées à un certain espace de caractéristiques pour la reconnaissance de visage.

Les méthodes basées sur les descripteurs

Pour montrer chaque sommet d'un nuage de points du visage 3D, les descripteurs généraux sont des méthodes pour enlever localement des caractéristiques de forme de son voisinage.

Les méthodes basées sur la fusion des caractéristiques multiples

Toutes ses caractéristiques locales mentionnées ci-dessus peuvent être fusionnées pour améliorer les performances de la reconnaissance de visage 3D.

Chapitre 1 : Concepts de base pour la reconnaissance de visage 3D

1.7 Conclusion

Dans ce chapitre nous avons présenté pour quoi on a choisi le visage et les définitions des deux applications de la biométrie l'identification et la vérification, les différents problèmes de reconnaissance de visages 2D et ensuite présente les opportunités de reconnaissance de visage 3D. Donnée un aperçu sur les techniques de mesure de leurs performances. Nous terminons ce chapitre par un état de l'art des techniques de reconnaissance de visages 3D selon qu'elles soient locales et globales.

Chapitre 2

Réseau de neurone à convolution

Chapitre 2 : Réseau de neurone à convolution

2.1 Introduction

Convolution neurale Networks (CNN) a révolutionné le domaine de la vision par ordinateur avec leurs performances exceptionnelles dans les tâches de vision. Il s'agit de fonctions non linéaires conçues pour modéliser un œil humain. Reconnaissant leur importance, des recherches ont été menées pour améliorer leurs rendements. Au fil du temps, l'exactitude du système a augmenté. [1] Les CNN sont un type particulier de réseaux de neurones applicables facilement à des images pour capter spatialement de l'information et de type perceptron multi-couches (Multi Layer Perceptron : MLP). De plus, de par leur structure convolutive, ils permettent de prendre en entrée des données de grande dimension ce qui est une limite du perceptron multicouches [2]

2.2 LES COUCHES CNN

Une architecture CNN est formée par un empilement de couches de traitement indépendantes. On distingue deux parties, une première partie que l'on appelle la partie convolutive du modèle et la seconde partie, que l'on va appeler la partie classification du modèle :

- La couche de convolution (CONV) qui traite les données d'un champ récepteur.
- La couche de correction (ReLU), souvent appelée par abus 'ReLU' en référence à la fonction d'activation (Unité de rectification linéaire).
- La couche de pooling (POOL), qui permet de compresser l'information en réduisant la taille de l'image intermédiaire (souvent par sous-échantillonnage).

Chapitre 2 : Réseau de neurone à convolution

- La couche FC (fully Connected) calcule les scores de classe, ce qui donne un volume de taille où chacun des 10 nombres correspond à un score de classe. Comme pour les réseaux neuronaux ordinaires et comme le nom l'indique, chaque neurone de cette couche est connecté à tous les neurones du volume précédent. [3]

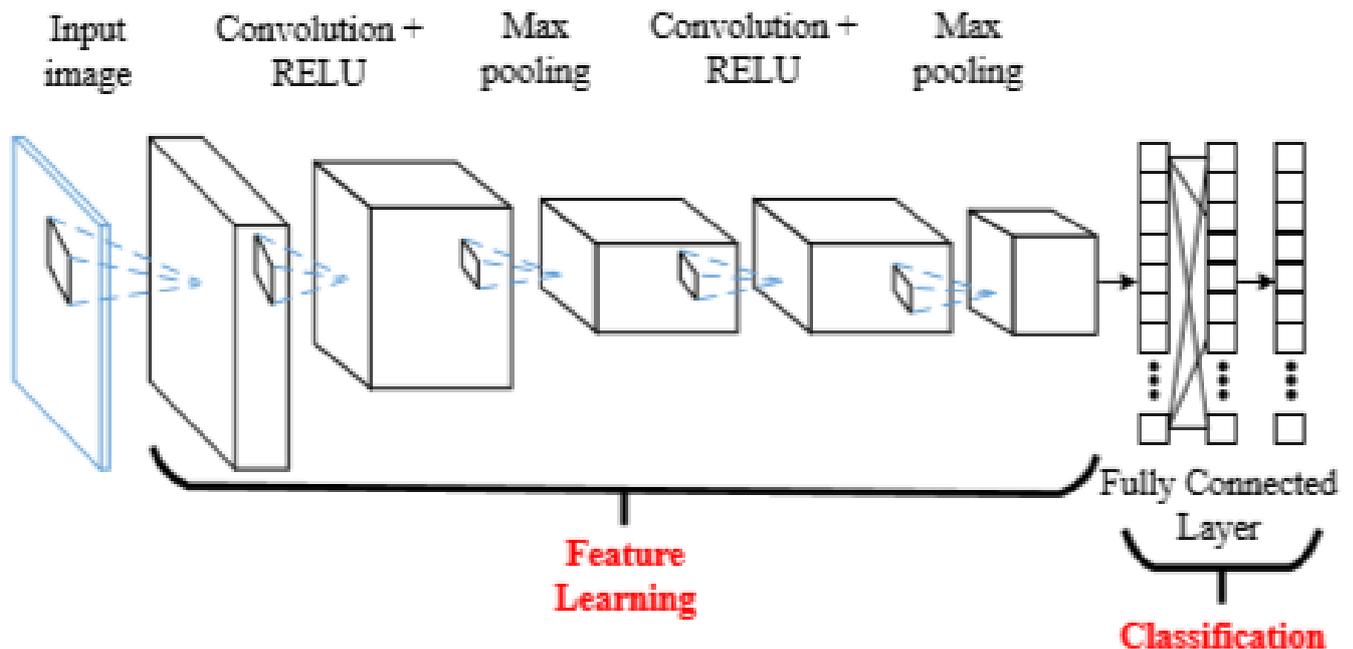


Figure 2.1 :Convolutional neural networks (CNN) [3]

2.3 L'étape de convolution

La couche CONV est le noyau de base d'un réseau neuronal conventionnel. Les paramètres de la couche CONV consistent en un ensemble de filtres K pouvant être appris, où chaque filtre a une largeur et une hauteur, et est presque toujours carré. Ces filtres sont petits (en matière de dimensions spatiales) mais s'étendent sur toute la profondeur du volume. [4]

Chapitre 2 : Réseau de neurone à convolution

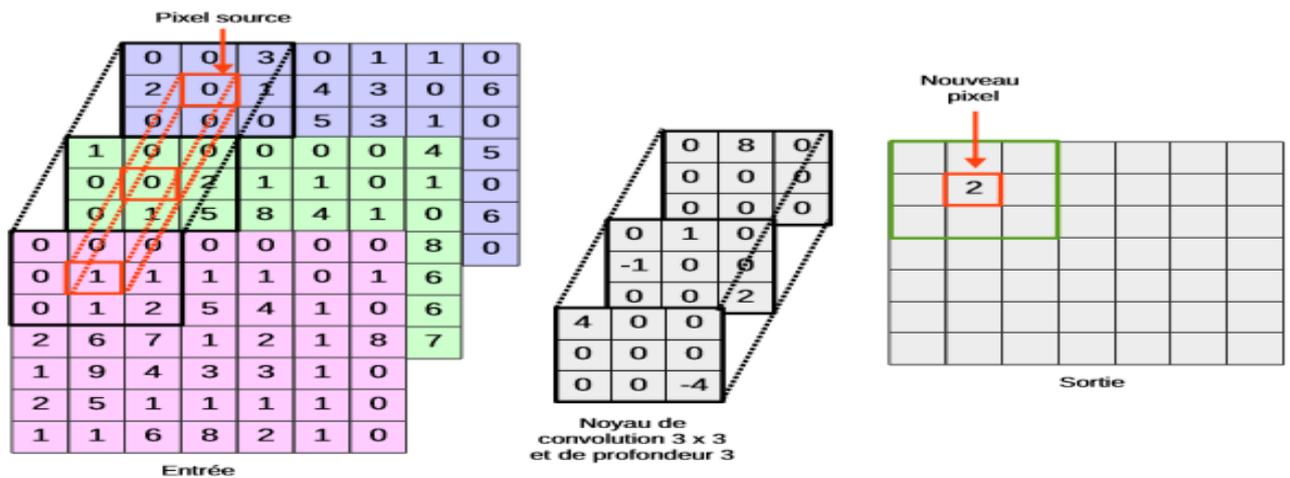


Figure 2.2 L'opération de convolution [2]

Trois hyper paramètres permettent de dimensionner le volume de la couche de convolution (aussi appelé volume de sortie) :

➤ 2.3.1 Profondeur

Profondeur' de la couche est le nombre de noyaux de convolution (ou nombre de neurones associés à un même champ récepteur).

➤ 2.3.2 Foulée (Stride)

Foulée est le nombre de pixels par lequel nous glisser notre filtre sur l'image. Lorsque la foulée est 1 alors nous déplacer les filtres d'un pixel. Lorsque la foulée est 2, les filtres saut 2 pixels à la fois. Avoir une plus grande foulée produira des cartes de caractéristiques plus petites .

➤ 2.3.3 Remplissage nul (Zero-padding)

il est commode de mettre des zéros à la frontière du volume d'entrée. La taille de ce 'zero-padding' est le troisième hyperparamètre. Cette marge permet de contrôler la dimension spatiale du volume de sortie. En particulier, il est parfois souhaitable de conserver la même surface que celle du volume d'entrée . [3]

Chapitre 2 : Réseau de neurone à convolution

2.4 Présentation du non linéarité (ReLU)

ReLU pour une opération non linéaire. La sortie est :

$$G(x) = \max(0, x)$$

Un élément important dans l'ensemble du processus est l'Unité linéaire rectifiée ou Relu. Les mathématiques derrière ce concept sont assez simples encore une fois: chaque fois qu'il y a une valeur négative dans un pixel, on la remplace par un 0. [5] .

Souvent, la correction Relu est préférable, mais il existe d'autre forme

- La correction par *tangente hyperbolique* $f(x)=\tanh(x)$,
- La correction par la *tangente hyperbolique saturante*: $f(x)=|\tanh(x)|$,
- La correction par la *fonction sigmoïde* $f(x)=\frac{1}{1+e^{-x}}$. $f(x)=\frac{1}{1+e^{-x}}$

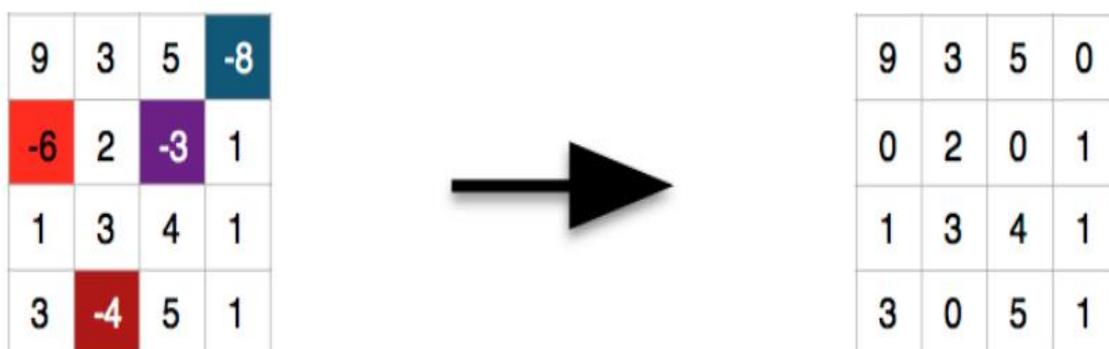


Figure 2.3 :ReLU operation (Image recueillie à partir d'Internet).

Chapitre 2 : Réseau de neurone à convolution

2.5 L'étape Pooling

L'idée principale de la mise en commun est l'échantillonnage descendant afin de réduire la complexité pour d'autres couches. Dans le domaine du traitement d'images, il peut être considéré comme similaire à la réduction de la résolution. Le regroupement n'affecte pas le nombre de filtres. Max Pooling est l'un des types les plus courants de méthodes de pooling. Il divise l'image en rectangles de sous-régions, et ne retourne que la valeur maximale de l'intérieur de cette sous-région. [6]

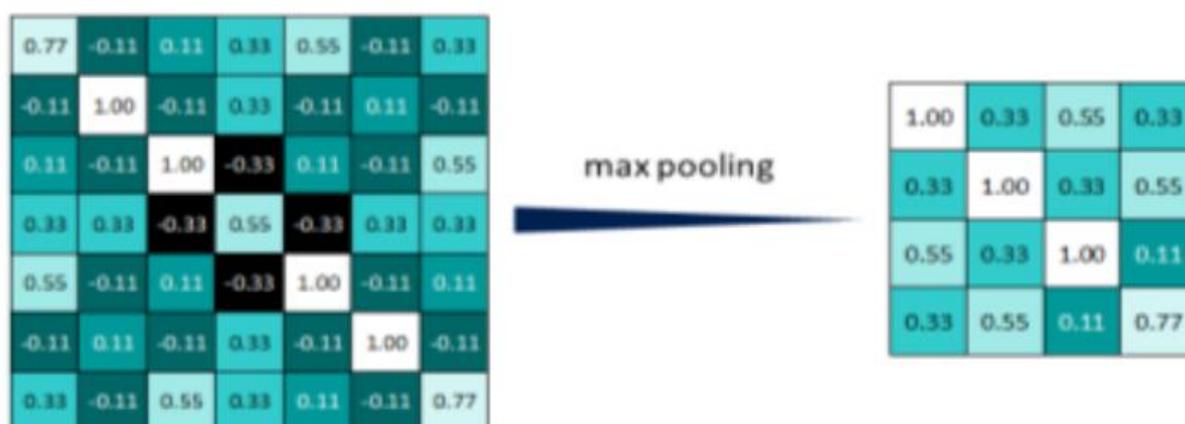


Figure 2.4 : Pooling avec un filtre 2x2 et un pas de 2 (Image recueillie à partir d'Internet).

2.6 Couches totalement connecté (Fully Connected Layer)

La couche fully-connected constitue toujours la dernière couche d'un réseau de neurones, convolutif ou non – elle n'est donc pas caractéristique d'un CNN . Chaque caractéristique de la couche spatiale finale est connectée à chaque État caché de la première couche entièrement connectée. Cette couche fonctionne exactement de la même manière qu'un réseau de flux

Chapitre 2 : Réseau de neurone à convolution

continus traditionnels. Dans la plupart des cas, on peut utiliser plus d'une couche entièrement connectée pour augmenter la puissance des calculs vers la fin. Les connexions entre ces couches sont exactement structurées comme un réseau de feed-forward traditionnel. Comme les couches entièrement connectées sont étroitement connectées, la grande majorité des paramètres se trouvent dans les couches . [7]

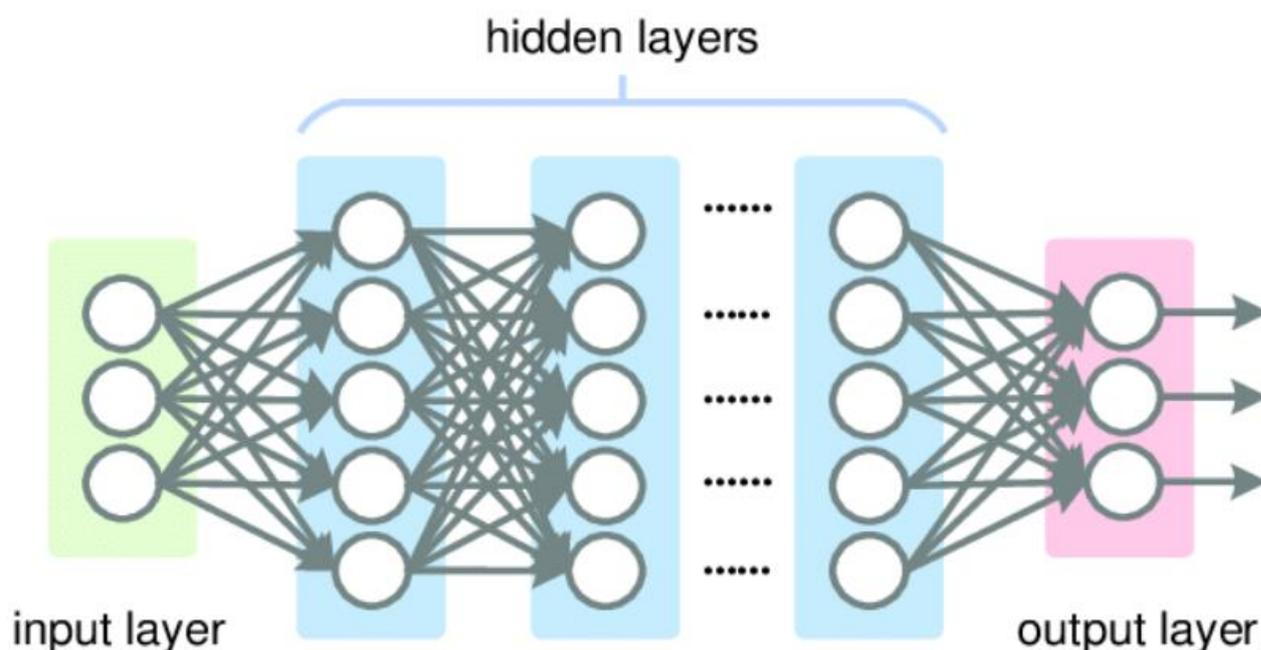


Figure 2.5 : Couches totalement connecté (Image recueillie à partir d'Internet).

2.7 Les applications de CNN :

❖ vision par ordinateur :

1 la reconnaissance de visages :

La reconnaissance faciale constitue une série de problèmes connexes. [8]

- Identifier tous les visages de l'image

Chapitre 2 : Réseau de neurone à convolution

- Se concentrer sur chaque visage malgré un mauvais éclairage ou une pose différente
- Identification des caractéristiques uniques
- Comparaison des caractéristiques identifiées avec la base de données existante et détermination du nom de la personne

2 Analyser des documents :

Les réseaux neuronaux convolutionnels peuvent également être utilisés pour l'analyse documentaire. Cela n'est pas seulement utile pour l'analyse manuscrite, mais a également un intérêt majeur dans les reconnaissances. Pour qu'une machine puisse scanner l'écriture d'un individu, puis la comparer à la base de données larges qu'elle possède, elle doit exécuter presque un million de commandes par minute. Il est dit avec l'utilisation de CNN et de nouveaux modèles et algorithmes. [9]

❖ Traitement du langage naturel

Nous explorons comment les CNN ont été utilisés dans la reconnaissance de la parole au cours des dernières années. Récemment, les CNN ont également été appliqués aux tâches de classification des phrases, de catégorisation des sujets, d'analyse des sentiments et bien d'autres.

Reconnaissance de la parole :

Les réseaux neuronaux convolutions ont été utilisés récemment dans la reconnaissance de la parole et ont donné de meilleurs résultats par rapport aux réseaux neuronaux profonds (DNN). En 2015, les chercheurs de Microsoft Corporation ont indiqué les plus domaines dans lesquels CNN donne de meilleurs résultats que DNN. Ils sont :

- la robustesse du bruit
- Reconnaissance vocale à distance .

Chapitre 2 : Réseau de neurone à convolution

2.8 Conclusion

Dans ce chapitre, nous avons d'expliquer les concepts principaux de réseau de neurone à convolution (CNN) . on distingue quatre types de couche . premièrement la couche de convolution Ce qui se passe à cette couche est détecte les features es utilisant le filtrage par convolution . Deuxième opération est la fonction nonlinéarité (ReLU) pour remplacer toutes les valeurs de pixels négatives par zéro. La troisième opération est pooling elle reduire le nombre de paramètre du réseau , et on applique une méthode de max pooling . et la couche finale entièrement connectée pour la classification . on parler aussi sur les applications de réseaux de neurones convolution . CNN est meilleur que les autres méthodes d'apprentissage profond dans les applications relatives à la vision par ordinateur et au traitement du langage naturel, car il atténue la plupart des problèmes traditionnels.

Chapitre 3
Conception du
système et résultats

Chapitre 3: Conception du système et résultats

3.1 Introduction

La reconnaissance de visage est un problème important, qui trouve des applications dans différents domaines. Plusieurs méthodes traditionnelles ont été utilisées dans la reconnaissance faciale. Au nombre de ces méthodes les réseaux de neurones à convolution CNN.

3.2 Notre algorithme d'identification du visage 3D

Dans ce travail, nous présentons un système (figure 3.1) d'identification du visage 3D basé sur CNN. Dans la première étape, une phase de prétraitement est nécessaire pour le remplissage des trous et le rééchantillonnage des données 3D.

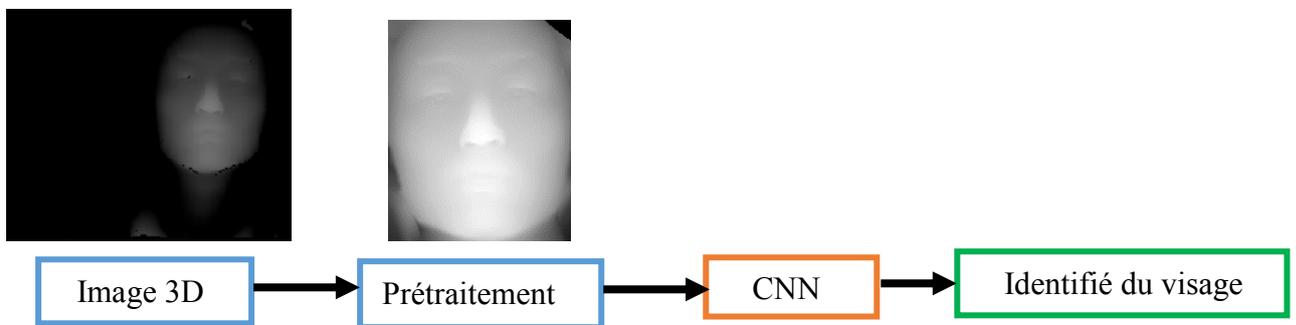


Figure 3.1 : Schéma général de notre système.

3.2.1 Prétraitement d'image

prétraitement est une phase importante dans le processus globale d'identification, C'est une méthode simple qui augmente en général les performances du système . On suppose qu'un visage est décrit par un nuage de points 3D capturés par un scanner laser 3D. Chaque nuage de points est constitué de milliers de points dans l'espace 3D. Ces points discrets décrivent approximativement la surface de la face. Chaque point est décrit avec des coordonnées spatiales 3D et les coordonnées de couleurs RGB correspondantes. Ce prétraitement comporte deux étapes principales, l'alignement des surfaces de visage 3D et l'acquisition de la profondeur et de l'intensité des images.

Dans un premier temps, l'information de courbure locale en chaque point, appelée indice de forme (shape index), est calculée. L'indice de forme est proposé par Dorai et Jain et représente les informations de forme locale des points. L'indice de forme au point p est calculé en utilisant la valeur maximale ($K_{max}(p)$) et la valeur minimale ($K_{min}(p)$) de la courbure locale.

Chapitre 3: Conception du système et résultats

$$\text{Indice de forme } (P) = \frac{1}{2} \frac{1}{\pi} \tan^{-1} \frac{K_{\max}(p) + K_{\min}(p)}{K_{\max}(p) - K_{\min}(p)} \quad (3.1)$$

L'indice de forme est indépendant des coordonnées du système et peut s'adapter à la forme de la surface. La valeur zéro représente une calotte sphérique. Les points dont l'indice de forme est de l'ordre de 0,85 à 1 sont groupés pour former des zones possibles de la pointe du nez. Comme il peut y avoir de petites régions causées par les cheveux ou d'autres facteurs, les petites régions (<70 point) sont enlevées. Après cela, les régions possibles d'extrémité du nez sont étiquetées comme S_{Int} . Notons que ces régions comprennent les zones autour des oreilles, les yeux, les lèvres et le menton. Après cela, deux modèles de pointe du nez sont utilisés. Tout d'abord, un modèle est utilisé pour faire correspondre le visage afin de déterminer quelles sont les régions dans S_{Int} . Qui ressemblent localement à une pointe du nez. Ensuite, pour les régions où le modèle de pointe du nez est sélectionné, nous sélectionnons un modèle de la plus grande pointe du nez est marqué pour sélectionner le réel de pointe du nez p_{nt} .

Une sphère de rayon r_l centrée à la pointe du nez p_{nt} est ensuite utilisée pour recadrer le visage 3D (figure 3.2 (b)) et son visage 2D correspondant (figure 3.3 (b)). Une valeur $r_l = 80$ mm a été choisie. Une fois que le visage est coupé, les points aberrants sont éliminés. Un point aberrant d'observation est défini dans [48]. Les points aberrants sont ceux dont la valeur est supérieure à un seuil d_t de l'un de ses 8 voisins. d_t est calculée automatiquement à l'aide:

$$d_t = \mu + 0.6 \beta \quad (3.2)$$

Où: μ est la valeur moyenne des points voisins, β et son écart-type

La suppression des pointes peut entraîner des trous dans le visage 3D (figure 3.2 (c)) qui sont remplis en utilisant l'interpolation cubique pour combler les données manquantes. Ceci apparaît souvent dans les yeux et les zones de la bouche. Après avoir éliminé les pointes aberrantes, les visages sont acquis à des distances différentes du capteur 3D. Ils ont donc des résolutions différentes. Nous normalisons les images du visage 3D et 2D sur une grille carrée uniforme de 1 mm, afin de nous assurer que tous les visages ont la même résolution. Après découpage, élimination des pointes aberrantes, remplissage des trous et normalisation, nous alignons tous les visages 3D et 2D avec le premier visage 3D (expression neutre) en utilisant l'algorithme ICP (Iterative Closest Point). [9]

Chapitre 3: Conception du système et résultats

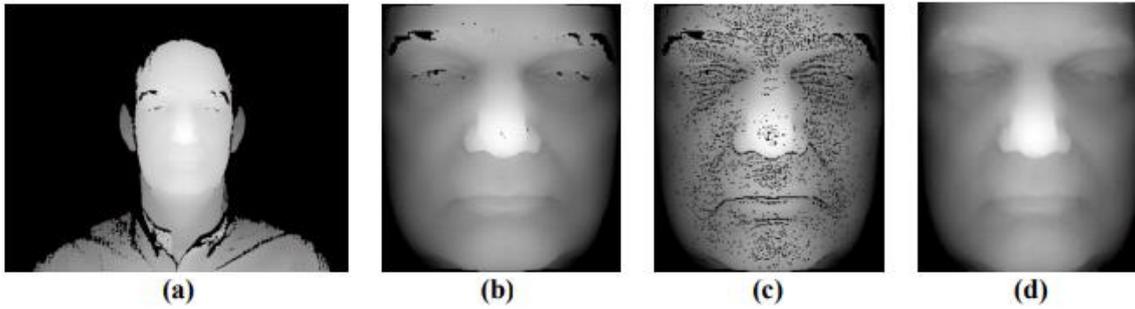


Figure 3.2: Prétraitement de l'image de profondeur: (a) données 3D, (b) image 3D coupée, (c) élimination des pointes aberrantes, (d) image de profondeur après étape de prétraitement. [9]

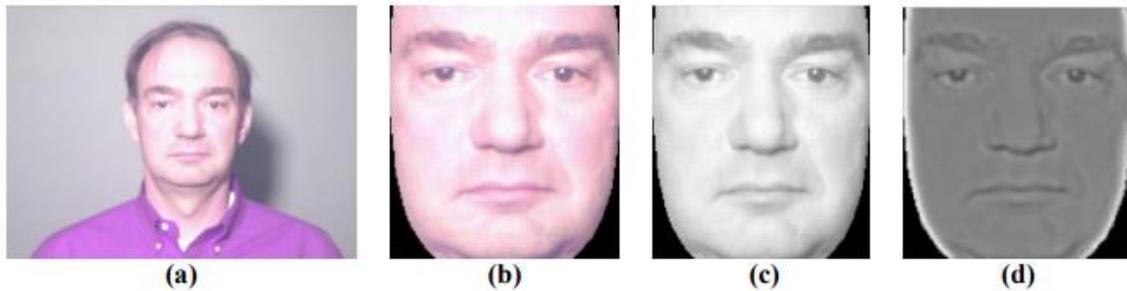


Figure 3.3: Prétraitement de l'image 2D: (a) données 2D, (b) image 2D coupée, (c) image d'intensité, (d) normalisation photométrique par l'approche PS (preprocessing sequence approach). [9]

Suivant alignement par l'ICP, une sphère de rayon $r_2 = 75$ mm centrée sur le bout du nez est utilisée pour recadrer le visage 3D et 2D (figure 3.2.d) et (figure 3.3 .b). Pour l'image couleur RVB (figure 3.3 .b), nous avons utilisé l'image d'intensité (figure 3.3. c). l'influence de la variation d'illumination.

3.3. CNN

Pour reconnaître les visages, nous utilisons des réseaux de neurones à convolution dans la deuxième phase de notre système.

Le premier élément à prendre en compte est le choix d'architecture du CNN. Les architectures de réseaux neuronaux sont des combinaisons de couches qui se transmettent les sorties les unes après les autres.

Nous avons choisi une architecture initiale. Des résultats de performance très différents sont obtenus en fonction du choix et de la séquence des couches constituant les architectures de réseau neuronal. Pour améliorer le modèle et trouver une bonne architecture, nous avons entamé un processus d'expérimentation et de test avec diverses combinaisons.

Chapitre 3: Conception du système et résultats

Notre architecture est constituée de huit couches en tous. Dont sept couches de convolution utilisant des filtres de nombre et tailles différentes et on applique une et une normalisation par lot (BatchNormalization), ensuite des corrections ReLU sont appliquées pour éliminer les valeurs négatives.

À la fin, couche totalement connecté pour calculer le taux reconnaissance.

Les paramètres étudiés dans ce travail sont :

1. Nombre des filtres dans la couche de convolution.
2. Taille des filtres dans la couche de convolution.
3. Le type de Pooling (Max Pooling ou Moyen Pooling).

Le tableau 3.1 présente l'architecture utilisée:

Chapitre 3: Conception du système et résultats

Paramètre des couches	Couches
Image d'entrée (256*256)	1
Convolution2D(nombre des filtres, taille de filtre)	
Normalisation des lots	
ReLU	
Pooling (taille de filtre * taille de filtre)	2
Convolution2D(nombre des filtres, taille de filtre)	
Normalisation des lots	
ReLU	
Pooling (taille de filtre * taille de filtre)	3
Convolution2D(nombre des filtres, taille de filtre)	
Normalisation des lots	
ReLU	
Pooling (taille de filtre * taille de filtre)	4
Convolution2D(nombre des filtres, taille de filtre)	
Normalisation des lots	
ReLU	
Pooling (taille de filtre * taille de filtre)	5
Convolution2D(nombre des filtres, taille de filtre)	
Normalisation des lots	
ReLU	
Pooling (taille de filtre * taille de filtre)	6
Convolution2D(nombre des filtres, taille de filtre)	
Normalisation des lots	
ReLU	
Pooling (taille de filtre * taille de filtre)	7
Convolution2D(nombre des filtres, taille de filtre)	
Normalisation des lots	
ReLU	
Couche totalement connecté ()	8

Tableau 3.1: Architecture initiale de notre CNN

3.4 Base de données CASIA 3D

Pour tester notre système d'identification du visage. Nous utilisons la base de données CASIA 3D.

Entre Août et Septembre 2004, une base de données de visage 3D constituée de 4624 scans de 123 personnes est construite par l'utilisation d'un scanner 3D sans contact (Minolta Vivid 910), comme le montre la figure 3.4 chaque sujet ayant 37 ou 38 images avec les variations individuelles de poses, les expressions et les illuminations, les

Chapitre 3: Conception du système et résultats

variations combinées des expressions sous illumination et pose sous des expressions (figure 3.4, 3.5, 3.6 et 3.7). [9]

À partir de chaque individu, une image couleur 2D et une surface triangulaire 3D du visage sont également générées. On a visé à construire une base de données complète du visage en 3D, qui est en outre amenée à être une plate-forme publique pour tester les algorithmes de reconnaissance faciale 3D ou autres.

Dans notre travail, nous avons étudié les variations d'éclairage, expressions et les changements combinés dans l'expression sous éclairage. Par conséquent, nous avons utilisé 15 images pour chaque personne.

Nous avons utilisé un protocole d'évaluation contient 100 personnes. Pour l'apprentissage de CNN nous utilisons 5 images ou 10 images et pour le test nous utilisons 5 images.



Figure 3.4 : Capture des scènes de la base de données CASIA visages 3D. [9]

Chapitre 3: Conception du système et résultats



Figure 3.5 : Variations d'éclairage de la base de données CASIA visage 3D [9].



(a) Smile

(b) Laugh

(c) anger

(d) surprise

(e) eye close

Figure 3.6 : Variations d'expressions de la base de données CASIA visage 3D.

Chapitre 3: Conception du système et résultats

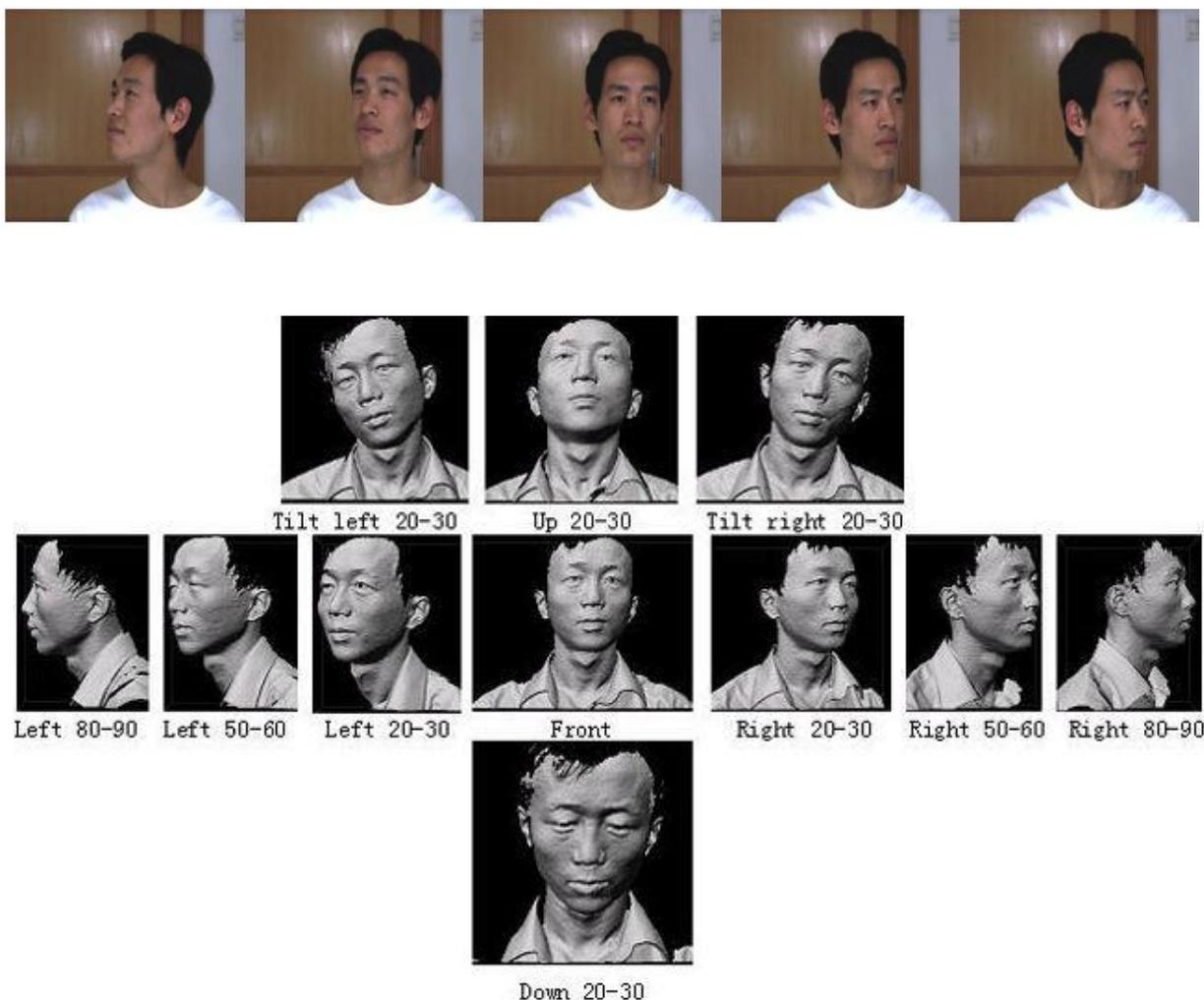


Figure 3.7 : Variations des poses de la base de données CASIA visage 3D. [9]

3.5 Résultats

L'efficacité de notre système d'identification du visage est évaluée dans cette section sur la base de données CASIA 3D.

3.1 Identification uni-modale de visage 3D

Dans cette section, nous comparons l'identification uni-modale de visage 3D dans la base CASIA 3D avec l'utilisation de 5 et 10 images pour l'apprentissage de CNN. Le tableau 3.1 montre le taux de reconnaissance (TR) en utilisant Max Pooling et la taille des filtres dans la phase de convolution varient entre 2 et 6.

Chapitre 3: Conception du système et résultats

	Paramètre Couches		Paramètre Couches		Paramètre Couches	
	CNN 3D Visage	Image d'entrée (256*256)	1	Image d'entrée (256*256)	1	Image d'entrée (256*256)
Conv2D(nombre 8, taille 2)		Conv2D(nombre 8, taille 3)		Conv2D(nombre 8, taille 4)		
Normalisation des lots		Normalisation des lots		Normalisation des lots		
ReLU		ReLU		ReLU		
Max Pooling (2 * 2)		2	Max Pooling (2 * 2)	2	Max Pooling (2 * 2)	2
Conv2D(nombre 16,taille 2)			Conv2D(nombre 16,taille 3)		Conv2D(nombre 16,taille 4)	
Normalisation des lots			Normalisation des lots		Normalisation des lots	
ReLU			ReLU		ReLU	
Max Pooling (2 * 2)	3	Max Pooling (2 * 2)	3	Max Pooling (2 * 2)	3	
Conv2D(nombre 32, taille 2)		Conv2D(nombre 32, taille 3)		Conv2D(nombre 32, taille 4)		
Normalisation des lots		Normalisation des lots		Normalisation des lots		
ReLU		ReLU		ReLU		
Max Pooling (2 * 2)	4	Max Pooling (2 * 2)	4	Max Pooling (2 * 2)	4	
Conv2D(nombre 64, taille 2)		Conv2D(nombre 64, taille 3)		Conv2D(nombre 64, taille 4)		
Normalisation des lots		Normalisation des lots		Normalisation des lots		
ReLU		ReLU		ReLU		
Max Pooling (2 * 2)	5	Max Pooling (2 * 2)	5	Max Pooling (2 * 2)	5	
Conv2D(nombre 128, taille 2)		Conv2D(nombre 128, taille 3)		Conv2D(nombre 128, taille 4)		
Normalisation des lots		Normalisation des lots		Normalisation des lots		
ReLU		ReLU		ReLU		
Max Pooling (2 * 2)	6	Max Pooling (2 * 2)	6	Max Pooling (2 * 2)	6	
Conv2D(nombre 256, taille 2)		Conv2D(nombre 256, taille 3)		Conv2D(nombre 256, taille 4)		
Normalisation des lots		Normalisation des lots		Normalisation des lots		
ReLU		ReLU		ReLU		
Max Pooling (2 * 2)	7	Max Pooling (2 * 2)	7	Max Pooling (2 * 2)	7	
Conv2D(nombre 256,taille 2)		Conv2D(nombre 256,taille 3)		Conv2D(nombre 256,taille 4)		
Normalisation des lots		Normalisation des lots		Normalisation des lots		
ReLU		ReLU		ReLU		
Couche totalement connecté	8	Couche totalement connecté	8	Couche totalement connecté	8	
TR 5images	76.40		76.00		74.00	
TR 10images	92.80		90.80		91.60	

Chapitre 3: Conception du système et résultats

	Paramètre Couches		Paramètre Couches	
	CNN 3D Visage	Image d'entrée (256*256)	1	Image d'entrée (256*256)
Conv2D(nombre 8, taille 5)		Conv2D(nombre 8, taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		
Max Pooling (2 * 2)		2	Max Pooling (2 * 2)	2
Conv2D(nombre 16,taille 5)			Conv2D(nombre 16,taille 6)	
Normalisation des lots			Normalisation des lots	
ReLU			ReLU	
Max Pooling (2 * 2)	3	Max Pooling (2 * 2)	3	
Conv2D(nombre 32, taille 5)		Conv2D(nombre 32, taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		
Max Pooling (2 * 2)	4	Max Pooling (2 * 2)	4	
Conv2D(nombre 64, taille 5)		Conv2D(nombre 64, taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		
Max Pooling (2 * 2)	5	Max Pooling (2 * 2)	5	
Conv2D(nombre 128, taille 5)		Conv2D(nombre 128, taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		
Max Pooling (2 * 2)	6	Max Pooling (2 * 2)	6	
Conv2D(nombre 256, taille 5)		Conv2D(nombre 256, taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		
Max Pooling (2 * 2)	7	Max Pooling (2 * 2)	7	
Conv2D(nombre 256,taille 5)		Conv2D(nombre 256,taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		
Couche totalement connecté	8	Couche totalement connecté	8	
TR 5images	75.40		72.60	
TR 10images	91.20		93.80	

Tableau 3.2 : Taux de reconnaissance (TR) en utilisant Max Pooling et la taille des filtres dans la phase de convolution varient entre 2 et 6 (5 images et 10 images pour l'apprentissage de CNN).

Chapitre 3: Conception du système et résultats

Le tableau 3.2 montre que :

- Le taux de reconnaissance change en parallèle avec la taille des filtres et quand utilisant le Max pooling.
- Quand on a utilisée 5 images le pourcentage varie entre (72 % et 76 %). où il a atteint la limite la plus élevée estimée par TR= 76.40%
- La taille 2 donne le meilleur résultat.
- Et pour 10 images on a vécu une croissance de ce pourcentage qui a été entre (90% et 93%) où il a atteint un maximum de TR= 93.80 % (la taille 6 donner le meilleur résultat).

Le tableau 3.2 montre taux de reconnaissance (TR) en utilisant Moyen Pooling et la taille des filtres dans la phase de convolution varient entre 2 et 6 (5 images et 10 images pour l'apprentissage de CNN).

	Paramètre Couches		Paramètre Couches		Paramètre Couches	
	Image d'entrée (256*256)		Image d'entrée (256*256)		Image d'entrée (256*256)	
CNN 3D Visage	Conv2D(nombre 8, taille 2)	1	Conv2D(nombre 8, taille 3)	1	Conv2D(nombre 8, taille 4)	1
	Normalisation des lots		Normalisation des lots		Normalisation des lots	
	ReLU		ReLU		ReLU	
	Moyen Pooling (2 * 2)		Moyen Pooling (2 * 2)		Moyen Pooling (2 * 2)	
	Conv2D(nombre 16,taille 2)	2	Conv2D(nombre 16,taille 3)	2	Conv2D(nombre 16,taille 4)	2
	Normalisation des lots		Normalisation des lots		Normalisation des lots	
	ReLU		ReLU		ReLU	
	Moyen Pooling (2 * 2)		Moyen Pooling (2 * 2)		Moyen Pooling (2 * 2)	
	Conv2D(nombre 32, taille 2)	3	Conv2D(nombre 32, taille 3)	3	Conv2D(nombre 32, taille 4)	3
	Normalisation des lots		Normalisation des lots		Normalisation des lots	
	ReLU		ReLU		ReLU	
	Moyen Pooling (2 * 2)		Moyen Pooling (2 * 2)		Moyen Pooling (2 * 2)	
	Conv2D(nombre 64, taille 2)	4	Conv2D(nombre 64, taille 3)	4	Conv2D(nombre 64, taille 4)	4
	Normalisation des lots		Normalisation des lots		Normalisation des lots	
	ReLU		ReLU		ReLU	
Moyen Pooling (2 * 2)	Moyen Pooling (2 * 2)		Moyen Pooling (2 * 2)			
Conv2D(nombre 128, taille 2)	5	Conv2D(nombre 128, taille 3)	5	Conv2D(nombre 128, taille 4)	5	
Normalisation des lots		Normalisation des lots		Normalisation des lots		
ReLU		ReLU		ReLU		
Moyen Pooling (2 * 2)		Moyen Pooling (2 * 2)		Moyen Pooling (2 * 2)		

Chapitre 3: Conception du système et résultats

	<table border="1"> <tr><td>Moyen Pooling (2 * 2)</td><td rowspan="4">6</td></tr> <tr><td>Conv2D(nombre 256, taille 2)</td></tr> <tr><td>Normalisation des lots</td></tr> <tr><td>ReLU</td></tr> <tr><td>Moyen Pooling (2 * 2)</td><td rowspan="4">7</td></tr> <tr><td>Conv2D(nombre 256,taille 2)</td></tr> <tr><td>Normalisation des lots</td></tr> <tr><td>ReLU</td></tr> <tr><td>Couche totalement connecté</td><td>8</td></tr> </table>	Moyen Pooling (2 * 2)	6	Conv2D(nombre 256, taille 2)	Normalisation des lots	ReLU	Moyen Pooling (2 * 2)	7	Conv2D(nombre 256,taille 2)	Normalisation des lots	ReLU	Couche totalement connecté	8	<table border="1"> <tr><td>Moyen Pooling (2 * 2)</td><td rowspan="4">6</td></tr> <tr><td>Conv2D(nombre 256, taille 3)</td></tr> <tr><td>Normalisation des lots</td></tr> <tr><td>ReLU</td></tr> <tr><td>Moyen Pooling (2 * 2)</td><td rowspan="4">7</td></tr> <tr><td>Conv2D(nombre 256,taille 3)</td></tr> <tr><td>Normalisation des lots</td></tr> <tr><td>ReLU</td></tr> <tr><td>Couche totalement connecté</td><td>8</td></tr> </table>	Moyen Pooling (2 * 2)	6	Conv2D(nombre 256, taille 3)	Normalisation des lots	ReLU	Moyen Pooling (2 * 2)	7	Conv2D(nombre 256,taille 3)	Normalisation des lots	ReLU	Couche totalement connecté	8	<table border="1"> <tr><td>Moyen Pooling (2 * 2)</td><td rowspan="4">6</td></tr> <tr><td>Conv2D(nombre 256, taille 4)</td></tr> <tr><td>Normalisation des lots</td></tr> <tr><td>ReLU</td></tr> <tr><td>Moyen Pooling (2 * 2)</td><td rowspan="4">7</td></tr> <tr><td>Conv2D(nombre 256,taille 4)</td></tr> <tr><td>Normalisation des lots</td></tr> <tr><td>ReLU</td></tr> <tr><td>Couche totalement connecté</td><td>8</td></tr> </table>	Moyen Pooling (2 * 2)	6	Conv2D(nombre 256, taille 4)	Normalisation des lots	ReLU	Moyen Pooling (2 * 2)	7	Conv2D(nombre 256,taille 4)	Normalisation des lots	ReLU	Couche totalement connecté	8
Moyen Pooling (2 * 2)	6																																						
Conv2D(nombre 256, taille 2)																																							
Normalisation des lots																																							
ReLU																																							
Moyen Pooling (2 * 2)	7																																						
Conv2D(nombre 256,taille 2)																																							
Normalisation des lots																																							
ReLU																																							
Couche totalement connecté	8																																						
Moyen Pooling (2 * 2)	6																																						
Conv2D(nombre 256, taille 3)																																							
Normalisation des lots																																							
ReLU																																							
Moyen Pooling (2 * 2)	7																																						
Conv2D(nombre 256,taille 3)																																							
Normalisation des lots																																							
ReLU																																							
Couche totalement connecté	8																																						
Moyen Pooling (2 * 2)	6																																						
Conv2D(nombre 256, taille 4)																																							
Normalisation des lots																																							
ReLU																																							
Moyen Pooling (2 * 2)	7																																						
Conv2D(nombre 256,taille 4)																																							
Normalisation des lots																																							
ReLU																																							
Couche totalement connecté	8																																						
TR 5images	77.20	81.80	76.60																																				
TR 10images	91.40	91.80	91.60																																				

	Paramètre Couches		Paramètre Couches	
	CNN 3D Visage	Image d'entrée (256*256)	1	Image d'entrée (256*256)
Conv2D(nombre 8, taille 5)		Conv2D(nombre 8, taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		
Moyen Pooling (2 * 2)		2	Moyen Pooling (2 * 2)	2
Conv2D(nombre 16,taille 5)			Conv2D(nombre 16,taille 6)	
Normalisation des lots			Normalisation des lots	
ReLU			ReLU	
Moyen Pooling (2 * 2)		3	Moyen Pooling (2 * 2)	3
Conv2D(nombre 32, taille 5)			Conv2D(nombre 32, taille 6)	
Normalisation des lots	Normalisation des lots			
ReLU	ReLU			
Moyen Pooling (2 * 2)	4	Moyen Pooling (2 * 2)	4	
Conv2D(nombre 64, taille 5)		Conv2D(nombre 64, taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		
Moyen Pooling (2 * 2)	5	Moyen Pooling (2 * 2)	5	
Conv2D(nombre 128, taille 5)		Conv2D(nombre 128, taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		

Chapitre 3: Conception du système et résultats

	<table border="1"> <tbody> <tr> <td>Moyen Pooling (2 * 2)</td> <td rowspan="4" style="text-align: center; vertical-align: middle;">6</td> </tr> <tr> <td>Conv2D(nombre 256, taille 5)</td> </tr> <tr> <td>Normalisation des lots</td> </tr> <tr> <td>ReLU</td> </tr> <tr> <td>Moyen Pooling (2 * 2)</td> <td rowspan="4" style="text-align: center; vertical-align: middle;">7</td> </tr> <tr> <td>Conv2D(nombre 256, taille 5)</td> </tr> <tr> <td>Normalisation des lots</td> </tr> <tr> <td>ReLU</td> </tr> <tr> <td>Couche totalement connecté</td> <td style="text-align: center; vertical-align: middle;">8</td> </tr> </tbody> </table>	Moyen Pooling (2 * 2)	6	Conv2D(nombre 256, taille 5)	Normalisation des lots	ReLU	Moyen Pooling (2 * 2)	7	Conv2D(nombre 256, taille 5)	Normalisation des lots	ReLU	Couche totalement connecté	8	<table border="1"> <tbody> <tr> <td>Moyen Pooling (2 * 2)</td> <td rowspan="4" style="text-align: center; vertical-align: middle;">6</td> </tr> <tr> <td>Conv2D(nombre 256, taille 6)</td> </tr> <tr> <td>Normalisation des lots</td> </tr> <tr> <td>ReLU</td> </tr> <tr> <td>Moyen Pooling (2 * 2)</td> <td rowspan="4" style="text-align: center; vertical-align: middle;">7</td> </tr> <tr> <td>Conv2D(nombre 256, taille 6)</td> </tr> <tr> <td>Normalisation des lots</td> </tr> <tr> <td>ReLU</td> </tr> <tr> <td>Couche totalement connecté</td> <td style="text-align: center; vertical-align: middle;">8</td> </tr> </tbody> </table>	Moyen Pooling (2 * 2)	6	Conv2D(nombre 256, taille 6)	Normalisation des lots	ReLU	Moyen Pooling (2 * 2)	7	Conv2D(nombre 256, taille 6)	Normalisation des lots	ReLU	Couche totalement connecté	8
Moyen Pooling (2 * 2)	6																									
Conv2D(nombre 256, taille 5)																										
Normalisation des lots																										
ReLU																										
Moyen Pooling (2 * 2)	7																									
Conv2D(nombre 256, taille 5)																										
Normalisation des lots																										
ReLU																										
Couche totalement connecté	8																									
Moyen Pooling (2 * 2)	6																									
Conv2D(nombre 256, taille 6)																										
Normalisation des lots																										
ReLU																										
Moyen Pooling (2 * 2)	7																									
Conv2D(nombre 256, taille 6)																										
Normalisation des lots																										
ReLU																										
Couche totalement connecté	8																									
TR 5images	71.40	74.20																								
TR 10images	91.40	91.80																								

Tableau 3.3 : Taux de reconnaissance (TR) en utilisant Moyen Pooling et la taille des filtres dans la phase de convolution varient entre 2 et 6 (5 images et 10 images pour l'apprentissage de CNN).

Chapitre 3: Conception du système et résultats

Dans le troisième tableau, on note

- Dans le cas de l'utilisation de 5 images on a remarqué un changement de TR (Taux de reconnaissance) où nous obtenons un taux entre 71% et 81%.
- La taille 3 nous a donné le meilleur résultat est 81.80%.
- Et pour le 2^{ème} cas de 10 images le pourcentage était convergé à 91%.
- Lors de l'utilisation de 10 images, le résultat était beaucoup mieux que 5 images.

3.4.2 Identification multimodales (2D + 3D) de visage

Le tableau présente le taux de reconnaissance (TR) de visage 3D et 2D+3D.

Ce tableau monte que :

- Lorsque nous avons utilisé l'image 3D, le résultat était 93.80 . Soit nous avons utilisé (2D+3D) et le résultat était meilleur 96% .
- La vérification de visage 2D+3D est toujours meilleure que la vérification de visage 3D.
- L'identification multimodales (2D+3D) de visage beaucoup mieux que la Identification uni-modale de visage 3D.

Chapitre 3: Conception du système et résultats

	Image 3D		Image 2D+3D	
	Paramètre	Couches	Paramètre	Couches
CNN 3D Visage Et CNN 2D+3D Visage	Image d'entrée (256*256*4)	1	Image d'entrée (256*256*4)	1
	Conv2D(nombre 8, taille 6)		Conv2D(nombre 8, taille 6)	
	Normalisation des lots		Normalisation des lots	
	ReLU		ReLU	
	Max Pooling (2 * 2)	2	Max Pooling (2 * 2)	2
	Conv2D(nombre 16,taille 6)		Conv2D(nombre 16,taille 6)	
	Normalisation des lots		Normalisation des lots	
	ReLU		ReLU	
Max Pooling (2 * 2)	3	Max Pooling (2 * 2)	3	
Conv2D(nombre 32, taille 6)		Conv2D(nombre 32, taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		
Max Pooling (2 * 2)	4	Max Pooling (2 * 2)	4	
Conv2D(nombre 64, taille 6)		Conv2D(nombre 64, taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		
Max Pooling (2 * 2)	5	Max Pooling (2 * 2)	5	
Conv2D(nombre 128, taille 6)		Conv2D(nombre 128, taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		
Max Pooling (2 * 2)	6	Max Pooling (2 * 2)	6	
Conv2D(nombre 256, taille 6)		Conv2D(nombre 256, taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		
Max Pooling (2 * 2)	7	Max Pooling (2 * 2)	7	
Conv2D(nombre 256,taille 6)		Conv2D(nombre 256,taille 6)		
Normalisation des lots		Normalisation des lots		
ReLU		ReLU		
	Couche totalement connecté	8	Couche totalement connecté	8
TR 10images	93.80		96	

Tableau 3.4 : Taux de reconnaissance (TR) de visage 3D et 2D+3D.

Chapitre 3: Conception du système et résultats

3.5 Conclusion

Dans ce chapitre, nous avons présenté les résultats obtenus en effectuant une étude comparative tout d'abord sur les deux systèmes d'identification uni-modale de visage 3D et multimodales (2D et 3D). Et finalement nous avons conclu que l'identification multimodales (2D et 3D) donne meilleurs résultats que d'identification uni-modale de visages 3D.

*Conclusion
Générale*

Conclusion générale

la reconnaissance faciale est l'une des rares méthodes biométriques qui possèdent les mérites d'une grande précision et d'une faible intrusion. Il a la précision d'une approche physiologique sans être intrusif. Pour cette raison, depuis le début des années, la reconnaissance faciale a attiré l'attention des chercheurs dans les domaines de la sécurité, de la psychologie et du traitement d'images, à la vision par ordinateur. De nombreux algorithmes ont été proposés pour la reconnaissance faciale.

Dans Cette mémoire , nous sommes intéressés au problème de la reconnaissance faciale dont le but d'implémenter un algorithme puissant pour identifier l'individu par l'intermédiaire de son visage. en utilisant les réseaux de neurones à convolution.

A travers le premier chapitre, nous avons noté que les problèmes et difficultés de reconnaissance de visage 2D qui sont Changement d'illumination (éclairage), Pose , Expressions faciales , et Occlusions , et aussi État de l'art des différentes approches récentes en reconnaissance de visages 3D .

Dans la deuxième étape , nous avons discuté des notions fondamentales des réseaux de neurones à convolution en présentant les différents types de couches utilisées: la couche convolutionnelle, la couche de rectification RELU , la couche de pooling et la couche fully connected . et les applications de CNN .

Dans la dernière étape , nous avons présenté les résultats obtenus en effectuant une étude des différentes étapes de notre système en utilisant la base de données CASIA 3D . Pour la phase de reconnaissance on a utilisé le CNN , afin de vérifier l'identité de la personne en entrant.

Les résultats de cette mémoire montrent :

- Lors de l'utilisation de 10 images, le résultat était beaucoup mieux que 5 images.
- La vérification de visage 2D+3D est toujours meilleure que la vérification de visage 3D.

- L'identification multimodales (2D+3D) de visage beaucoup mieux que la Identification uni-modale de visage 3D.
- D'après les résultats obtenus et présentés dans ce chapitre nous remarquons que le meilleur TR dans la base d'évaluation est obtenu pour L'identification multimodales (2D+3D) de visage avec un TR = 96% .

Perspectives :

Les perspectives d'évolution de ce travail sont :

- Appliquer notre système sur d'autre base de données plus grand que la base de données CASIA.

Bibliographie

Bibliographie

[I] HU, H., SHAH, S. A. A., BENNAMOUN, M., & MOLTON, M. (2017, NOVEMBER). 2D AND 3D FACE RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK. IN TENCON 2017-2017 IEEE REGION 10 CONFERENCE (PP. 133-132). IEEE.

[II] KAMENCAY, P., BENČO, M., MIŽDOŠ, T., & RADIL, R. (2017). A NEW METHOD FOR FACE RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK.

[III] HU, H., SHAH, S. A. A., BENNAMOUN, M., & MOLTON, M. (2017, NOVEMBER). 2D AND 3D FACE RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK. IN TENCON 2017-2017 IEEE REGION 10 CONFERENCE (PP. 133-132). IEEE.

[IV] [HTTPS://WWW.FUTURA-SCIENCES.COM/TECH/DEFINITIONS/INTELLIGENCE-ARTIFICIELLE-DEEP-LEARNING-17262/](https://www.futura-sciences.com/tech/definitions/intelligence-artificielle-deep-learning-17262/)

[1] PERRONNIN, F., & DUGELAY, J. L. (2002). INTRODUCTION A LA BIOMETRIE-AUTHENTIFICATION DES INDIVIDUS PAR TRAITEMENT AUDIO-VIDEO. TRAITEMENT DU SIGNAL, 19(4).

[2] R. HEITMEYER, "BIOMETRIC IDENTIFICATION PROMISES FAST AND SECURE PROCESSING OF AIRLINE PASSENGERS," ICAO JOURNAL., VOL. 55, 2000]

[3] [HTTPS://WWW.LEBIGDATA.FR/RECONNAISSANCE-FACIALE-TOUT-SAVOIR](https://www.lebigdata.fr/reconnaissance-faciale-tout-savoir)

[4] A.CHOUCHANE. ANALYSE D'IMAGES D'EXPRESSIONS FACIALES ET ORIENTATION DE LA TETE BASEE SUR LA PROFONDEUR. THESE PRESENTEE EN VUE DE L'OBTENTION DU DIPLOME DE DOCTORAT LMD EN GENIE ELECTRIQUE, 2016

[5] A. CHAARI : « NOUVELLE APPROCHE D'IDENTIFICATION DANS LES BASES DE DONNEES BIOMETRIQUES BASEE SUR UNE CLASSIFICATION NON SUPERVISEE ». THESE DE DOCTORAT, UNIVERSITE D'EVRY VAL D'ESSONNE. SOUTENUE LE 06/10/ 2009.

[6] BUYSENS, PIERRE. FUSION DE DIFFERENTS MODES DE CAPTURE POUR LA RECONNAISSANCE DU VISAGE APPLIQUEE AUX E_TRANSACTIONS. DISS. 2011.

[7] T. C. Faltemier. Flexible and robust 3D face recognition. PhD thesis, The University of Notre Dame, 2007

[8] MORIZET, N., EA, T., ROSSANT, F., AMIEL, F., & AMARA, A. (2006). REVUE DES ALGORITHMES PCA LDA ET EBGMM UTILISES EN RECONNAISSANCE 2D DU VISAGE POUR LA BIOMETRIE. P1-11. INSTITUT SUPERIEUR D' ELECTRONIQUE DE PARIS (ISEP), DEPARTEMENT D' ELECTRONIQUE

[9] OUAMANE, A. (2015). RECONNAISSANCE BIOMETRIQUE PAR FUSION MULTIMODALE DU VISAGE 2D ET 3D (DOCTORAL DISSERTATION, UNIVERSITE MOHAMED KHIDER-BISKRA).

[10] M. KIRBY AND L. SIROVICH. APPLICATION OF THE KARHUNEN-LOEVE PROCEDURE FOR THE CHARACTERIZATION OF HUMAN FACES. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, PAMI-12(1) :103–108, JANUARY 1990

[11] D. J. KRIEGMAN, J. P. HESPANHA, AND P. N. BELHUMEUR. EIGENFACES VS. FISHERFACES : RECOGNITION USING CLASS-SPECIFIC LINEAR PROJECTION. IN EUROPEAN CONFERENCE ON COMPUTER VISION, PAGES I:43–58, 1996

- [12] M. S. BARTLETT, J. R. MOVELLAN, AND T. J. SEJNOWSKI. FACE RECOGNITION BY INDEPENDENT COMPONENT ANALYSIS. *TRANSACTIONS ON NEURAL NETWORKS*, AUGUST 15 2002
- [13] I. BUCIU AND I. PITAS. APPLICATION OF NON-NEGATIVE AND LOCAL NON NEGATIVE MATRIX FACTORIZATION TO FACIAL EXPRESSION RECOGNITION. IN *INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION*, PAGES I : 288–291, 2004
- [14] Y. WANG, Y. JIA, C. HU, AND M. TURK. NON-NEGATIVE MATRIX FACTORIZATION FRAME WORK FOR FACE RECOGNITION. *INTERNATION JOURNAL OF PATTERN RECOGNITION AND ARTIFICIAL INTELLIGENCE*, 19(4) :495–511, 2005.
- [15] M. VISANI, C. GARCIA, AND J. M. JOLION. NORMALIZED RADIAL BASIS FUNCTION NETWORKS AND BILINEAR DISCRIMINANT ANALYSIS FOR FACE RECOGNITION. IN *IEEE CONFERENCE ON ADVANCED VIDEO AND SIGNAL BASED SURVEILLANCE*, PAGES 342– 347, 2005.
- [16] H. CEVIKALP, M. NEAMTU, M. WILKES, AND A. BARKANA. DISCRIMINATIVE COMMON VECTORS FOR FACE RECOGNITION. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 27(1) :4–13, 2005
- [17] A. PENTLAND, B. MOGHADDAM, AND T. STARNER. VIEW-BASED AND MODULAR EIGENSPACES FOR FACE RECOGNITION. IN *IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION*, 1994.
- [18] H. KIM, H. PARK, AND H. ZHA. DISTANCE PRESERVING DIMENSION REDUCTION FOR MANIFOLD LEARNING. IN *INTERNATIONAL CONFERENCE ON DATA MINING*. SIAM, 2007.
- [19] S. BISWAS, K. W. BOWYER, AND PATRICK J. FLYNN. MULTIDIMENSIONAL SCALING FOR MATCHING LOW-RESOLUTION FACIAL IMAGES. IN *BIOMETRICS : THEORY, APPLICATIONS AND SYSTEMS*, 2010.
- [20] M. H. YANG. FACE RECOGNITION USING EXTENDED ISOMAP. IN *INTERNATIONAL CONFERENCE ON IMAGE PROCESSING*, PAGES 117–120, 2002
- [21] G. HAGEN, T. SMITH, A. BANASUK, R.R. COIFMAN, AND I. MEZIC. VALIDATION OF LOW-DIMENSIONAL MODELS USING DIFFUSION MAPS AND HARMONIC AVERAGING. IN *IEEE CONFERENCE ON DECISION AND CONTROL*, 2007

- [22] G. HAGEN, T. SMITH, A. BANASUK, R.R. COIFMAN, AND I. MEZIC. VALIDATION OF LOW-DIMENSIONAL MODELS USING DIFFUSION MAPS AND HARMONIC AVERAGING. IN IEEE CONFERENCE ON DECISION AND CONTROL, 2007
- [23] S.PANG AND N.KASABOV. INVESTIGATING LLE EIGENFACE ON POSE AND FACE IDENTIFICATION. IN JUN WANG, ZHANG YI, JACEK M. ZURADA, BAO-LIANG LU, AND HUIJUN YIN, EDITORS, INTERNATIONAL SYMPOSIUM ON NEURAL NETWORKS, VOLUME 3972 OF LECTURE NOTES IN COMPUTER SCIENCE, PAGES 134–139, 2006
- [24] X. HE, S. YAN, Y. HU, P. NIYOGI, AND H. ZHANG. FACE RECOGNITION USING LAPLACIAN FACES. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 27(3) :328–340, 2005.
- [25] Y. PANG, Z. LIU, AND Y. SUN. SUBSPACE LEARNING BASED ON LAPLACIAN EIGENMAPS AND LDA FOR FACE RECOGNITION. INTERNATIONAL JOURNAL ON INFORMATION ACQUISITION, 3(1) :45–51, 2006.
- [26] B. RADUCANU AND F. DORNAIKA. DYNAMIC FACIAL EXPRESSION RECOGNITION USING LAPLACIAN EIGENMAPS-BASED MANIFOLD LEARNING. IN INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, PAGES 156–161. IEEE, 2010.
- [27] H. KIM, H. PARK, AND H. ZHA. DISTANCE PRESERVING DIMENSION REDUCTION FOR MANIFOLD LEARNING. IN INTERNATIONAL CONFERENCE ON DATA MINING. SIAM, 2007
- [28] Q. WANG AND J. LI. COMBINING LOCAL AND GLOBAL INFORMATION FOR NONLINEAR DIMENSIONALITY REDUCTION. NEUROCOMPUTING, 72(10-12) :2235–2241, 2009
- [29] M. N. TELI. DIMENSIONALITY REDUCTION USING NEURAL NETWORKS. TECHNICAL REPORT, APRIL 02 2008
- [30] M. K. FLEMING AND G. W. COTTRELL. CATEGORIZATION OF FACES USING UNSUPERVISED FEATURE EXTRACTION. IN IEEE INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS, VOLUME II, PAGES II–65–II–70, SAN DIEGO, 1990. IEEE. UCSD
- [31] S. LAWRENCE, C. L. GILES, A. C. TSOI, AND A. D. BACK. FACE RECOGNITION : A CONVOLUTIONAL NEURAL-NETWORK APPROACH. IEEE TRANSACTIONS ON NEURAL NETWORKS, 8(1) :98–113, JANUARY 1997.

- [32] S. DUFFNER AND C. GARCIA. FACE RECOGNITION USING NON-LINEAR IMAGE RECONSTRUCTION. IN I-LIDS : BAG AND VEHICLE DETECTION CHALLENGE, PAGES 459–464, 2007.
- [33] R. BRUNELLI AND T. POGGIO. FACE RECOGNITION : FEATURES VERSUS TEMPLATES. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 15(10) :1042– 1052, 1993
- [34] B. HEISELE, P. HO, J. WU, AND T. POGGIO. FACE RECOGNITION : COMPONENT-BASED VERSUS GLOBAL APPROACHES. COMPUTER VISION AND IMAGE UNDERSTANDING, 91(1– 2) :6– 21, JULY/AUGUST 2003.
- [35] J. R. PRICE AND T. F. GEE. FACE RECOGNITION USING DIRECT, WEIGHTED LINEAR DISCRIMINANT ANALYSIS AND MODULAR SUBSPACES. PATTERN RECOGNITION, 38(2) :209– 219, FEBRUARY 2005
- [36] T. F. COOTES, G. J. EDWARDS, AND C. J. TAYLOR. ACTIVE APPEARANCE MODELS . IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 23(6) :681–685, JUNE 2001
- [37] D. HUANG, ROBUST FACE RECOGNITION BASED ON THREE DIMENSIONAL DATA, THÈSE DE L'ÉCOLE CENTRAL DE LYON SPECIALITE: INFORMATIQUE, 2011.
- [38] S. JAHANBIN, H. CHOI, Y. LIU AND A. C. BOVIK. THREE DIMENSIONAL FACE RECOGNITION USING ISOGEODESIC AND ISO-DEPTH CURVES. IN IEEE INTERNATIONAL CONFERENCE ON BIOMETRICS: THEORY, APPLICATIONS AND SYSTEMS, PP. 1-6, OCT. 2008.
- [39] GANDIKOTA, ROHIT, AND DEEPAK MISHRA. "HOW YOU SEE ME: UNDERSTANDING CONVOLUTIONAL NEURAL NETWORKS." TENCON 2019-2019 IEEE REGION 10 CONFERENCE (TENCON). IEEE, 2019. .
- [40] CHABOT, FLORIAN. ANALYSE FINE 2D/3D DE VEHICULES PAR RESEAUX DE NEURONES PROFONDS. DISS. 2017.
- [41] KAMENCAY, PATRIK, ET AL. "A NEW METHOD FOR FACE RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK." (2017).

[42] ROSEBROCK, ADRIAN. DEEP LEARNING FOR COMPUTER VISION WITH PYTHON: IMAGENET BUNDLE. PYIMAGESEARCH, 2017.]

[43] <https://www.datasciencetoday.net/>

[44] ALBAWI, SAAD, TAREQ ABED MOHAMMED, AND SAAD AL-ZAWI. "UNDERSTANDING OF A CONVOLUTIONAL NEURAL NETWORK." 2017 INTERNATIONAL CONFERENCE ON ENGINEERING AND TECHNOLOGY (ICET). IEEE, 2017.

[45] AGGARWAL, CHARU C. "NEURAL NETWORKS AND DEEP LEARNING." SPRINGER 10 (2018): 978-3.

[46] BHANDARE, A., BHIDE, M., GOKHALE, P., & CHANDAVARKAR, R. (2016). APPLICATIONS OF CONVOLUTIONAL NEURAL NETWORKS. INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGIES, 7(5), 2206-2215.

[47] [HTTPS://WWW.FLATWORLDSOLUTIONS.COM/DATA-SCIENCE/ARTICLES/7-APPLICATIONS-OF-CONVOLUTIONAL-NEURAL-NETWORKS.PHP](https://www.flatworldsolutions.com/data-science/articles/7-applications-of-convolutional-neural-networks.php)

[48] F. R. AL-OSAIMI, M. BENNAMOUN, AND A. MIAN. SPATIALLY OPTIMIZED DATA-LEVEL FUSION OF TEXTURE AND SHAPE FOR FACE RECOGNITION. IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 21, NO. 2, 2012.