



Université Mohamed Khider de Biskra
Faculté des Sciences et de la Technologie
Département de génie électrique

MÉMOIRE DE MASTER

Sciences et Technologies
Télécommunication
Réseaux et Télécommunication

Réf. : Entrez la référence du document

Présenté et soutenu par :
MARREF Djouhaine Messaoud LEHRAKI Firess

Le : mercredi 22 juin 2022

Estimation de fatigue et somnolence des conducteurs

Jury :

Mr. BAARIR ZINE-Eddine	Pr	Université de Biskra	Président
Mr. OUAMANE Abdelmalik	Pr	Université de Biskra	Examineur
Mr. OUAFI Abdelkrim	Pr	Université de Biskra	Encadreur



Université Mohamed Khider de Biskra
Faculté des Sciences et de la Technologie
Département de génie électrique

MÉMOIRE DE MASTER

Sciences et Technologies
Télécommunication
Réseaux et Télécommunication

Réf. : Entrez la référence du document

Estimation de fatigue et somnolence des conducteurs

Le :

Présenté par :

MARREF Djouhaine Messaoud
LEHRAKI Firess

Avis favorable de l'encadreur :

OUAFI Abdelkrim

Signature Avis favorable du Président du Jury

Cachet et signature

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Remerciement

J'adresse mes sincères remerciements à tous les intervenants et toutes les

Personnes qui par leurs paroles, leurs contributions dans ce travail, leurs écrits, leurs conseils et leurs critiques ont guidé mes réflexions et ont accepté de répondre à mes questions durant mes recherches. Je ne peux achever ce projet, sans exprimer mes sincères gratitudes à tous les professeurs de notre faculté, pour leur dévouement et leur assistance tout au long de mon étude.

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.

Dédicace

*Au nom de dieu le clément et le miséricordieux louange à ALLAH le
tout puissant.*

*Je dédie ce modeste travail en signe de respect, reconnaissance et de
remerciement :*

A mes parents :

*Sans leurs soutiens et leurs conseils, mes accomplissements
n'auraient pas eu lieu, ils ont été derrière moi dans chacun de mes
pas tout au long de ma vie, ma plus profonde gratitude leurs ai
exprimé, aucun mot ne pourrait qualifier l'estime que je leur porte ni
le bien qu'ils m'ont fait, apporter et donner.*

*A ma belle famille Mes frères et soeurs que j'aime en particulier ma
belles soeurs DR. Faiza L et DR. Nadia L merci pour votre soutien et
vos encouragements*

A mon cher Oubai, qui est allongé à l'hôpital pour se faire soigner

*A mon amies et toutes les personnes qui ont contribué de près ou de
loin à la réalisation de ce travail.*

Firas lehraki

Dédicaces

Je dédie cet événement marquant a mon père, celui qui
m'a toujours encouragé et soutenu dans mes dur
moments.

A ma vie, mon adorable mère, a celle qui est
toujours présente et continue de l'être pour faire mon
bonheur. Merci pour ce sacrifice afin que tes enfants
grandissent

Merci tout
Simplement d'être...ma mère.
A mon frère, mon bras droit Lokman (Kimo)

Mes sœurs, Rahma et son époux Riad et leurs enfants Ghaith et Malik
merci d'être la grande sœur que tu es.

Hibat errahmen ma deuxième maman, c'est vers toi ma sœur que je me
tourne quand ça va mal. Tu es très importante pour moi

A ma jumelle, Jina le plus beau cadeau que mes parents m'aient fait, c'était
toi

A toutes ma famille, merci d'être les meilleures personnes au monde.

Marref Djouhain messaoud

SOMMAIRE	
LIST DES FIGURES	
LIST DES TABLEAUX	
INTRODUCTION GENERALE	1
CHAPTER I: Les généralités des expressions faciales	
1. Introduction	2
2. Historique sur la reconnaissance de l'expression faciale	2
3. Définition de visage	3
3.1. Les étapes de la reconnaissance Faciale	3
3.3.a Détection de visage	4
3.3.b. Extraction des caractéristiques du visage	4
3.3.c. La reconnaissance de visage	4
4. Les difficultés de la reconnaissance des visages	5
5. Différents types des expressions faciales	5
6. Système de reconnaissance des expressions faciales	6
7. Les 7 expressions faciales universelles	7
7.1. La joie (joy)- facile à reconnaître, difficile à feindre	7
7.2. La tristesse (sad)	8
7.3. La colère (angry)	8
7.4. Le mépris (contempt)	9
7.5. Le dégoût (disgust)	9
7.6. La Peur (afraid)	10
7.7. La surprise (surprise)	11
8. La somnolence	11
8.1. Relation entre la somnolence et la conduite	11
8.2. Approches de détection de somnolence	12
8.2.1. Approches axées sur les signaux physiologiques	12
8.2.2. Approches axées sur le traitement d'image	13
9. Techniques de reconnaissance faciale	13
10. Avantages et inconvénients de la reconnaissance faciale	14
11. Conclusion	14
CHAPITRE II : Les différentes méthodes de la reconnaissance des expressions facile	
1. Introduction	15

2. Le système de reconnaissance des expressions faciales	15
2.1 L'Acquisition de visage	16
2.2 Prétraitements	16
2.3 Détection du visage	16
2.4. Extraction des données caractéristique	17
2.4.1. Les différentes méthodes développées pour la reconnaissance faciale	17
a. Méthodes globales	17
b. Méthodes locales	18
Algorithme LBP (Local Binary Pattern)	19
Histogram of oriented gradients	20
Calcul de vecteur des caractéristiques Histogrammes	20
c. Méthodes hybrides	21
3. Synthèse des travaux développés pour la reconnaissance d'expressions faciales	22
4. Réseaux de neurones	23
4.1. Définition	23
4.2. Topologie	23
4.3. L'apprentissage en profondeur (deep Learning)	23
4.3.1 Définition	24
4.3.2 Pourquoi le choix deep Learning	25
4.4. Les réseaux de neurones convolutifs	25
4.4.1. Modèle génératif	26
4.4.2. Architecture du système	27
4.4.3 Les différentes couches de CNN	28
a) L'étape de convolution	28
b) Les différentes convolutions	29
c) Fonctions d'activation	30
d) L'étape de pooling	30
e) La couche totalement connectée	32
5. Conclusion	32
CHAPITRE III: Implémentation et résultats expérimentaux	
1. Introduction	33
2. Environnement de développement	33

2.1. Kaggel	33
2.2. Pourquoi les GPU?	34
2.3. Python	34
3. Bibliothèques utilisées	35
3.1. OpenCV (Open Source Computer Vision Library)	35
3.2. Numpy	35
3.3. Matplotlib	35
3.4. Keras	35
3.5. Tensor Flow	35
3.6. Pandas	36
4. BDD	36
5. Implémentation et résultats	37
5.1. Formation du model avec keras	38
6. Résultats expérimentaux	38
6.1. Notre modèle	38
6.2. Les résultats de l'entraînement du modèle	40
6.3. Les résultats affichés	40
6.4. Le teste	41
6.4.1. L'œil fermé	41
6.4.2. L'œil ouverte	43
7. Conclusion	45
Conclusion générale	46
Références bibliographiques	47
Résumé	

Liste Des Figures :

Chapitre I	
Figure 1: Architecture d'un système de reconnaissance des expressions faciales	6
Figure 2: Les différentes méthodes de détection de somnolence	12
Figure 3: illustration des capteurs utilisés dans les méthodes axées sur la performance du conducteur	13
Figure 4 : Classification des algorithmes principaux utilisés en reconnaissance faciale	14
Chapitre 2	
Figure 5: Architecture d'un système de reconnaissance des expressions faciales	15
Figure 6: Opérateur LBP de base	19
Figure 7: les filtres qui utilisent dans le calcul de gradient verticale et horizontale	20
Figure 8: Classification des algorithmes principaux utilisés en reconnaissance Faciale	21
Figure 9 : Topologie des Réseaux de neurones artificiels	23
Figure 10 : La relation entre l'intelligence artificielle, le ML et le deep Learning	24
Figure 11 : Schéma illustratif de DL avec plusieurs couches	24
Figure 12 : Comparaison entre la machine Learning et le Deep Learning	25
Figure 13 : Différents modèles du Deep Learning	26
Figure 14 : Architecture générale d'un CNN	27
Figure 15 : Architecture proposée	28
Figure 16 : Phase de convolution	29
Figure 17: différents types de convolutions	29
Figure18: Exemple de description de la couche pooling	31
Figure 19 : Calcul du pooling sur une image 4×4	31
Figure 20 : Mise à plat des images	32
Chapitre III	
Figure 21: Environnement de KAGGEL	33
Figure 22: Le logo de python	34
Figure 23: Annotations d'images de l'ensemble de données proposé	37
Figure 24 : Résultat du modèle obtenu faciale	39
Figure 25 : Résultat de l'entraînement du modèle avec les images	40

Figure 26: Définissons les classes de la BDD de test et d'entraînement.	40
Figure 27 : Les états obtenus de la BDD de test et d'entraînement	41
Figure 28 : L'affichage des images	41
Figure 29 : Les résultats de teste: Œil fermer	42
Figure 30 : Les résultats de teste: œil ouverte	43
Figure 31 : Résultat de test	44

Liste des tableaux :

Tableau	Page
Tableau 1 : Différentes classifications d'expressions faciales	5
Tableau 2: Synthèse des travaux développés	22

Introduction

Générale

Introduction générale

Les expressions faciales sont un signe visible d'un état émotionnel où psychologie humaine. La reconnaissance des expressions faciales est une application très pertinente dans de nombreux domaines, notamment : domaine médicale, domaine sécuritaire, ... Les expressions faciales varient d'une personne à l'autre. C'est donc une tâche pas facile a effectuée. La détection automatique des expressions faciales suit un schéma commun. Le processus de reconnaissance des expressions faciales est constitué des étapes suivantes : Extraction des données, tests d'apprentissage, classifications et décisions.

Plusieurs travaux ont été développés dans le domaine de l'apprentissage automatique des expressions faciales comme les algorithmes d'extraction fonction (statistique ou structurelle) et le classificateur par apprentissage profond, en particulier les réseaux de neurones convolutifs (CNN). Il est apparu principalement pour résoudre les problèmes d'apprentissage automatique, un des principaux facteurs de succès de ces méthodes est la quantité de données d'entraînement.

Les systèmes de détection de la somnolence de conducteur disponibles actuellement appartiennent généralement à deux catégories : les systèmes très coûteux dédiées pour les voitures de haut de gamme et les systèmes abordables et peu robustes. La somnolence du conducteur est l'une des principales causes d'accidents de la route dans lesquels de nombreuses personnes décèdent où se blessent chaque année.

L'objectif de notre travail est de concevoir un système de détection de la somnolence du conducteur afin de l'aider à éviter les accidents causés par la collision. Cette détection se fait en fonction d'états des yeux.

Ce mémoire est organisé en trois chapitres comme suit :

- Dans le premier chapitre, nous présenterons un aperçu général sur la reconnaissance des expressions faciales.
- Le deuxième chapitre, mettra en évidence, les différentes approches de la reconnaissance des expressions faciales
- Dans le troisième chapitre nous présenterons l'implémentation et les résultats expérimentaux de notre système.

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 capables reconnaître les expressions faciales dans des vidéos ou des images. La 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.

2. Historique sur la reconnaissance de l'expression faciale :

L'expression faciale est un aspect important du comportement et de la communication non verbale, où le changement dans le visage, est perceptible visuellement, dû à l'activation (volontaire ou non) de l'un ou de plusieurs des 44 muscles composant le visage (250000 expressions possibles). L'expression faciale, déjà étudiée par Darwin et Duchenne de Boulogne au dix-huitième siècle, a joué un rôle majeur dans la recherche sur les émotions depuis les travaux de Sylvan Tomkins dans les années 1960. Ses élèves Paul Ekman et Carroll Izard ont défendu l'idée d'un nombre limité d'émotions de base auxquelles sont associées des expressions faciales automatiques, universelles et innées [1][2].

Durant la seconde moitié du dix-huitième siècle, le neurologue Duchenne de Boulogne réalise une série d'expériences sur l'expression faciale de l'émotion. Il utilise la photographie et la stimulation électrique des muscles de la face pour mettre en évidence les mouvements associés à l'expression des émotions. Il remarque notamment que les sourires exprimant une joie sincère se différencient des sourires volontaires par la contraction d'orbiculaire oculi, un muscle situé autour des yeux.

Des recherches menées dans les années 1980 par Paul Ekman et son équipe ont permis de confirmer et de compléter ces résultats. Ekman a mis en évidence le fait que nous sommes pour la plupart incapables de contracter volontairement l'orbiculaire oculi, et que ceux qui le peuvent n'arrivent généralement pas à contracter ce muscle de chaque côté au même moment.

En outre, les sourires de Duchenne ne sont généralement associés à une activité symétrique dans le lobe frontal, considérée comme un signe d'affection positif [3].

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 pourraient nuire, lutter contre la fraude des passeports, identifier les personnes disparues, etc...

3. Définition de visage

Le visage peut être défini comme : « Une structure tridimensionnelle possédant une configuration "externe" – le contour de la face modelé par les saillies osseuses et souligné par les cheveux – où s'inscrit une configuration "interne" formée par l'assemblage des traits. Certaines particularités s'y ajoutent comme la pilosité, les lunettes, la texture, la coloration de la peau... » [4]

- **Reconnaissance Faciale**

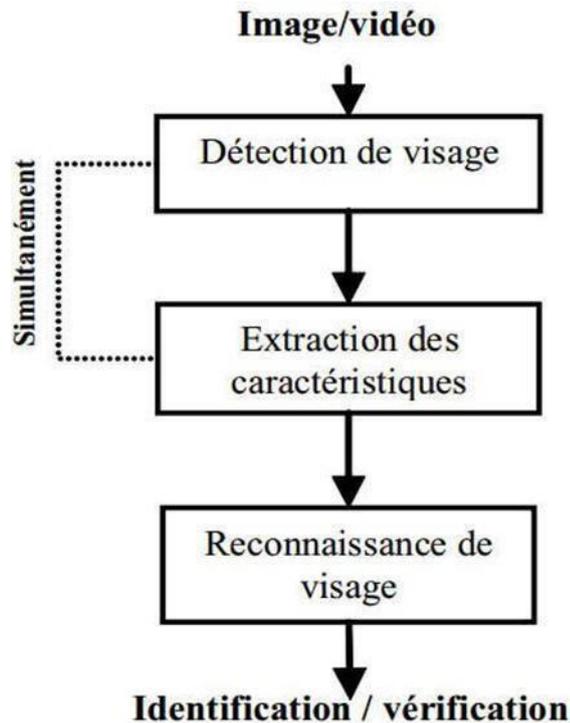
C'est une tâche que les humains effectuent naturellement. La grande disponibilité d'ordinateurs puissants et peu onéreux ainsi que des systèmes informatiques embarqués ont suscité un énorme intérêt dans le traitement automatique des images et des vidéos numériques au sein de nombreuses applications (la surveillance, IHM ...), en tant qu'une des technologies biométriques de base, ceci étant dû aux avancées rapides dans des technologies telles que les appareils photo numériques, Internet et les dispositifs mobiles, le tout associé à des besoins en sécurité sans cesse en augmentation. [5]

La détection de visage dans l'image est un traitement indispensable et crucial avant la phase de reconnaissance. En effet, le processus de reconnaissance de visages ne pourra jamais devenir intégralement automatique s'il n'a pas été précédé par une étape de détection efficace.

On peut classifier les systèmes de reconnaissance du visage en deux grandes catégories selon la source de capture de l'image : reconnaissance du visage dans une séquence vidéo ou bien à partir d'images fixes.

3.1. Etapes de la reconnaissance Faciale

La reconnaissance automatique de visage s'effectue en trois étapes principales : détection de visages, extraction et normalisation des caractéristiques du visage, identification et/ou vérification



a. Détection de visage

L'efficacité des systèmes biométriques basés sur l'authentification de visage dépend essentiellement de la méthode utilisée pour localiser le visage dans l'image ; tels que les yeux, le nez, la bouche, etc. Cependant, les solutions doivent gérer la variabilité des conditions d'acquisition de la vie quotidienne, notamment : La pose ; La présence ou absence des composantes structurales (la barbe) ; Les occultations (visages peuvent masquer d'autres visages) ; L'éclairage, les caractéristiques de l'appareil photographique...etc.

b. Extraction des caractéristiques du visage

L'extraction des caractéristiques telles que les yeux, le nez... est une étape de prétraitement nécessaire à la reconnaissance faciale. On peut distinguer deux pratiques différentes : la première repose sur l'extraction des régions entières du visage, elle est souvent implémentée avec une approche globale de reconnaissance de visage. La deuxième pratique extrait des points particuliers des différentes régions caractéristiques du visage, tels que la bouche et du nez. Elle est utilisée avec une méthode locale de reconnaissance et aussi pour l'estimation de la pose du visage.

c. La reconnaissance de visage

Le module de reconnaissance exploite les caractéristiques du visage ainsi extraites pour créer une signature numérique qu'il stocke dans une base de données. Ainsi, à chaque visage de la base est associée une signature unique qui caractérise la personne correspondante. La reconnaissance d'un visage requête est obtenue par l'extraction de la signature requête

Chapitre I : Les généralités des expressions faciales

correspondante et sa mise en correspondance avec la signature la plus proche dans la base de données. La reconnaissance dépend du mode de comparaison utilisé : vérification ou identification [6].

4. Difficultés de la reconnaissance des visages

- ❖ Changement d'illumination les variations d'éclairage rendent la tâche de reconnaissance de visage très difficile.
- ❖ Variation de pose est considérée comme un problème majeur pour les systèmes de reconnaissance faciale. Cependant, lorsque la rotation est supérieure à 30°, la normalisation géométrique n'est plus possible.
- ❖ Expressions faciales la déformation du visage qui est due aux expressions faciales est localisée principalement sur la partie inférieure du visage. Elle entraîne forcément une diminution du taux de reconnaissance.
- ❖ Les vrais jumeaux il est peu probable que la vérification automatique de visage, ne pourra jamais détecter les différences très subtiles qui existent entre les jumeaux. (Problème Apple face détection lock).

5. Différents types des expressions faciales

Le tableau représenté ci-dessous (tableau 1) une classification des expressions faciales selon certains chercheurs à base d'émotion et d'inclusion [7]

Théoriciens	Émotions de base	Base d'inclusion
Plutchik	Acceptation, colère, anticipation, dégoût, joie, peur, tristesse, surprise	Relation aux processus biologiques adaptatifs
Arnold	Colère, aversion, courage, abattement, désir, désespoir, peur, haine, espoir, amour, tristesse	Relation aux tendances d'action
Frijda	Désir, bonheur, intérêt, surprise, émerveillement, chagrin	Désir, bonheur, intérêt, surprise, émerveillement, Chagrin
McDougall	Colère, dégoût, exaltation, peur, soumission, émotion tendre, émerveillement	Relation à l'instinct
Ekman, Friesen Et Ellsworth	Colère, dégoût, peur, joie, tristesse, surprise	Expressions faciales universelles

Tableau 1: Différentes classifications d'expressions faciales .

6. Système de reconnaissance des expressions faciales

Les expressions faciales sont l'un des moyens de communication non verbaux résultant d'une réaction souvent involontaire, résultant principalement de contractions musculaires modifiant les traits du visage tels que la bouche, les sourcils. C'est un moyen essentiel de transmission d'informations sociales entre humains, d'une durée de quelques secondes. Un système de reconnaissance des expressions faciales est constitué de trois étapes principales.

La première étape c'est la détection du visage où déterminer la zone du visage. La deuxième étape c'est l'extraction de donnée à partir du visage d'individu pour extraire les caractéristiques du visage tels que la position de la bouche, les yeux, et les sourcils et donner une meilleure description de l'expression. La dernière étape est la classification qui permet de classer spécifiquement l'expression appropriée pour l'image à partir des caractéristiques extraites dans l'étape précédente.

Le schéma suivant illustre un système complet de reconnaissance des expressions faciales:

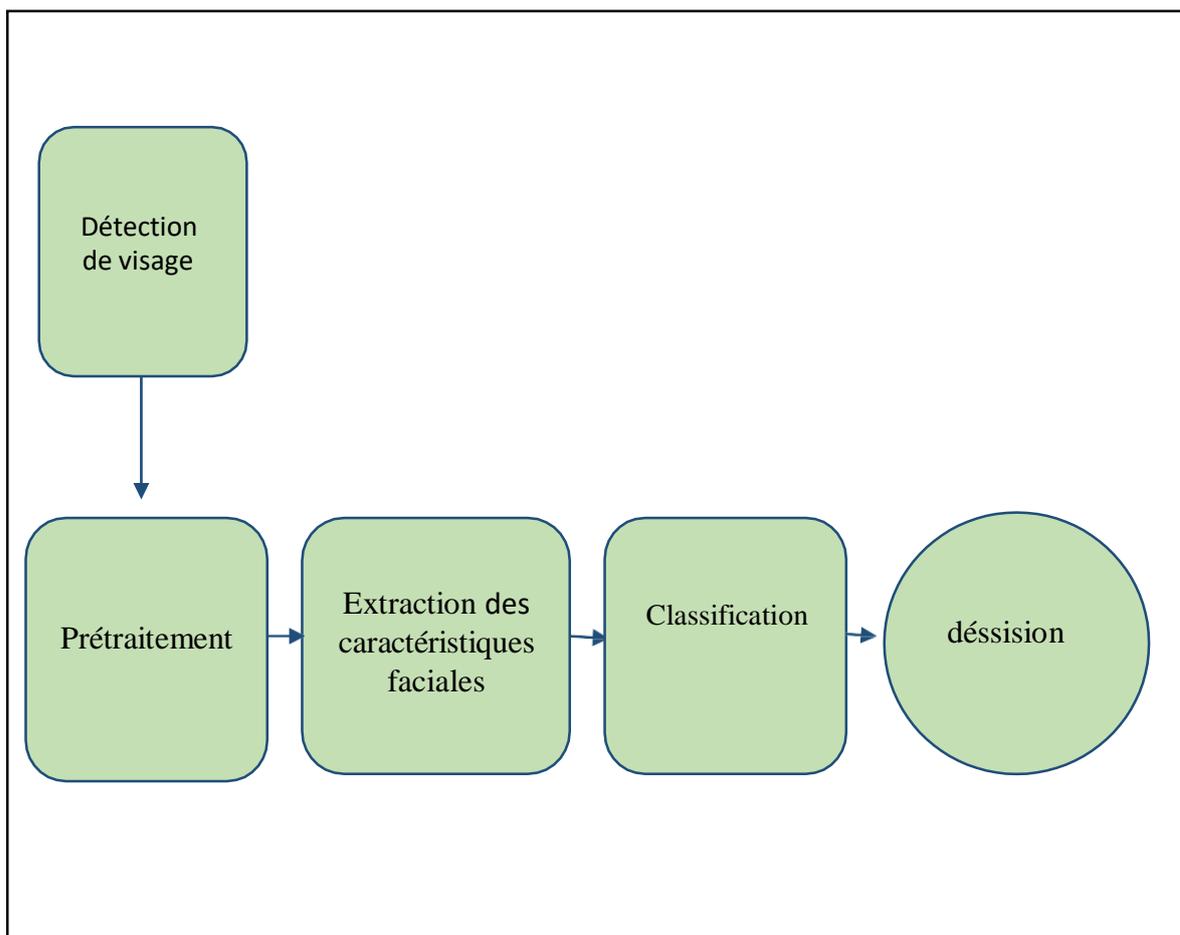
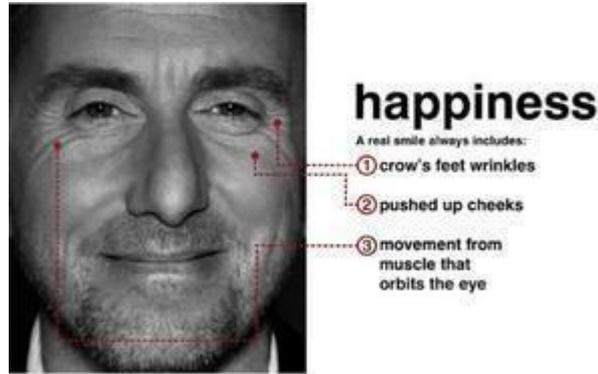


Figure1 : Architecture d'un système de reconnaissance des expressions faciales.

7. Les 7 expressions faciales universelles

La joie (joy)- facile à reconnaître, difficile à feindre



C'est peut-être trivial, mais lorsqu'une personne ressent de la joie, elle sourit. L'inverse, c'est à dire essayer de faire croire à quelqu'un qu'on est joyeux, est beaucoup plus difficile comme je l'expliquais dans l'article : le sourire, l'arme ultime pour être apprécié.

Cette situation m'est récemment apparue lorsque je proposais à l'une de mes amies une séance piscine avec le beau temps de ces derniers jours. Sur le principe, elle était d'accord et souriait mais dans l'absolument je sentais que son sourire était faux. Après s'être accordé sur le RDV, le lapin anticipé n'a pas tardé à apparaître quelques heures avant d'enfiler le maillot (normalement l'excuse justement, c'est : "je peux pas, j'ai piscine" . Bref passons, comme le laissait suggérer son visage, elle n'était pas du tout enthousiaste à aller faire trempette.

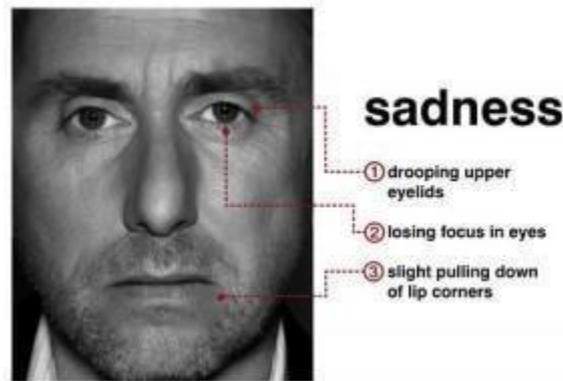
Un sourire délibérément faussé apparaît souvent dans ce type de situation, quand vous voulez faire plaisir à quelqu'un et ne pas le vexer. Mais également lorsque vous êtes en désaccord avec votre interlocuteur et que vous faites semblant de l'être.

Parfois, situation inverse, on peut observer une micro expression de joie, alors qu'elle ne devrait "normalement" pas apparaître. Cette expression naît quand une personne est fière d'elle lors d'une réprimande par exemple.

Je ne connais pas le milieu carcéral. Pourtant, dans l'un des épisodes de la série Lie to Me, l'acteur principal interroge un tueur en série présumé en mentionnant ses précédents meurtres. Le criminel nie les accusations mais se fait au final démasqué par le sentiment de joie (quelques millièmes de seconde) qui apparaît sur son visage, traduisant la fierté de cette entreprise pour la moins sinistre.

Les indices : le sourire et les plis au niveau des yeux prouvant ainsi que le sourire est authentique.

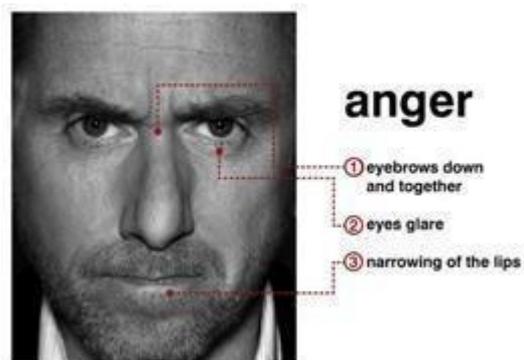
La tristesse (sad) :



Il est parfois important de savoir repérer la tristesse derrière un sourire. Que ce soit pour avoir une oreille attentive ou bien pour repérer la détresse dans laquelle se trouve un proche. Beaucoup de personnes peuvent faire semblant d'être heureuses car elles ne veulent pas entacher la bonne humeur de leur entourage. Pourtant, il est parfois nécessaire d'observer cette tristesse et d'apporter l'aide adéquate. Au sein même d'une conversation, changer de la tristesse en de la joie peut relancer votre discussion avec un nouvel élan.

Les indices : les paupières supérieures et les coins externes des sourcils s'affaissent. La personne donnera l'impression d'avoir le regard dans le vide. Également, les coins des lèvres ont tendance à s'affaisser.

La colère (angry) :

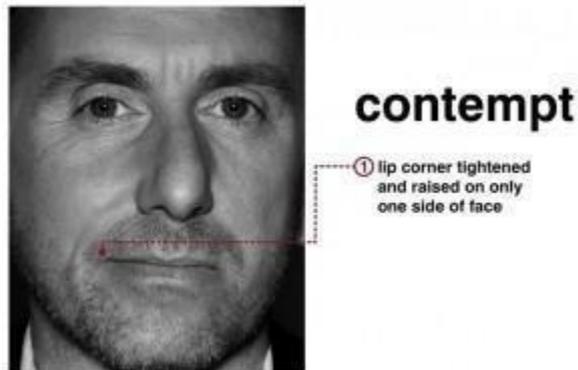


La colère est en général facile réopérer. Mais parfois elle peut survenir d'un coup en pleine conversation banale, s'en crier "GARE", et avec toutes les conséquences qui suivent et qui auraient pu être repérées aux premiers abords.

Chapitre I : Les généralités des expressions faciales

Les Indices : les lèvres sont pressées l'une contre l'autre. Les sourcils se pressant vers le nez. Les rides entre deux sourcils sont apparentes et la plus grande caractéristique restant le regard fixe.

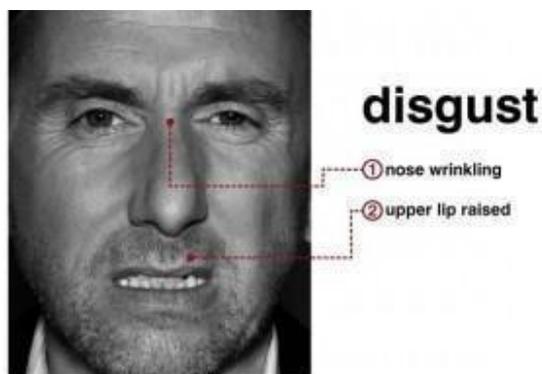
Le mépris (contempt) :



C'est l'expression pour laquelle mes poils se hérissent tellement elle m'exaspère. Et pourtant sur les tests de détection de micro-expressions en début d'article, elle est la plus facile à repérer. Et pour cause, il s'agit de la seule expression parmi ces 7 émotions qui est asymétrique. Ceci se voit très bien sur l'illustration. Cette expression est unilatérale sur l'une ou l'autre des faces (gauche ou droite). Toutes les autres expressions, que ce soit la colère ou la surprise, ont des éléments caractéristiques symétriques sur le visage suivant l'axe vertical du nez.

Les indices : un côté de la lèvre est légèrement tiré vers l'arrière (ou remonté). Ce mouvement de la lèvre peut être très subtil, apparaître et disparaître très rapidement.

Le dégoût (disgust) :



Chapitre I : Les généralités des expressions faciales

Si vous pouvez observer cette expression sur le visage de votre interlocuteur, il s'agit d'un très bon indice pour savoir si ce que vous racontez plaît ou non.

Dernièrement, j'ai pu observer cette expression chez une trentaine de personnes autour de moi pendant quelques minutes. C'est en me retournant dans la salle de cinéma lors de la projection du film 127. Ceux qui l'ont vu comprendront de quoi je parle. Pour les autres, je le conseille (ame sensible tout de même s'abstenir) pour aux amoureux de la nature et du dépassement de soi.

Les indices : la lèvre supérieure remonte laissant généralement apparaître les dents. Cet indice s'accompagne généralement d'un pli au niveau du nez.

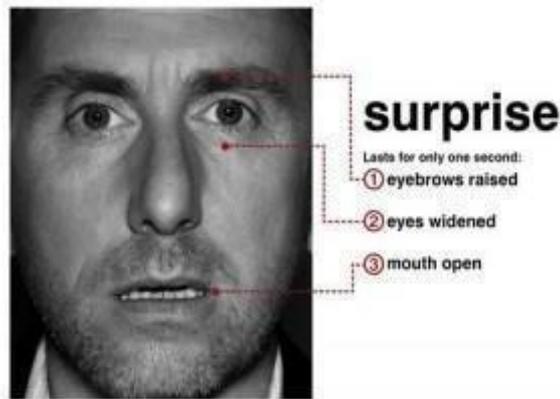
La Peur (afraid) :

C'est une expression que l'on connaît tous. Si vous voulez l'observer je vous invite tout simplement à surprendre l'un de vos amis. Sur son visage se dessineront pendant quelques millisecondes les expressions suivantes.



Les indices : Les lèvres s'étirent à l'horizontal vers les oreilles. Les paupières inférieures se tendent et les paupières supérieures sont remontées. Les sourcils remontent également comme s'ils s'étaient donné rendez-vous pour se rejoindre.

La surprise (surprise)



Dans le test, je mélange souvent les deux micro expressions : la surprise et la peur. Les éléments faciaux impliqués sont similaires et les contextes parfois aussi. C'est une micro-expression très importante pour déterminer le mensonge.

Si vous accusez votre conjoint de quelque chose et qu'ensuite a l'air surpris de votre remarque, il se peut que votre accusation soit tout de même correcte. Pour le savoir, il suffit de prêter attention à la durée de cette expression. En effet, il est facile de simuler la surprise.

Mais ce que les gens ne savent pas, c'est que la surprise simulée s'identifie facilement. En effet, si quelqu'un fait semblant d'être surpris, son expression faciale durera plus d'une seconde. A l'inverse, la vraie surprise est très rapide (moins d'une seconde) avant de revenir à un état neutre. Elle est très brève. Si quelqu'un garde ses sourcils vers le haut plus d'une seconde, il ment, sans aucun doute.

Les indices : la bouche est entrouverte, les yeux bien ouverts et les sourcils relevés. Le tout, et c'est le plus important, dure moins d'une seconde. [8]

8. La somnolence

Les différentes études qui se sont intéressées aux effets dynamiques ou statiques de la fatigue sur le visage sont résumées dans le tableau ci-dessous. Aucune d'entre elles n'a toutefois répondu à la question fondamentale à savoir s'il est possible ou non de détecter une personne fatiguée, à l'œil nu, sur la base des traits de son visage.

Relation entre la somnolence et la conduite

La somnolence est définie comme une diminution du niveau de conscience dépeint par la somnolence et la difficulté à rester alerte, mais la personne se réveille avec une simple excitation provoquée par les stimuli. Cela peut être dû à une absence de repos, à des

médicaments, à un abus de substances ou à un problème cérébral. La somnolence est principalement le résultat de la fatigue qui peut être à la fois mentale et physique. D'autre part, n'importe quelle activité poursuivie suffisamment longtemps, causera à une difficulté à maintenir des performances qualifiées, y compris la conduite automobile. Et pour arrêter ou réduire le nombre des accidents, l'état de somnolence du conducteur devrait être surveillé en permanence.

Approches de détection de somnolence

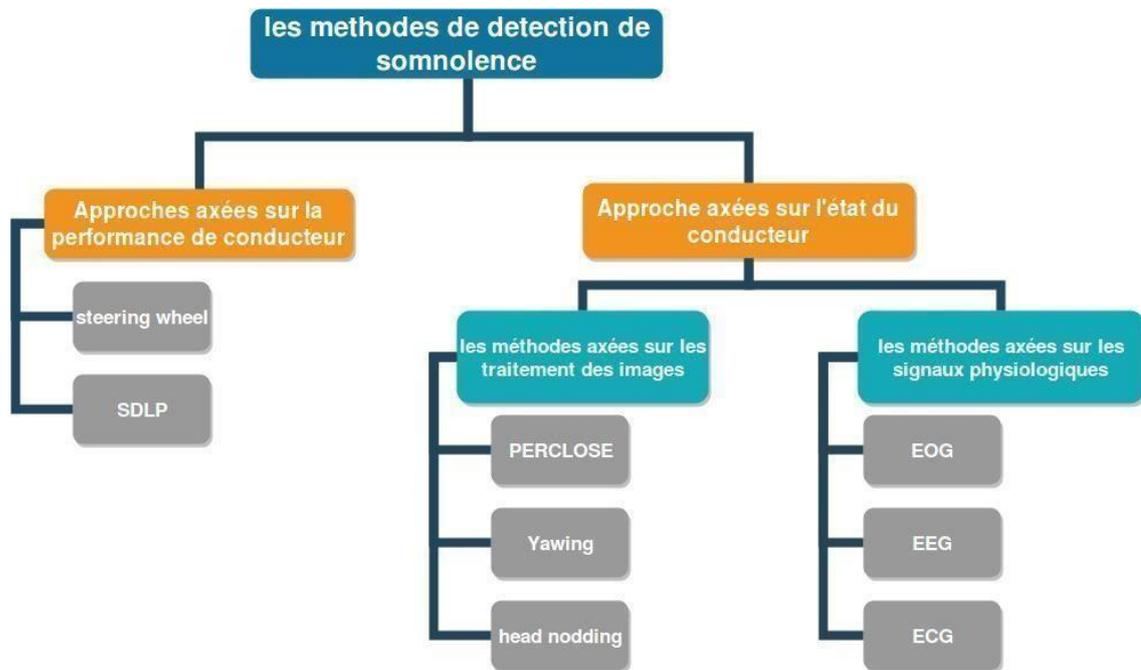


Figure 2: Différentes méthodes de détection de somnolence.

Approches axées sur les signaux physiologiques

Quand une personne se sent somnolente ou fatiguée, de nombreux changements s'apparaissent sur son corps. Ces changements incluent les changements de la vitesse du rythme cardiaque, l'état du cerveau et l'état des muscles. Basant sur ces changements, plusieurs recherches et approches sont proposées. Dans ces approches généralement utilisent des capteurs portés par le conducteur comme on a indiqué dans la figure 2.3, Ces capteurs permettent à capter plusieurs types des signaux EEG (l'état du cerveau), EOG (mouvement des yeux) et ECG (rythme cardiaque). Dans Quelques recherches, ils ont utilisé les signaux EOG pour surveiller l'état du conducteur pendant la conduisent. Et Ceci est fait en classifier les mouvements des yeux à l'aide des signaux vertical et horizontal de l'EOG obtenu à partir des capteurs. La classification se fait selon l'un des algorithmes d'apprentissage SVM [9], PNN (Prototype Near est Neighbor) et MLP (Multi-Layer Perceptron) [10].

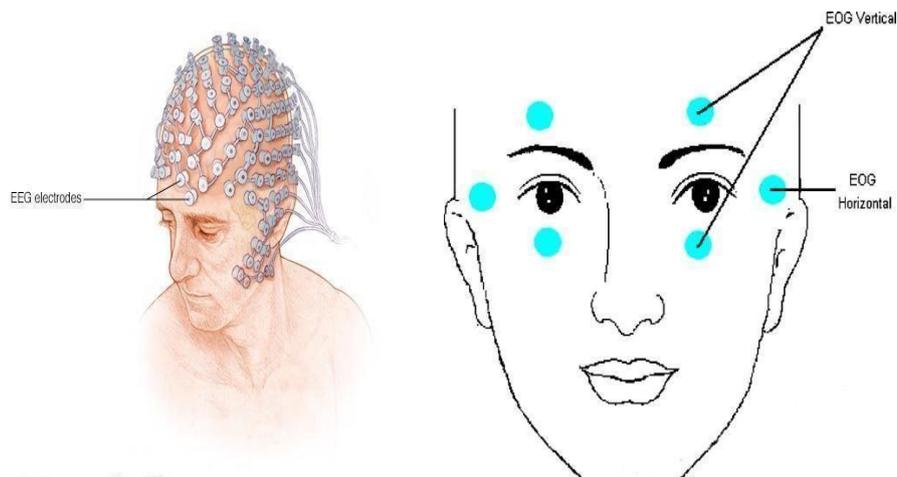


Figure 3: illustration des capteurs utilisés dans les méthodes axées sur la performance du Conducteur [10].

Les signaux ECG sont aussi utilisés pour détecter l'état du conducteur. Ceci se fait à l'aide de variabilité du rythme cardiaque, dans lequel l'augmentation du rythme cardiaque est une indication de diminution du niveau de conscience, ce qui peut se produire au cours de la somnolence du conducteur. L'EEG est également l'un des meilleurs signaux qui aident à étudier l'état du conducteur. L'EEG a différentes bandes de fréquences où la bande delta entre 0.5 et 4 Hz correspond à l'activité du sommeil, la bande thêta entre 4 et 8 Hz correspond à la somnolence, la bande alpha entre 8 et 13 Hz correspond à relaxation et la bande beta entre 13 et 25 Hz correspond à la vigilance [10].

Approches axées sur le traitement d'image

Les approches précédentes sont classifiées comme des approches intrusives car elles sont influencées par l'attention du conducteur. Maintenant, en raison du développement technologique, les recherches sont orientées vers les méthodes basées sur les traitements des images où on utilise une caméra pour extraire les symptômes de somnolence du conducteur comme les mouvements des yeux, les bâillements fréquents, les mouvements de tête.

9. Techniques de reconnaissance faciale

Plusieurs méthodes de reconnaissance de visages ont été proposées ces dernières années, suivant deux grands axes : la reconnaissance à partir d'images fixes et la reconnaissance à partir de séquence d'images (vidéo). De manière générale, on distingue trois catégories de méthodes : les méthodes globales, les méthodes locales et les méthodes hybrides [11].

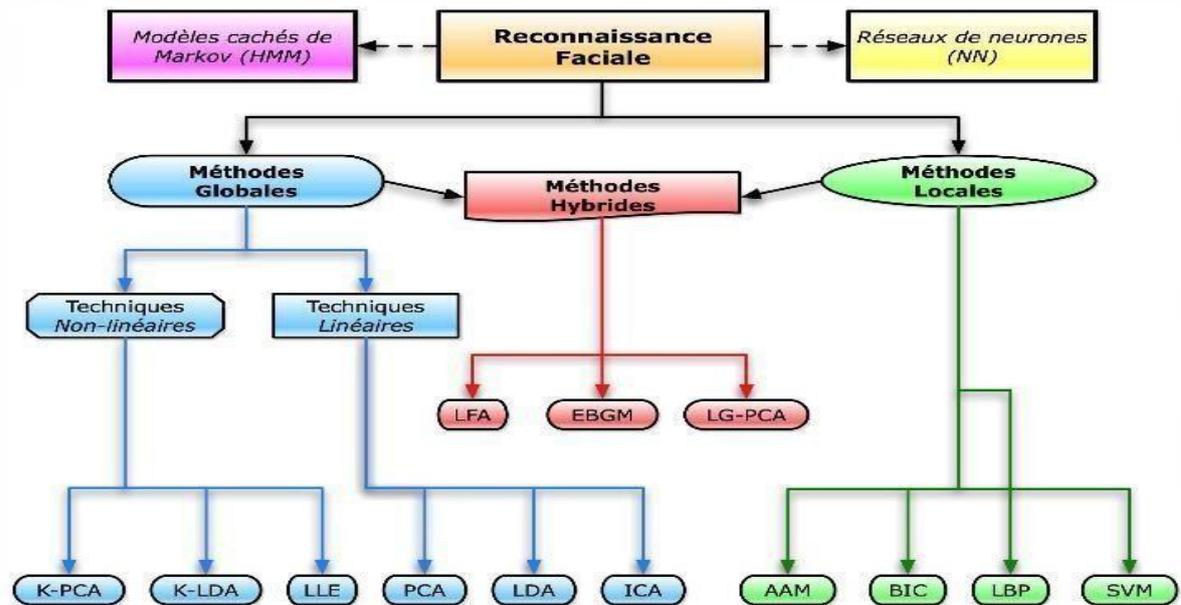


Figure 4 : Classification des algorithmes principaux utilisés en reconnaissance faciale

10. Avantages et inconvénients de la reconnaissance faciale

Plusieurs facteurs rendent la modalité visage attractive pour une utilisation à grande échelle, elle est acceptable par les personnes vu que c'est une partie apparente du corps, facilement vérifiable par n'importe quel opérateur pour avoir une décision, intrusive en raison de la vérification aisée. De plus, les capteurs d'images sont moins chers sur le marché [12]. En dépit de ses avantages, elle présente des inconvénients qui influent sur la qualité de la reconnaissance. On peut citer les aspects suivants :

- Changement d'illumination.
- Expressions faciales.
- Présence ou absence des composantes structurales, telles que : la barbe, la moustache et les lunettes.
- Les vrais jumeaux qui ont le même indicatif d'ADN

11. Conclusion

Dans ce chapitre, nous avons apporté les principaux aspects de la reconnaissance faciale et leur importance dans l'évolution et l'architecture générale d'un système de reconnaissance de visages et nous avons fait un survol de quelques méthodes de reconnaissance des expressions faciales, les méthodes globales, locales et hybrides.

Dans la partie suivante, nous allons présenter la reconnaissance des expressions faciales et les différentes méthodes qui permettent leurs extractions.

1. Introduction

Le système visuel est l'ensemble des organes participant à la réception visuelle, de la rétine aux différentes parties du cortex visuel. Son rôle est de percevoir et d'interpréter les deux images en deux dimensions acquises par nos yeux, mais également l'image en trois dimensions de la scène reconstruite en intégrant des connaissances a priori

Les déformations des traits du visage représentent l'information sur l'expression du visage, qu'a reclassée précédemment dans six émotions universelles dégoût, joie, peur, tristesse, surprise, avec le développement de la vision artificielle et de la communication homme-machine, nous avons découvert qu'il existe plusieurs techniques qui tentent d'identifier une personne avec une précision et des performances élevées.

Le problème de la reconnaissance de visages peut être formulé comme suit : étant données une ou plusieurs images d'un visage, la tâche est de trouver ou de vérifier l'identité d'une personne par comparaison de son visage à l'ensemble des images de visage stockées dans une base de données

2. Système de reconnaissance des expressions faciales

Le système de reconnaissance exploite les caractéristiques du visage ainsi extraites pour créer une signature numérique qu'il stocke dans une base de données. Ainsi, à chaque visage de la base est associée une signature unique qui caractérise la personne correspondante. La reconnaissance d'un visage requête est obtenue par l'extraction de la signature requête correspondante et sa mise en correspondance avec la signature la plus proche dans la base de données.

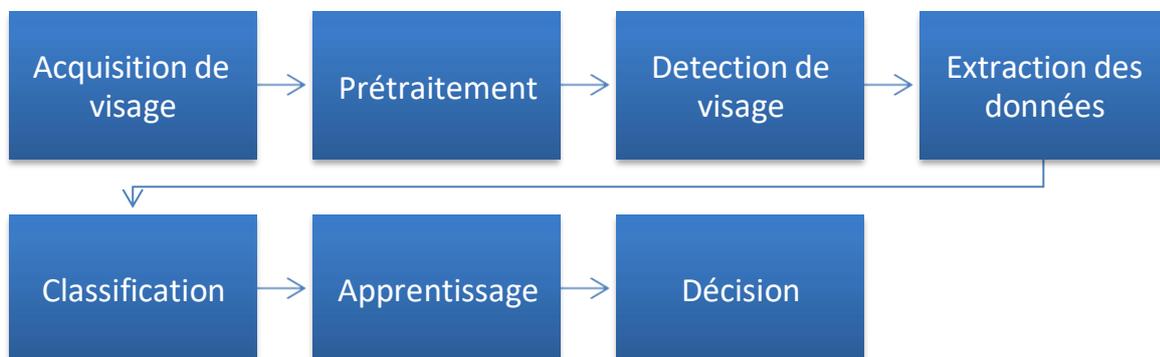


Figure 5: Architecture d'un système de reconnaissance des expressions faciales.

2.1 Acquisition de visage

Cette étape consiste à extraire l'image de l'utilisateur du monde extérieur dans un état statique à l'aide d'un appareil photo ou dynamique à l'aide d'une caméra. Après, l'image extraite sera digitalisée ce qui donne lieu à une représentation bidimensionnelle au visage, caractérisée par une matrice de niveaux de gris. L'image dans cette étape est dans un état brut ce qui engendre un risque de bruit qui peut dégrader les performances du système [13]

2.2 Prétraitements

Le rôle de cette étape est d'éliminer les parasites causés par la qualité des dispositifs optiques ou électroniques lors de l'acquisition de l'image en entrée, dans le but de ne conserver que les informations essentielles et donc préparer l'image à l'étape suivante. Elle est indispensable car on ne peut jamais avoir une image sans bruit à cause du background et de la lumière qui est généralement inconnue. Il existe plusieurs types de traitement et d'amélioration de la qualité de l'image, telle que : la normalisation, l'égalisation et le filtre médian. Cette étape peut également contenir la détection et la localisation du visage dans une image, surtout là où le décor est très complexe [14].

2.3 Détection du visage

Après avoir capturé la scène contenant un visage, la deuxième étape consiste à extraire l'image. Cela peut se faire par détection de la couleur de la peau, ou par des méthodes détectant les différentes caractéristiques du visage par des descripteurs locaux (ada_boost). Cette étape est autant plus délicate autant que l'image acquise contient plusieurs objets de visage ou un fond non uniforme qui crée une texture perturbant la bonne segmentation du visage. Cette étape est dépendante de la qualité des images acquises [15].

Après la segmentation du visage, on peut filtrer ou améliorer la qualité par des prétraitements qui sont appliqués au visage extrait. On peut effectuer des normalisations géométriques et photométriques. Ces prétraitements sont nécessaires pour éliminer ou limiter les variations de pose ou d'illumination. Un prétraitement photométrique tend à uniformiser l'éclairage dans une image et ainsi minimiser l'influence de l'illumination. Cela peut être effectué soit par des méthodes simples telle que l'égalisation d'histogramme, une correction gamma ou par des méthodes plus complexes tel la méthode retinex. Une normalisation géométrique est un ajustement du visage pour qu'il ait une dimension donnée et qu'il soit horizontal [16].

2.4 Extraction des données caractéristique

Une fois que vous avez identifié la position du visage dans l'image, l'étape suivante est l'extraction des caractéristiques. Il s'agit de l'étape la plus importante d'un système de reconnaissance des expressions faciales. L'extraction d'informations à partir du visage capturé peut être obtenue en décrivant des points d'intérêt liés aux caractéristiques faciales telles que les yeux, les sourcils, la bouche, etc. de sorte qu'un vecteur représente toutes les données.

2.4.1 Différentes méthodes développées pour la reconnaissance faciale

Ils existent différentes approches qui ont été développées dans le domaine de reconnaissance des expressions faciales. Elles peuvent être séparées en deux grandes familles à savoir : les méthodes globales (ou holistiques) et les méthodes locales. [17]

a. Méthodes globales

Elles sont basées sur des techniques d'analyse statistique bien connues. Il n'est pas nécessaire de repérer certains points caractéristiques du visage (comme les centres des yeux, les narines, le centre de la bouche, etc.) à part pour normaliser les images. Dans ces méthodes, les images de visage (qui peuvent être vues comme des matrices de valeurs de pixels) sont traitées de manière globale et sont généralement transformées en vecteurs, plus faciles à manipuler. L'avantage principal des méthodes globales est qu'elles sont relativement rapides à mettre en œuvre et que les calculs de base sont d'une complexité moyenne. En revanche, elles sont très sensibles aux variations d'éclairage, de pose et d'expression faciale.

Ceci se comprend aisément puisque la moindre variation des conditions de l'environnement ambiant entraîne des changements inéluctables dans les valeurs des pixels qui sont traités directement. Ces méthodes utilisent principalement une analyse de sous-espaces dévisages. L'utilisation de techniques de modélisation de sous-espace fait avancer la technologie de reconnaissance faciale de manière significative.

Nous pouvons distinguer deux types de techniques parmi les méthodes globales : les techniques linéaires et les techniques non linéaires.

Les techniques linéaires *projettent* linéairement les données d'une espace de grande dimension (par exemple, l'espace de l'image originale) sur un sous-espace de dimension inférieure. Malheureusement, ces techniques sont incapables de préserver les variations non convexes des variétés (géométriques donc au sens mathématique du terme) de visages afin

Dédifférencier des individus. Dans un sous-espace linéaire, les distances euclidiennes et plus généralement les distances de Mahalanobis, qui sont normalement utilisées pour faire comparer des vecteurs de données, ne permettent pas une bonne classification entre les classes de formes et entre les individus eux-mêmes. La technique linéaire la plus connue et sans aucun doute l'analyse en composantes principales (PCA), également appelée transformée de Karhunen-Loeve. Le PCA fut d'abord utilisé afin de représenter efficacement des images de visages humains. [18] [19] [20]

Bien que ces méthodes globales linéaires basées sur l'apparence évitent l'instabilité des toutes premières méthodes géométriques qui ont été mises au point, elles ne sont pas assez précises pour décrire les subtilités des variétés (géométriques) présentes dans l'espace de l'image originale. Ceci est dû à leurs limitations à gérer la non-linéarité en reconnaissance faciale : les déformations de variétés non linéaires peuvent être lissées et les concavités peuvent être remplies, causant des conséquences défavorables.

Afin de pouvoir traiter ce problème de non-linéarité sur la reconnaissance des expressions faciales, de telles méthodes linéaires ont été étendues à des techniques non linéaires basées sur la notion mathématique de noyau ("kernel") comme le Kernel PCA et le Kernel LDA [21].

b. Méthodes locales

Elles sont basées sur des modèles, utilisant des connaissances a priori que l'on possède sur la morphologie du visage et s'appuie en général sur des points caractéristiques de celui-ci. Kanade a présenté un des premiers algorithmes de ce type [22] en détectant certains points ou traits caractéristiques d'un visage puis en les comparant avec des paramètres extraits d'autres visages. Ces méthodes constituent une autre approche pour prendre en compte la non-linéarité en construisant une espace de caractéristiques locales et en utilisant des filtres d'images appropriés, de manière à ce que les distributions des visages soient moins affectées par divers changements.

Les approches bayésiennes, les machines à vecteurs de support (SVM) [23], la méthode des modèles actifs d'apparence (AAM) [24] ou encore la méthode "local binary pattern" (LBP) ont été utilisées dans ce but.

Toutes ces méthodes ont l'avantage de pouvoir modéliser plus facilement les variations de pose, d'éclairage et d'expression par rapport aux méthodes globales. Toutefois, elles sont plus lourdes à utiliser puisqu'il faut souvent placer manuellement un assez grand nombre de

Points sur le visage alors que les méthodes globales ne nécessitent de connaître que la position des yeux afin de normaliser les images, ce qui peut être fait automatiquement et de manière assez fiable par un algorithme de détection [25]

- **Algorithm LBP (Local Binary Pattern)**

Le descripteur de motif binaire local (LBP) est une méthode efficace d'extraction des caractéristiques faciales ; Il extrait des informations des valeurs de pixel voisines et développe l'histogramme de l'image. Il s'agit d'un opérateur non paramétrique et décrit la structure spatiale locale d'une image. Il calcule un code binaire à partir de dérivées binaires de pixels puis trouve la différence du pixel central avec ses pixels voisins, arrange ces différences d'une forme ordonnée et finalement ce motif de bits est converti en valeur décimale qui est le nouveau code LBP pour le Pixel central. L'opérateur LBP de base fonctionne pour 3×3 pixels est décrite par le diagramme suivant : [26] .

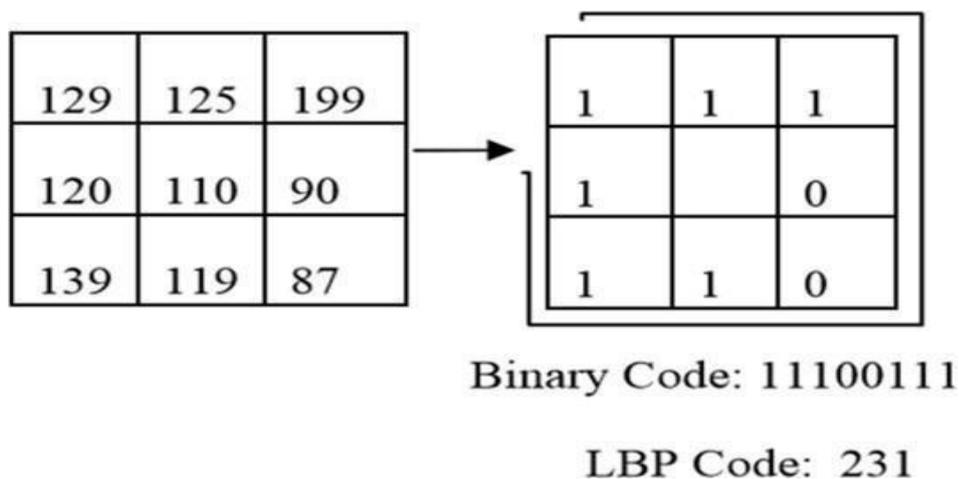


Figure 6 : Opérateur LBP de base.

La figure ci-dessus montre des valeurs d'échelle de gris de 3×3 pixels et le code LBP est calculé en utilisant la formule suivante :

$$LBP(x_c, y_c) = \sum_{n=0}^7 S(i_n - i_c) 2^n$$

$$S(x) = \begin{cases} 1 & \text{if } (x \geq 0) \\ 0 & \text{if } (x < 0) \end{cases}$$

Ici x_c et y_c montrent la position du pixel central i_c et i_c sont des valeurs d'échelle de gris des pixels environnants et du pixel central respectivement. Après l'étiquetage de l'image avec des codes LBP, image histogramme est générée qui permet de reconnaître micro-motifs dans l'image comme les yeux, le nez et les lèvres... etc.

✓ **Avantage :**

Tolérance vis-à-vis des changements d'illumination et simplicité.

✓ **Inconvénient :**

L'opérateur produit un Histogramme assez long et n'est pas trop robuste sur l'image plane.

[14]

- **Histogram of oriented gradients (HOG)**

Histogram of oriented gradients (HOG) est une caractéristique utilisée en vision par ordinateur pour la détection d'objet. Cette technique calcule des histogrammes locaux de l'orientation du gradient sur une grille dense, c'est-à-dire sur des zones régulières entières sur l'image.

Cette caractéristique a été utilisée pour la première fois par Navne et Dalal et Bill Triggs en 2005 pour détecter les piétons. Mais elle peut être utilisée pour détecter d'autres objets comme les voitures, Les avion, les visages...etc. Cette méthode est très simple et on peut la résumer en deux étapes principales : le calcul du vecteur des caractéristiques Histogrammes et la classification.

➤ **Calcul de vecteur des caractéristiques Histogrammes**

Comme nous avons dit. La méthode de HOG est destinée à la détection des piétons à l'origine. Pour cette raison Navne et Dalal et Bill Triggs [27] a utilisé une fenêtre de taille (64x128). Bien sûr, une fenêtre peut avoir n'importe quelle taille selon l'objet que l'on souhaite détecter. La première étape dans la réalisation de vecteur d'histogramme est le calcul du gradient vertical et horizontal. Ceci est facilement réalisé avec les filtres indiqués dans la figure

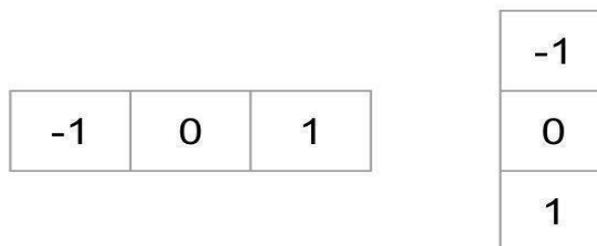


Figure 7 : les filtres qui utilisent dans le calcul de gradient verticale et horizontale

Quant à la deuxième étape, elle consiste à calculer la longueur et la direction du vecteur en fonction des formules et les résultats de la première étape.

C. Methods hybrids

Ce sont une combinaison des méthodes globales et locales en combinant la détection de caractéristiques géométriques (ou structurales) avec l'extraction de caractéristiques d'apparence locales. Elles permettent d'augmenter la stabilité de la performance de reconnaissance lors de changements de pose, d'éclairage et d'expressions faciales. [28].

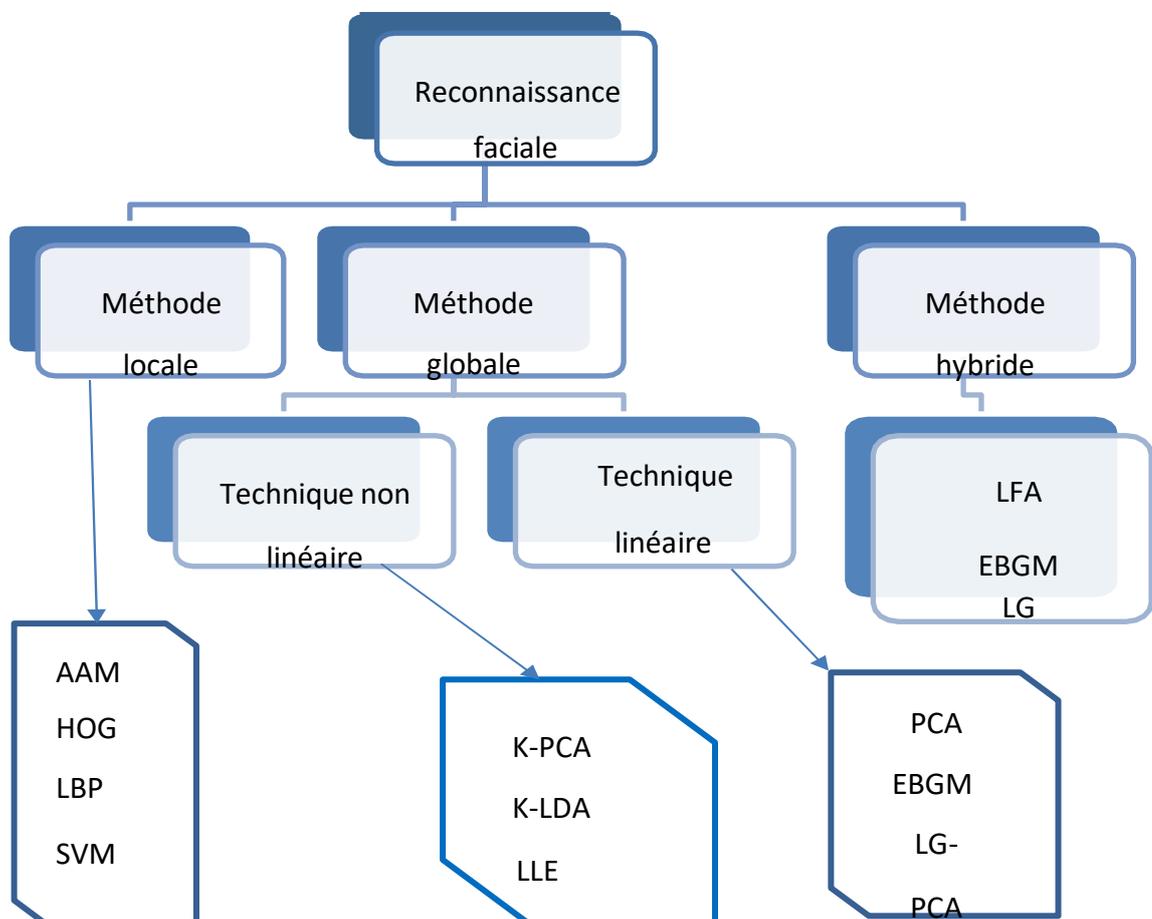


Figure 8 : Classification des algorithmes principaux utilisés en reconnaissance faciale .

3. Synthèse des travaux développés pour la reconnaissance d'expressions faciales

Le tableau suivant (tableau 2) présente quelques techniques récentes développées pour la reconnaissance des expressions faciales :

Références	Technique	Expression faciales	Bases de données	Type	Sujets	Performance
[29]	LPB, SVB	Six	CK+	images	50	97 %
[30]	EBGM, ELM	Six	CK+	images	40	96 %
[31], [32]	PCA, LDA	Six	CK+	193 images	9	92 % 75 %
[33]	SVM	Sourire spontané et sourire sournois	MMI	Vidéos	52	94 %
[34]	HOG, SVM	Six	CK+	images		90 %
[35]	HMMS	Sept	-Propre Base Cohn-Kanade	Vidéos	5 sujets 53 sujets	84,46 % 58,63 %
[36]	MAA	Sept	CMU	166 images	30 sujets	84,34 %
[37]	Codification Dérègles	Six		images	8 sujets	92 %
[38]	Multiboosts, SVM linéaire	Six	BU-3DDFE	Scans 3D	60 sujets	97,75 % 98,81 %
[39]	CNN	Sept	Fer 2013	images	35000 sujets	66.67 %
[40]	CNN	Six	CK+	images	123 sujets	91 %

Tableau 2: Synthèse des travaux développés.

4. Réseaux de neurones

4.1 Définition

Un réseau de neurones artificiels est un système dont la conception est à l'origine inspirée du fonctionnement des neurones biologiques, et qui par la suite s'est rapproché des méthodes statistiques. Les réseaux de neurones artificiels sont des réseaux fortement connectés par des processeurs élémentaires fonctionnant en parallèle. Chaque processeur élémentaire (neurone artificiel) calcule une sortie unique sur la base des informations qu'ils reçoivent.

4.2 Topologie

Chaque réseau de neurones est connecté entre eux de diverses manières. De la figure suivante, nous pouvons distinguer deux familles de réseaux de neurones : non bouclés ou statiques (a) et (b) et bouclés (dynamiques) (c) et (d).

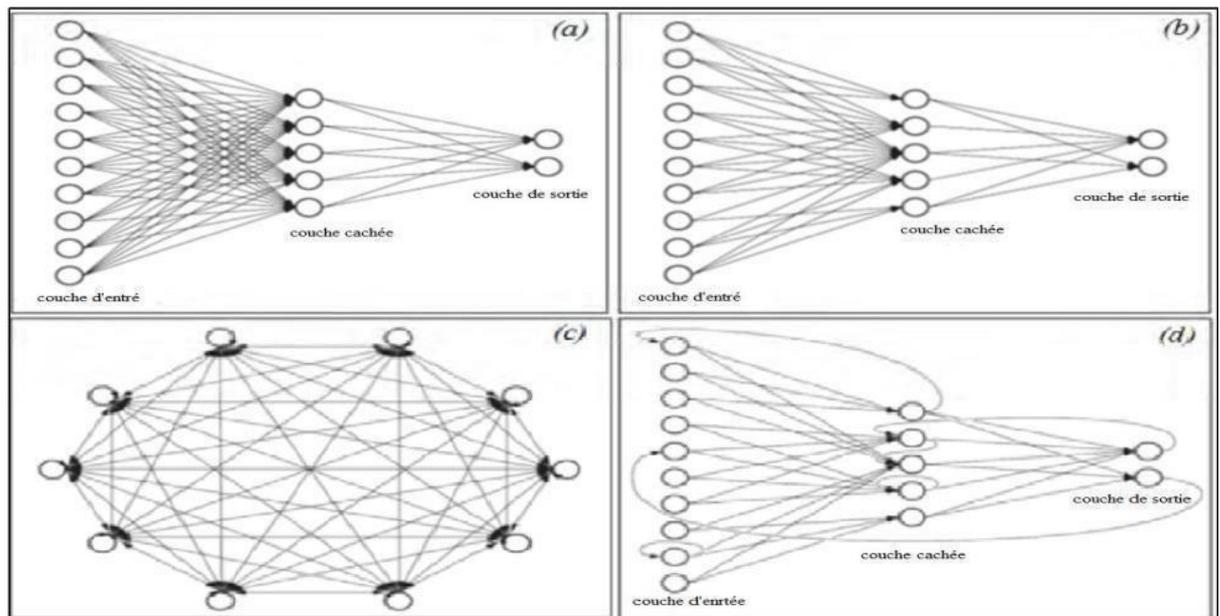


Figure 9: Topologie des Réseaux de neurones artificiels.[41]

4.3 L'apprentissage en profondeur (deep Learning) :

Le Deep Learning est un nouveau domaine de recherche de la machine Learning (ML), qui a été introduit dans le but de rapprocher le ML de son objectif principal à savoir : l'intelligence artificielle. Il concerne les algorithmes inspirés par la structure et du fonctionnement du cerveau. Ils peuvent apprendre plusieurs niveaux de représentation dans le but de modéliser des relations complexes entre les données.

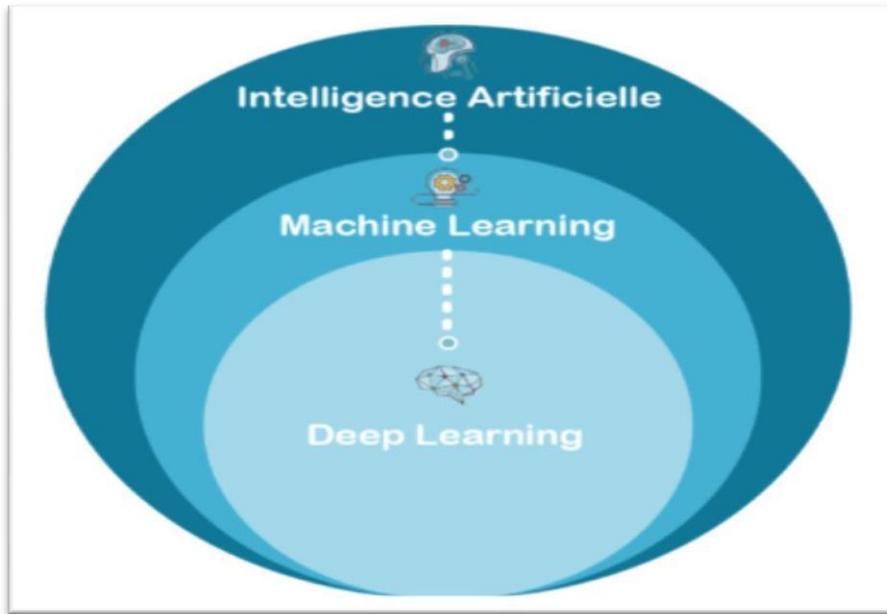


Figure 10: La relation entre l'intelligence artificielle, le ML et le deep Learning.

4.3.1 Définition

L'apprentissage en profondeur est un ensemble d'algorithmes d'apprentissage automatique qui tentent d'apprendre à plusieurs niveaux, correspondant à différents niveaux d'abstraction. Il a la capacité d'extraire des caractéristiques à partir des données brutes grâce aux multiples couches de traitement composé de multiples transformations linéaires et non linéaires et apprendre sur ces caractéristiques petites à petit à travers chaque couche avec une intervention humaine minimale [41][42][43][44][45][46] [47].

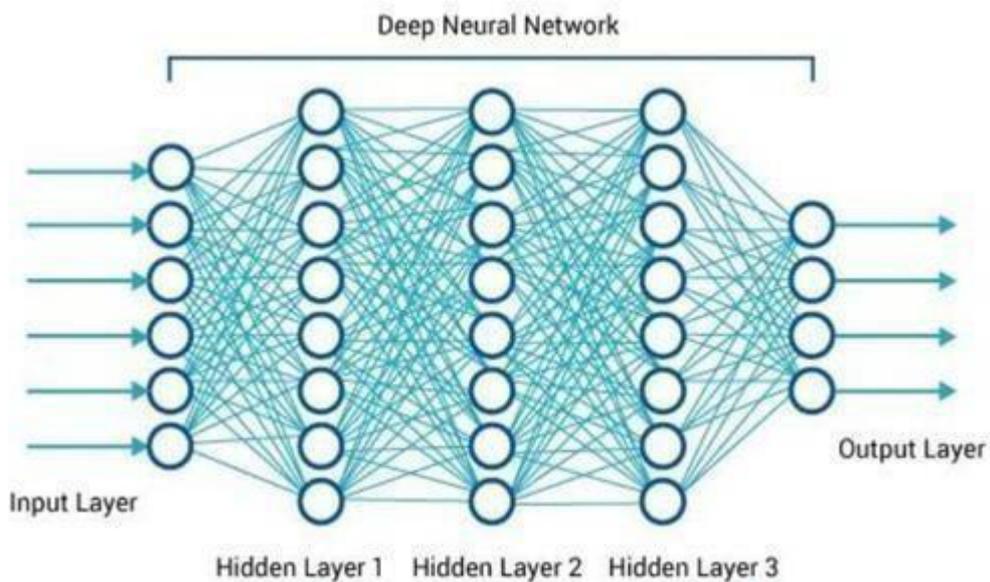


Figure 11 : Schéma illustratif de DL avec plusieurs couches [42].

4.3.2 Pourquoi le choix deep Learning :

Tout d'abord les différents algorithmes du deep Learning ne sont apparus qu'à l'échec de l'apprentissage automatique tentant de résoudre une grande variété de problèmes de l'intelligence artificielle (l'IA):

- Afin d'améliorer le développement des algorithmes traditionnels dans de telles tâches de l'IA.
- De développer une grande quantité de données telle que les bigdata.
- De s'adapter à n'importe quel type de problème.
- D'extraire les caractéristiques de façon automatique

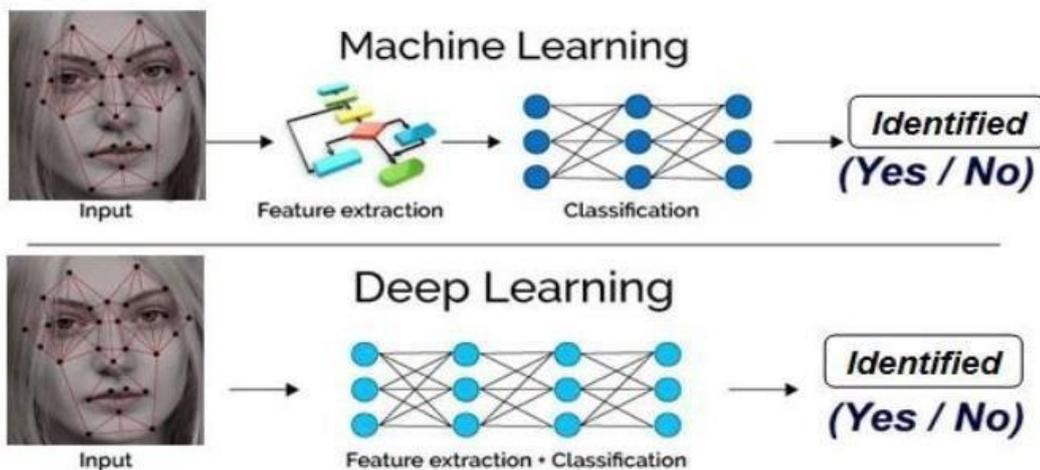


Figure 12 : Comparaison entre la machine Learning et le Deep Learning.

4.4 Réseaux de neurones convolutifs

Convolutional Neural Network (CNN) (réseaux de neurones convolutifs) sont un type de réseau de neurones spécialisés pour le traitement de données ayant une topologie semblable à une grille. Qui se sont avérés très efficaces dans des domaines tels que la reconnaissance et la classification d'images et vidéos. CNN a réussi à identifier les visages, les objets, panneaux de circulation et auto-conduite des voitures [48]. Récemment, les CNN ont été efficaces dans plusieurs tâches de traitement du langage naturel (telles que la classification des phrases) [49].[50][51].

Dans le ML, un réseau convolutif est un type de réseau de neurones feed-forward, il a été inspiré par des processus biologiques [52]. Il existe quatre (4) principales opérations illustrées dans le CNN à savoir :

- La couche convolution.

- La couche Rectifiée de Linear Unit.
- La couche Pooling.
- La couche entièrement connectée.

4.4.1 Modèle génératif

Si les modèles discriminatifs comme (CNN, RNN) sont utilisés pour prédire les données du label et de l'entrée, tant disque le modèle génératif décrit comment générer les données, il apprend et fait des prédictions en utilisant la loi de Bayes [53].

Cependant les modèles génératifs sont capables de bien plus que la simple classification comme par exemple générer de nouvelles observations.

Voici quelques exemples de modèle génératif:

- Boltzmann Machines [54]
- Restricte de Boltzmann Machines [55][56]
- Deep Belief Networks [57]
- Deep Boltzmann Machines [58][59]
- Generative Adversarial Networks •Generative Stochastic Networks[60]
- Adversarial autoencoders [61]

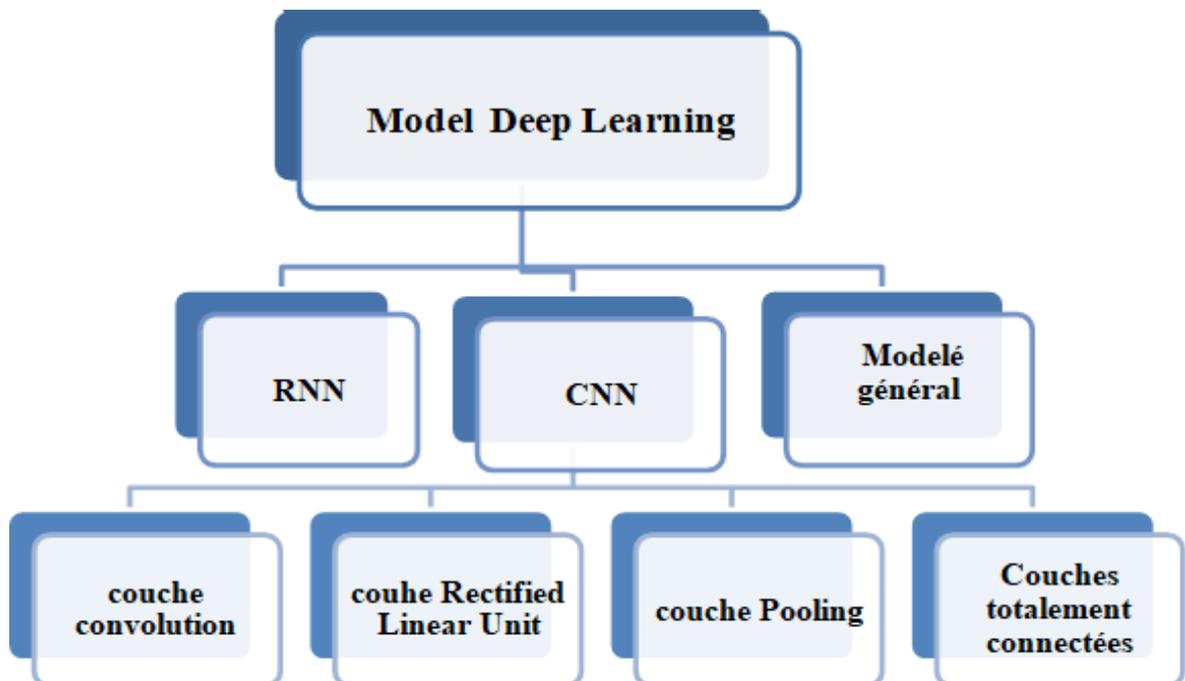


Figure 13: Différents modèles du Deep Learning.

4.4.2 Architecture du système

Nous allons commencer par entraîner un réseau de neurones convolutif des plus classiques, qui possède en entrée une base de données et paramètres entraînaibles et il consiste en des couches convolutives, dont à chaque couche convolutives est placée une couche de Max pooling et enfin une couche totalement connectée. Pour un apprentissage plus rapide, ReLusera la fonction d'activation.

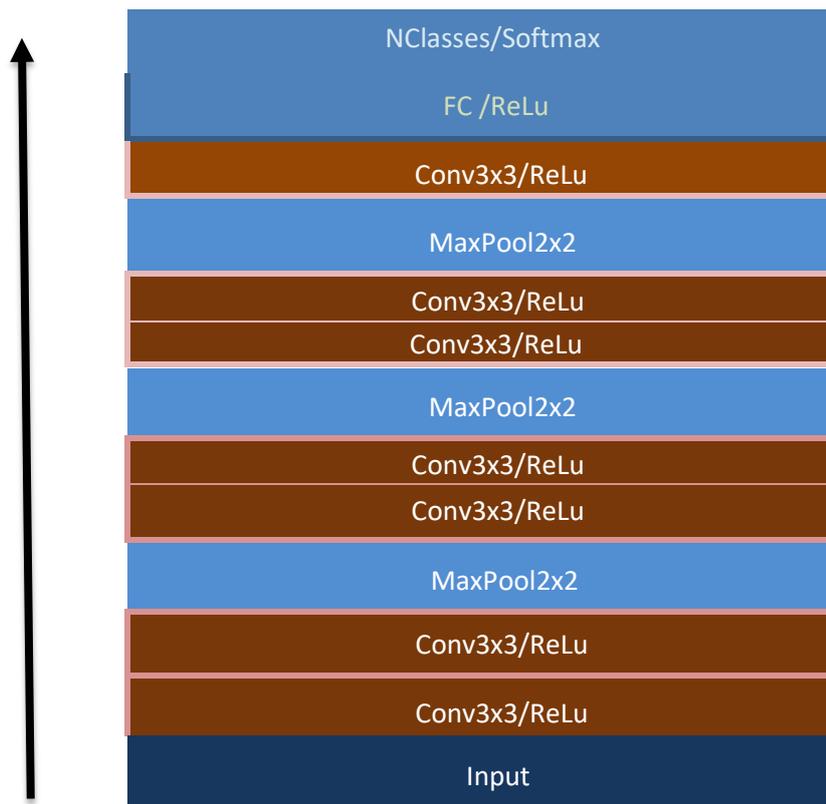


Figure 14 : Architecture générale d'un CNN.

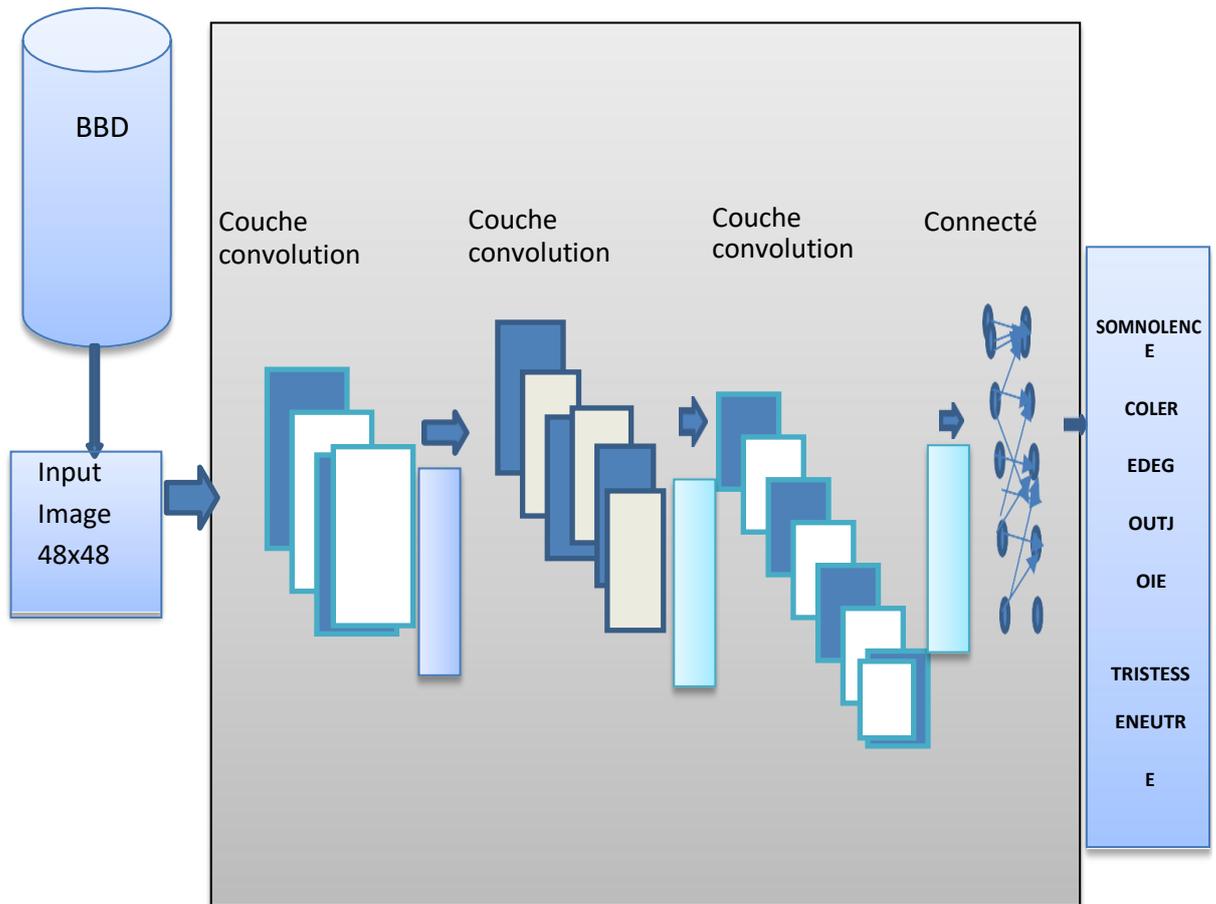


Figure 15 : Architecture proposée.

4.4.3 Différentes couches de CNN

Il existe quatre principales opérations dans le CNN, tout en ayant un rôle bien défini.

a) L'étape de convolution

La **convolution** est le cœur du réseau de neurones convolutif, comme vous vous en doutez. À l'origine, une convolution est un outil mathématique (on parle de produit de convolution) très utilisé en retouche d'image, car il permet d'en faire ressortir l'extraction des caractéristiques à partir des images d'entrées, afin d'appliquer un bon **filtre**. En fait, une convolution prend simplement en entrée une image et un filtre (qui est une autre image), effectue un calcul, puis renvoie une nouvelle image (généralement plus petite). [50]

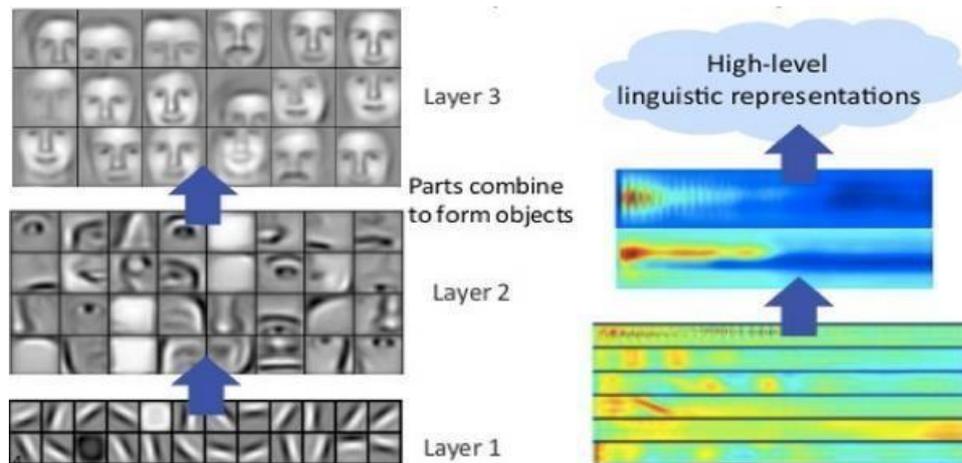


Figure 16 : Phase de convolution [50]

b) Les différentes convolutions

Il existe plusieurs types de convolutions, même si en général on utilise celle de base, il peut s'avérer utile de connaître les outils à notre disposition.

- La **convolution classique**, qui représente le décalage du noyau entre chaque calcul, et le **padding** qui est la manière dont on peut « dépasser » de l'image pour appliquer la convolution.
- La **dilate de convolution**, identique à la convolution à ceci près que le kernel est éclaté (on prend, par exemple, un pixel sur deux pour calculer la convolution). Il y a un paramètre supplémentaire : la dilation **rate**, qui est le nombre de pixels à ignorer.
- La **transpose convolution**, qui construit la sortie comme si on inversait une convolution sur l'image.
- La **séparable convolution**, qui est une convolution décomposable en convolutions plus simples.

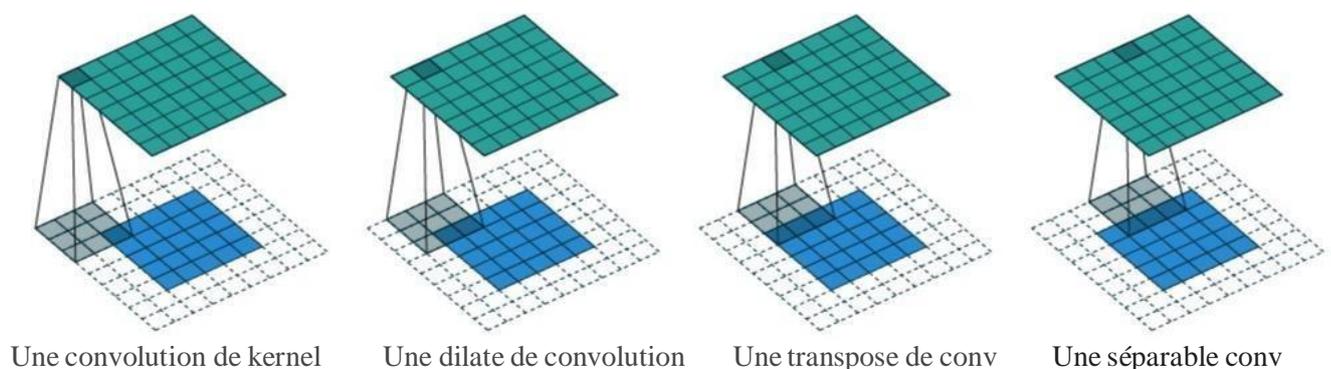


Figure 17 : les différentes convolution

c) Fonctions d'activation

Une fonction d'activation est une fonction mathématique appliquée à un signal en sortie d'un neurone artificiel. Il dérive de l'équivalent biologique qui signifie "**potentiel d'activation**", lorsque le seuil de stimulation aura été atteint entraîne une réponse du neurone. Son but principal est de pouvoir permettre aux réseaux de neurones d'apprendre des fonctions plus complexes qu'une simple régression linéaire, car le simple fait de multiplier les poids d'une couche cachée est juste une transformation linéaire .

- **Rectifie de Linear Unit (ReLU)**

Elle est utilisée après chaque opération de convolution, ou toutes les valeurs de pixels négatifs sont mises à zéro. Le but de ReLU est d'introduire la non-linéarité dans notre CNN, puisque la plupart des données du monde réel, puisque la plupart de caractéristiques appliquées à l'une des cartes d'entrée donne une carte de sortie qui est également appelée carte de caractéristiques rectifiées. [51]

d) L'étape de pooling

Elle permet de réduire la dimension de chaque carte de caractéristiques, mais conserve l'information la plus importante. IL peut être de différents types : max, moyenne, somme, etc. En cas de Pooling max, on définit un voisinage spatial (par exemple, une fenêtre 2×2) et de prendre le plus grand élément dans cette fenêtre.

La fonction de Pooling consiste à réduire progressivement la taille de la carte de caractéristiques rectifiée, en particulier, pooling:

- ✚ Rend les représentations d'entrée plus petites et plus faciles à gérer,
- ✚ Réduit le nombre de paramètres et les calculs dans le réseau [52],
- ✚ Rend le réseau invariant aux petites transformations, les distorsions et les translations dans l'image d'entrée (une faible distorsion en entrée ne change pas la sortie de la Pooling-car nous prenons le maximum/ valeur moyenne dans un voisinage local).
- ✚ Nous aide à arriver à une représentation presque invariante à l'échelle de notre image [53], [54].

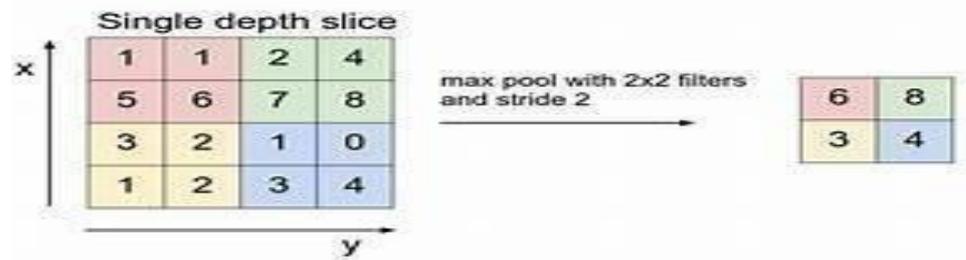


Figure18 : Exemple de description de la couche pooling [53].

Il existe plusieurs types de pooling :

- Le « max pooling », qui revient à prendre la valeur maximale de la sélection. C'est le type le plus utilisé, car il est rapide à calculer (immédiat), et permet de simplifier efficacement l'image
- Le « mean pooling » (ou average pooling), soit la moyenne des pixels de la sélection : on calcule la somme de toutes les valeurs et on divise par le nombre de valeurs. On obtient ainsi une valeur intermédiaire pour représenter ce lot de pixels
- Le « sum pooling », c'est la moyenne sans avoir divisé par le nombre des valeurs (on ne calcule que leur somme).

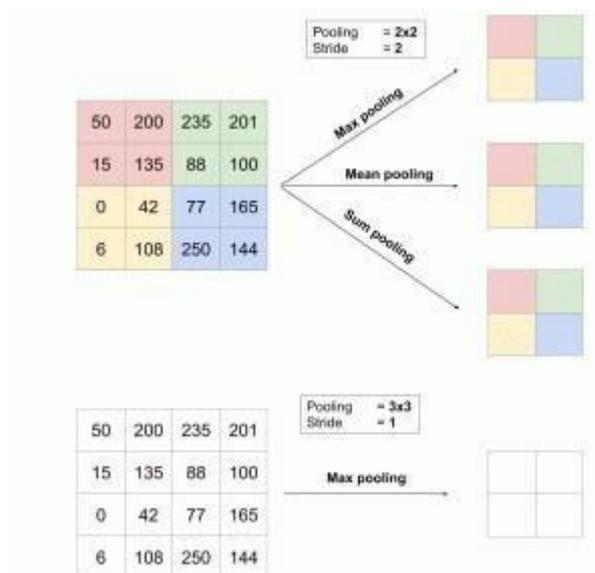


Figure 19: Calcul du pooling sur une image 4x4 [54].

• **La mise à-plat**

Il consiste simplement à mettre à bout toutes les images (matrices) que nous avons pour

en faire un (long) vecteur. Les pixels (en réalité ce ne sont plus des images ou des pixels, mais des matrices de nombres, donc les pixels sont ces nombres) sont récupérés ligne par ligne et ajoutés au vecteur final.

e) La couche totalement connectée

La couche entièrement connectée est un traditionnel perceptron multicouche (Multi Layer Perceptron) utilisant une fonction d'activation notamment appelée « soft max » dans la couche de sortie (d'autres classificateurs comme SVM peuvent également être utilisés). Le terme « entièrement connecté » implique que chaque neurone dans la couche précédente est connecté à chaque neurone sur la couche suivante.

La sortie des couches de convolution et de Pooling représente les fonctions de haut niveau de l'image d'entrée. Le but de la couche entièrement connectée est de pouvoir utiliser ces fonctions pour classer l'image d'entrée dans différentes classes en fonction de l'ensemble de données d'apprentissage. [55]

5. Conclusion

Dans ce chapitre, nous avons présenté quelques méthodes existantes sur la reconnaissance de l'expression faciale. Ainsi nous avons eu à définir qu'est-ce qu'une expression faciale et qu'elles sont les plus utilisés, et qu'est-ce que le Deep Learning et comment il se différencie des algorithmes de ML traditionnelles, Nous avons également parlé des trois familles majeures de modèle savoir les réseaux convolutifs (PRINCIPALEMENT : les concepts principaux de réseau de neurone à convolution (CNN) en termes simples.), les réseaux récurrents ainsi que les modèles génératifs.

Chapitre III: Implémentation et résultats expérimentaux

1. Introduction

L'objectif de ce chapitre est de présenter les étapes de l'implémentation de l'approche proposée dans le cadre d'un système de reconnaissance des expressions faciales. Nous commençons tout d'abord par la présentation des ressources, du langage et de l'environnement de développement que nous avons utilisé. Puis les étapes de la réalisation du modèle et on termine par les tests effectués.

Ce chapitre est composé de deux parties, l'implémentation du système et les résultats expérimentaux des tests.

2. Environnement de développement

2.1 Kaggle

Kaggle est une plateforme web qui accueille la plus grande communauté de Data Science au monde, avec plus de 536 000 membres actifs dans 194 pays et reçoit près de 150 000 soumissions par mois, et qui lui fournit des outils et des ressources puissants pour aider à atteindre tous les progrès de science des données. Kaggle, de la même manière que Data scientist, offre un environnement Jupyter Notebooks personnalisable et sans configuration. Sont accessibles gratuitement des GPU et une grande quantité de données et de codes publiés par la communauté [w1].

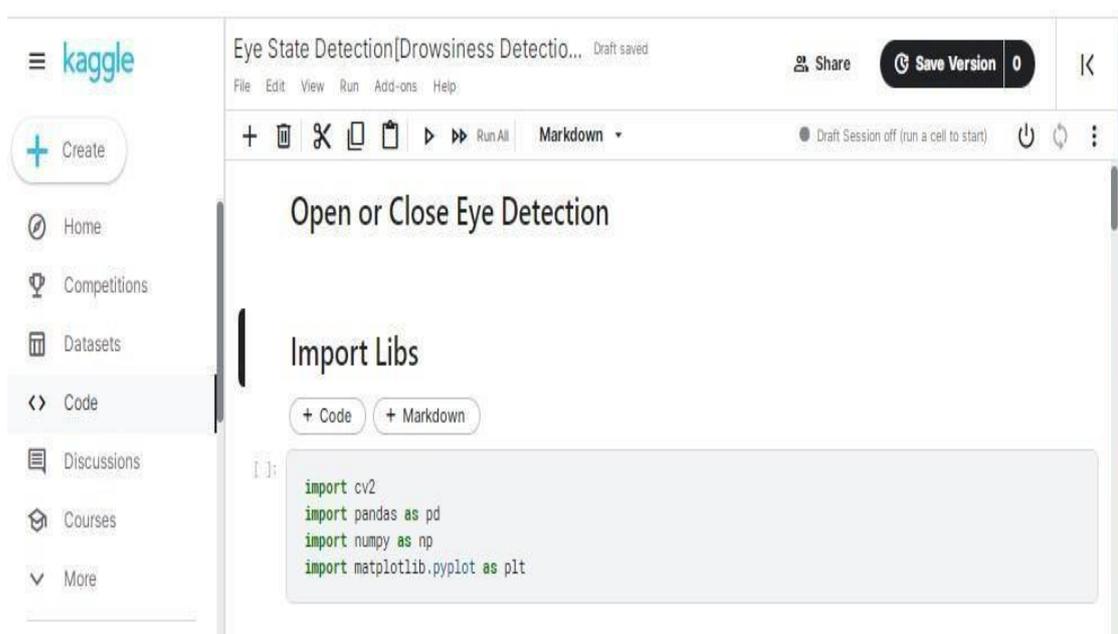


Figure 20: Environnement de KAGGLE

2.2 Pourquoi les GPU ?

Pour la plupart des approches de l'apprentissage en machine ou l'apprentissage en profondeur, les GPU sont essentiels en raison de la quantité de données sur laquelle vous devez opérer. Le calcul mathématique de Deep Learning sur un CPU peut prendre des mois ! Mais ces calculs peuvent être assignés à des GPU pour un calcul plus rapide. Il n'est pas recommandé de former un réseau de neurones sur un processeur, un GPU est nécessaire pour un calcul élevé.

2.3 Python

Python est un langage de programmation de haut niveau utilisé pour la programmation générale. Créé par Guido van Rossum et sorti en 1991, Python a une philosophie de conception qui met l'accent sur la lisibilité du code, notamment en utilisant des espaces importants.

Il fournit des constructions qui permettent une programmation claire à petite et à grande échelle. Python dispose d'un système de type dynamique et d'une gestion automatique de la mémoire. Il prend en charge de multiples paradigmes de programmation, y compris orientés objet, impératifs, fonctionnels et procéduraux, et dispose d'une bibliothèque standard vaste et complète. Les interpréteurs de Python sont disponibles pour de nombreux systèmes d'exploitation [w2].



Figure 21: Le logo de python.

3. Bibliothèques utilisées

3.1 OpenCV (Open Source Computer Vision Library)

Est une bibliothèque proposant un ensemble de plus de 2500 algorithmes de vision par ordinateur spécialisé dans le traitement d'images, accessible au travers d'API pour les langages C, C++, et Python. Elle est distribuée sous une licence BSD (libre) pour les plateformes Windows, GNU/Linux, Android et MacOS [w3], nous avons utilisé cette bibliothèque pour la détection du visage à partir des images introduites.

3.2 Numpy

Est une bibliothèque permettant d'effectuer des calculs numériques avec Python. Elle introduit une gestion facilitée des tableaux de nombres, des fonctions sophistiquées (diffusion), on peut aussi l'intégrer le code (C /C ++)et Fortran [w4] [w5]

3.3 Matplotlib

Est une bibliothèque de traçage pour le langage de programmation Python et son extension mathématique numérique NumPy. Il fournit une API orientée objet permettant d'incorporer des graphiques dans des applications à l'aide de kits d'outils d'interface graphique à usage général tels que Tkinter, wxPython, Qt ou GTK +.

3.4 Keras

Keras est une bibliothèque open source écrite en python et permettant d'interagir avec les algorithmes de réseaux de neurones profonds et de machine Learning, notamment Tensor flow et Theano. Elle a été initialement écrite par François Chollet [w6].

3.5 Tensor Flow

Tensor Flow est une bibliothèque open source, permettant d'exécuter des applications de machine learning et de deep learning. Cet outil dédié à l'apprentissage automatique a été développé par Google, et est fortement utilisé dans le domaine de l'intelligence artificielle (IA). Ainsi, des professionnels comme des novices peuvent créer des modèles de machine learning ou de deep learning pour optimiser les capacités de leur matériel [w7].

3.6 Pandas

Pandas est une bibliothèque open source sous licence BSD fournissant des structures de données hautes performances et faciles à utiliser, ainsi que des outils d'analyse des données pour le langage de programmation Python.

Pandas est un projet sponsorisé par Num FOCUS. Cela contribuera au succès du développement de pandas en tant que projet open source de classe mondiale et permettra de faire un don au projet [w8].

4. BDD

La détection des yeux et de leurs parties, l'estimation du regard et la fréquence des clignements des yeux sont des tâches importantes en vision par ordinateur. Au cours des dernières années, nous avons résolu ces problèmes dans le domaine du comportement du conducteur, ce qui entraîne l'acquisition de nombreuses données de test acquises en conditions réelles. Par conséquent, nous introduisons le MRL Eye Data set, l'ensemble de données à grande échelle d'images de l'œil humain. Cet ensemble de données contient des images infrarouges en basse et haute résolution, toutes capturées dans diverses conditions d'éclairage et par différents appareils. L'ensemble de données est adapté pour tester plusieurs fonctionnalités ou classificateurs pouvant être entraînés. Afin de simplifier la comparaison des algorithmes, les images sont divisées en plusieurs catégories, ce qui les rend également adaptées à l'entraînement et au test des classificateurs [w9].

Dans la base des données, nous avons annoté les propriétés suivantes (les propriétés sont indiquées dans l'ordre suivant) :

- ⇒ Identifiant du sujet: dans la base des données, nous avons collecté les données de 37 personnes différentes (33 hommes et 4 femmes);
- ⇒ ID d'image: l'ensemble de données se compose de 84 898 images.

Chapitre III: Implémentation et résultats expérimentaux

Exemples d'annotations d'images de l'ensemble de données proposé:

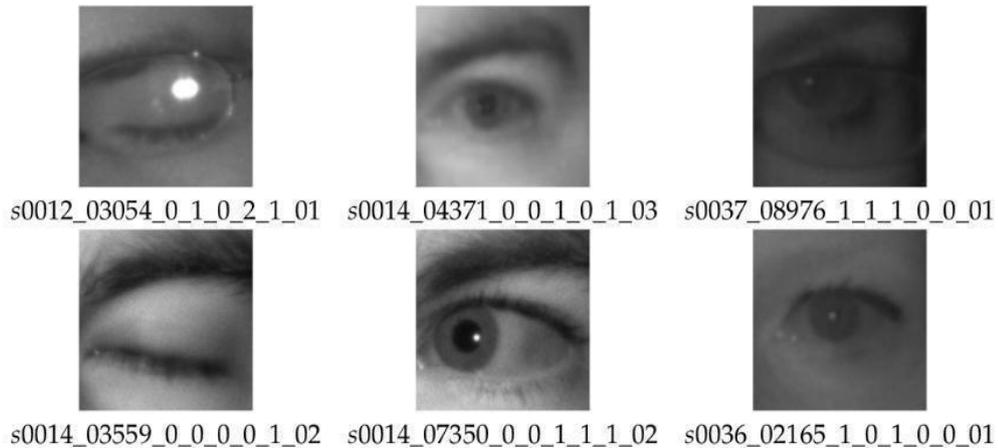


Figure 22: Annotations d'images de l'ensemble de données proposé

5. Implémentation et résultats

Cette section décrit les expériences réalisées et les résultats expérimentaux de notre système sur la reconnaissance des expressions faciales et dans notre cas, nous étudierons l'état de l'œil s'il est fermé ou ouvert.

Notre système est testé en utilisant une base de données affichée ci-dessus.

L'algorithme proposé est principalement réalisé en trois étapes comme ci-dessous :

- 1) Redimensionner les images d'entrée au format
- 2) Construire une structure CNN avec trois couches constituées respectivement de:
- 3) Après avoir extrait toutes les fonctionnalités, nous allons utiliser:

Flatten: elle permet d'aplatir les images

Dense: désigne le nombre que le veut sur une couche, elle prend toujours en entrée le résultat de l'image aplatit de la couche précédente.

Sofmax: qui est un classifieur qui désigne la distribution de la probabilité ou encore que la somme de toutes les probabilités de la sortie doit être égale à 1.

Chapitre III: Implémentation et résultats expérimentaux

5.1 Formation du model avec keras

- La première des choses à faire consiste à tout d'abord configurer l'environnement. Ensuite, installer toutes les bibliothèques citées ci-dessus:
- Importer les bibliothèques installées
- Ensuite, nous allons importer le type de modèle séquentiel de Keras. Il s'agit simplement d'un empilement linéaire de couches de réseaux de neurones à travers la commande: **From keras. Models import Sequential**
- Nous importerons les couches "principales" de Keras. Voici les couches utilisées dans presque tous les réseaux de neurones: `From keras. Layers import Dense, Dropout, Activation, Flatten`
- Pour un entraînement plus efficace sur les données d'image nous allons importer les couches CNN de Keras. Ce sont les couches convolutives qui nous aideront à former notre modèle. **From keras. layers import Convolution 2D, Max Pooling 2D**

Notre modèle contient:

- (1, 48, 48) qui correspond à la (profondeur, largeur, hauteur) de chaque image de notre base de données de MRL
- Conv (3, 3) indique que nous avons utilisé une matrice à trois dimensions.

6. Résultats expérimentaux

Dans cette partie nous nous intéressons aux résultats d'exécution de l'approche développée:

6.1 Notre modèle

Une fois que le modèle est formé, nous allons essayer de lui attribuer certains paramètres:

- ⇒ **Categorical_crossentropy** : permet la réduction de l'erreur tout en utilisant l'algorithme de la descente de gradient.
- ⇒ **Epoch**: désigne le nombre d'itération dans notre base de données.
- ⇒ **Optimiser**: permet de réduire le poids des erreurs.

Chapitre III: Implémentation et résultats expérimentaux

⇒ **Loss:** désigne le taux d'erreur.

⇒ **Accuracy:** désigne le taux de précision.

⇒ **History:** elle permet de donner au modèle toutes les images dont elle aura besoin en fonction du nombre d'itération (epochs) défini, dans le but de réduire le taux d'erreur.

```
Model: "model"
-----
Layer (type)                Output Shape                Param #
-----
input_1 (InputLayer)        [(None, 224, 224, 3)]      0
-----
conv1 (Conv2D)              (None, 112, 112, 32)       864
-----
conv1_bn (BatchNormalization) (None, 112, 112, 32)       128
-----
conv1_relu (ReLU)          (None, 112, 112, 32)       0
-----
conv_dw_1 (DepthwiseConv2D) (None, 112, 112, 32)       288
-----
conv_dw_1_bn (BatchNormaliza) (None, 112, 112, 32)       128
-----
conv_dw_1_relu (ReLU)      (None, 112, 112, 32)       0
-----
conv_pw_1 (Conv2D)          (None, 112, 112, 64)       2048
-----
conv_pw_1_bn (BatchNormaliza) (None, 112, 112, 64)       256
-----
conv_pw_1_relu (ReLU)      (None, 112, 112, 64)       0
-----
conv_pad_2 (ZeroPadding2D)  (None, 113, 113, 64)       0
-----
conv_dw_13 (DepthwiseConv2D) (None, 7, 7, 1024)         9216
-----
conv_dw_13_bn (BatchNormaliz) (None, 7, 7, 1024)         4096
-----
conv_dw_13_relu (ReLU)     (None, 7, 7, 1024)         0
-----
conv_pw_13 (Conv2D)         (None, 7, 7, 1024)         1048576
-----
conv_pw_13_bn (BatchNormaliz) (None, 7, 7, 1024)         4096
-----
conv_pw_13_relu (ReLU)     (None, 7, 7, 1024)         0
-----
global_average_pooling2d (Gl) (None, 1024)               0
-----
reshape_1 (Reshape)         (None, 1, 1, 1024)         0
-----
dropout (Dropout)           (None, 1, 1, 1024)         0
-----
flatten (Flatten)           (None, 1024)               0
-----
dense (Dense)               (None, 1)                  1025
-----
activation (Activation)     (None, 1)                  0
-----
Total params: 3,229,889
Trainable params: 3,208,001
Non-trainable params: 21,888
```

Figure 23: Résultat du modèle obtenu.

Chapitre III: Implémentation et résultats expérimentaux

6.2 Les résultats de l'entraînement du modèle:

Si nous faisons une comparaison de la figure 5 nous constatons qu'en traces deux itérations nous constatons que le taux d'erreur baisse tant dis que notre précision augmente cela signifie que notre modèle a été bien entrainé et réponds d'ailleurs à la définition du réseau de neurones qui disait que plus le réseau de neurones est profond meilleur sont ses performances.

```
new_model.fit(X,Y, epochs = 2, validation_split = 0.1)
#Note: Increase the number of epoch to get more appropriate result, accuracy.

Epoch 1/2
113/113 [=====] - 24s 164ms/step - loss: 0.0693 - accuracy: 0.9747 - val_loss: 0.8277 - val_a
ccuracy: 0.8425
Epoch 2/2
113/113 [=====] - 17s 146ms/step - loss: 0.0037 - accuracy: 0.9991 - val_loss: 0.0086 - val_a
ccuracy: 0.9975
<tensorflow.python.keras.callbacks.History at 0x7f0bc0511910>
```

Figure 24: Résultat de l'entraînement du modèle avec les images.

6.3 Les résultats affichés:

```
Datadirectory = '../input/mri-dataset/train'
Classes = ['Closed_Eyes', 'Open_Eyes']
# Datadirectory = 'dataset_new/train/'
# Classes = ['Closed', 'Open']
for category in Classes:
    path = os.path.join(Datadirectory, category)
    for img in os.listdir(path):
        img_array = cv2.imread(os.path.join(path,img), cv2.IMREAD_GRAYSCALE)
        backtorgb = cv2.cvtColor(img_array,cv2.COLOR_GRAY2RGB)
        plt.imshow(img_array, cmap="gray")
        plt.show()
        break
    break
```

Figure 25: Définissons les classes de la BDD de test et d'entraînement.

La figure désigne le résultat de la base de données et l'entraînement et phase de test.

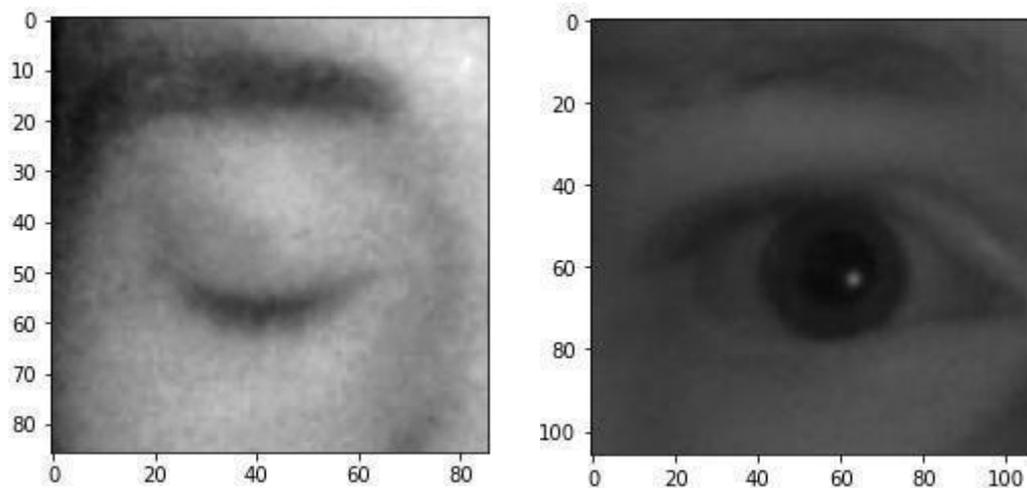


Figure 26: Les états obtenus de la BDD de test et d'entraînement.

6.4 Le teste

Et maintenant, en va tester si l'œil est fermé ou ouvert, comme la valeur varie de 0 à 1, nous l'œil est détecté comme fermé, alors la valeur sera plus proche de zéro. Et si l'œil est détecté comme ouvert, il sera proche de 1

6.4.2 L'œil fermé

```
img_array = cv2.imread('../input/mr1-dataset/train/Closed_Eyes/s0001_00006_0_0_0_0_01.png', cv2.IMREAD_GRAYSCALE)
backtorgb = cv2.cvtColor(img_array, cv2.COLOR_GRAY2BGR)
new_array = cv2.resize(backtorgb, (img_size, img_size))

X_input = np.array(new_array).reshape(1, img_size, img_size, 3)

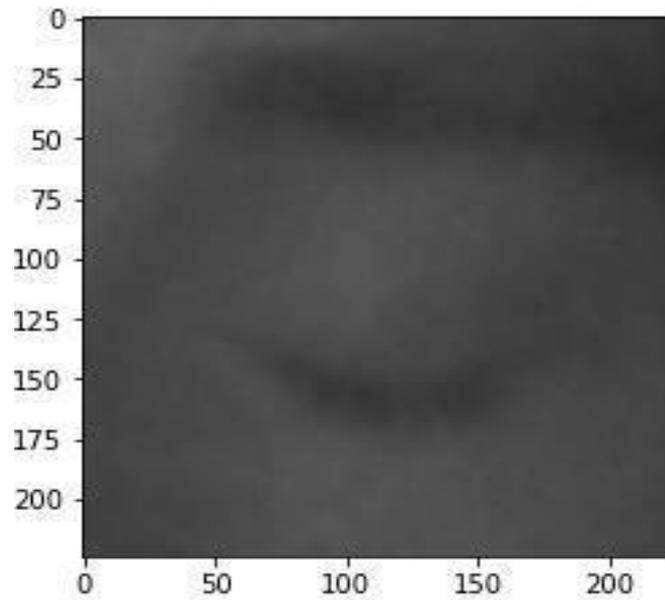
X_input.shape

(1, 224, 224, 3)

plt.imshow(new_array)
```

Figure 27: L'affichage des images

Chapitre III: Implémentation et résultats expérimentaux



```
prediction = new_model.predict(X_input)
```

```
prediction
```

```
array([[1.0847013e-11]], dtype=float32)
```

Figure 28: Les résultats de teste: Œil fermer

Puisque l'œil est fermé, nous avons obtenu la valeur de prédiction plus proche de 0, c'est-à-dire $1,084 \text{ e}^{-11}$.

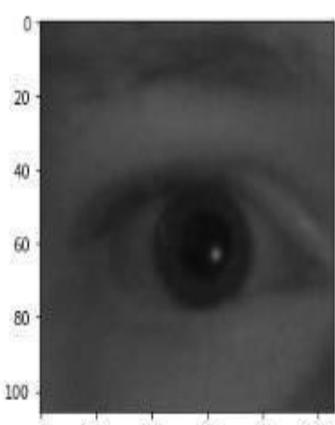
6.4.2 L'œil ouverte

```
img = cv2.imread('../input/mrl-dataset/train/Open_Eyes/s0034_00605_0_0_1_0_0_02.png')
```

+ Code + Markdown

```
plt.imshow(cv2.cvtColor(img, cv2.COLOR_BGR2RGB))
```

<matplotlib.image.AxesImage at 0x7feab6714b10>



The image shows a grayscale plot of an eye. The x-axis and y-axis both range from 0 to 100, with major ticks every 20 units. The eye is centered in the plot, with the pupil and iris clearly visible. The plot is displayed in a Jupyter Notebook environment, with the code above it and the plot title below it.

Figure 29: Les résultats de teste: œil ouverte

Nous avons d'abord lisez l'image pour tester est ce que l'œil est fermer ou ouverte.

```
faceCascade = cv2.CascadeClassifier(cv2.data.harcascades + 'haarcascade_frontalface_alt.xml')
eyeCascade = cv2.CascadeClassifier(cv2.data.harcascades + 'haarcascade_eye.xml')
```

```
gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
```

```
eyes = eyeCascade.detectMultiScale(gray, 1.1, 4)
```

```
for (x, y, w, h) in eyes:
    cv2.rectangle(img, (x,y), (x+w, y+h), (0, 255, 0), 1)
```

Chapitre III: Implémentation et résultats expérimentaux

- ✓ Ensuite, nous effectuons le processus de numérisation pour clarifier l'emplacement de la pupille de l'œil
- ✓ S'il est trouvé, il agrandira l'iris.
- ✓ S'il ne l'a pas trouvé ou ne l'a pas lu, il nous dira qu'il n'y a pas d'œil (no eye detected).

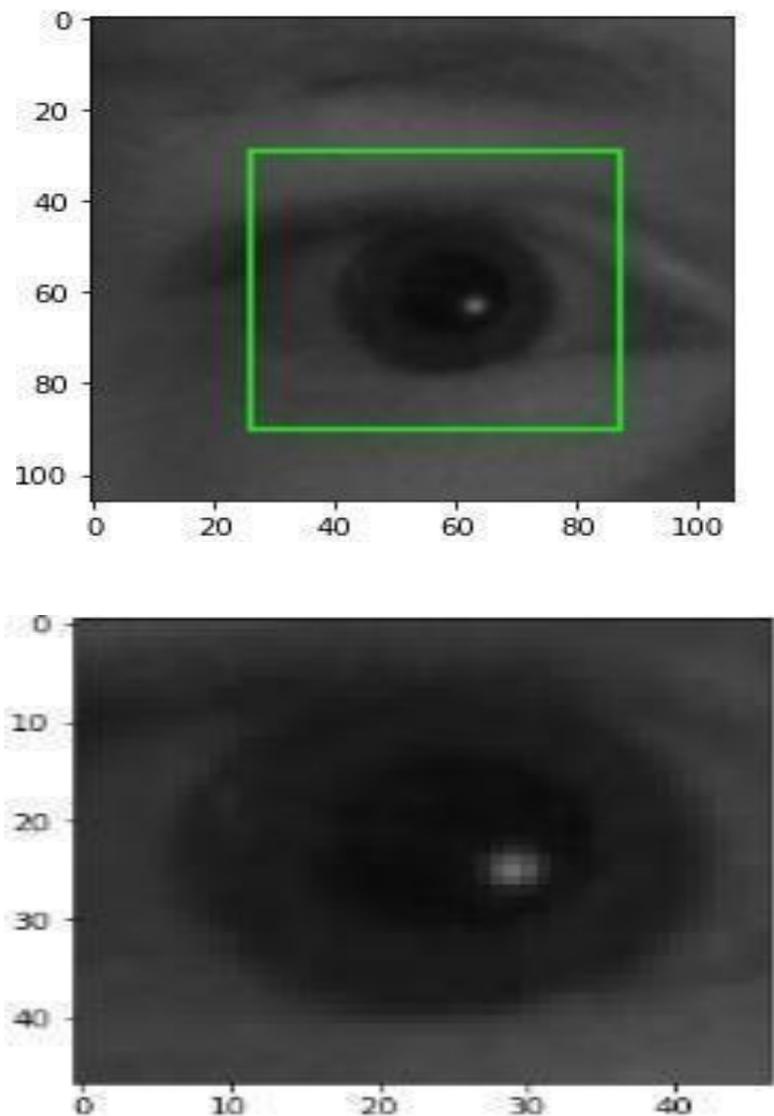


Figure 30: Résultat de test

Chapitre III: Implémentation et résultats expérimentaux

- ✓ Après avoir lu l'œil, il va calculer la précision ou le niveau d'ouverture des yeux

```
eyes_roi.shape  
  
(47, 47, 3)  
  
final_img = cv2.resize(eyes_roi, (224,224))  
final_img = np.expand_dims(final_img, axis=0)  
final_img = final_img/255.0  
  
new_model.predict(final_img)  
  
array([[1.]], dtype=float32)
```

- ✓ Puisque l'œil est ouvert, nous avons obtenu la valeur de prédiction égale à 1.

7. Conclusion:

Dans ce chapitre, nous avons présenté le système de détection de somnolence que nous avons réalisé dans ce projet. Ce système est basé sur l'état des yeux (ouvert ou fermé).

Il existe plusieurs méthodes de détection du visage et des yeux. L'utilisation de ces méthodes dépend du problème à résoudre. Dans notre travail, nous avons utilisé un système basé sur l'apprentissage en profondeur pour la détection de la somnolence (la détection des yeux), L'approche utilise une architecture de réseau de neurones convolutionnels.

Les tests effectués sur une base des données de taille grande étant encourageante.

Chapitre III: Implémentation et résultats expérimentaux

Bibliographies

- [1] https://fr.wikipedia.org/wiki/Expression_faciale.
- [2] DANALET, Antonin. Modèles de choix discrets pour la reconnaissance des expressions faciales statiques. 2007.
- [3] ABDAT, Faiza. Reconnaissance automatique des émotions par données multimodales : expressions faciales et signaux physiologiques. Université de Metz, France, 2010.
- [4] Y.Ma,X.Ding. Face Detection based on hierarchical support vector machines. s.l.: IEEE, 2002.
- [5] Hanane, OUAMANE. Identification de reconnaissance faciale avec des expressions. Mémoire de Fin d'Etudes de diplôme Master. Biskra : s.n. 07 Juin 2012.)
- [6] Ababsa, SouhilaGuerfi. Authentification d'individus par reconnaissance de caractéristiques biométriques liées aux visages 2D/3D. THÈSE pour obtenir le titre de Docteur de l'Université Evry Val d'Essonne. France : s.n. 2008.
- [7] Plutchik, R. (1980). A general psych evolutionary theory of emotion. In Theories of emotion (pp. 3-33). Academicpress)
- [8] <https://www.decodeurdunonverbal.fr/test-les-microexpressions-7-emotions-universelles-a-decoder/>
- [9] S. Hu and G. Zheng, "Driver Drowsiness Detection with Eyelid related Parameters by Support Vector Machine," Expert Systems with Applications, vol. 36, pp. 7651-7658, 2009..
- [10]P. Antoine, C. Sylvie and A. Caplier, "EOG-based drowsiness detection:Comparison between a fuzzy system and two supervised learning classifiers," The International Federation of Automatic Control, 2011
- [11] H. Ouamane et M. Benatia, « Identification De Reconnaissance Faciale Avec Des Expressions », Mémoire de Master en Electronique, Option : Télécommunication, Université Mohamed Khider Biskra, Soutenu le 07 Juin 2012.
- [12] H.Daoui et A.Elomaret Bentaouza Meriem Chahinez, Identification de personne grâce à la reconnaissance faciale par la géométrie du visage et la classification, mémoire de master, Mostaganem, 2014/2015.
- [13] D. Saigaa, Contribution à l'authentification d'individus par la reconnaissance de visages, thèse de Doctorat d'état en automatique, Université Mohamed Khieder Biskra Algérie. Novembre 2006).
- [14] Gonzalez, Rafael C « Digital image processing » volume2 1987.
- [15] Morizet, N. (2009). *Reconnaissance biométrique par fusion multimodale du visage et*

Bibliographies

- del'iris*(Thèsededoctorat).
- [16] Brown, G. D., Yamada, S., & Sejnowski, T. J. (2001). Independent component analysis at the neural cocktail party. *Trends in neurosciences*, 24(1), 54-63.
- [17] Sagi, B., Nemat-Nasser, S. C., Kerr, R., Hayek, R., Downing, C., & Hecht-Nielsen, R. (2001). A biologically motivated solution to the cocktail party problem. *Neural Computation*, 13(7), 1575-1602.
- [18] Sagi, B., Nemat-Nasser, S. C., Kerr, R., Hayek, R., Downing, C., & Hecht-Nielsen, R. (2001). A biologically motivated solution to the cocktail party problem. *Neural Computation*, 13(7), 1575-1602.
- [19] Kanade, T. (1973). *Picture processing system by computer complex and recognition of human faces*. Kyoto University, Japan (Thèsededoctorat).
- [20] Guo, G., Li, S. Z., & Chan, K. (2000, March). Face recognition by support vector machines. In *Proceedings fourth IEEE international conference on automatic face and gesture recognition (cat. no. PR00580)* (pp. 196-201).
- [21] Cootes, T. F., Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (6), 681-685.
- [22] Ahonen, T., Hadid, A., & Pietikäinen, M. (2004, May). Face recognition with local binary patterns. In *European conference on computer vision* (pp. 469-481). Springer, Berlin, Heidelberg.
- [23] Arca, S., Campadelli, P., & Lanza, R. (2006). A face recognition system based on automatically determined facial fiducial points. *Pattern recognition*, 39(3), 432-443.
- [24] Perlibakas, V. (2006). Face recognition using principal component analysis and log-gabor filters. *Ar Xiv preprint cs/0605025*
- [25] Qi, C., Li, M., Wang, Q., Zhang, H., Xing, J., Gao, Z., & Zhang, H. (2018). Facial expressions recognition based on cognition and mapped binary patterns. *IEEE Access*, 6, 18795-18803.
- [26] Ghimire, D., Lee, J., Li, Z. N., & Jeong, S. (2017). Recognition of facial expressions based on salient geometric features and support vector machines. *Multimedia Tools and Applications*, 76(6), 7921-7946.

Bibliographies

- [27] Lyons, M. J., Budynek, J., & Akamatsu, S. (1999). Automatic classification of single facial images. *IEEE transactions on pattern analysis and machine intelligence*, 21(12), 1357-1362.
- [28] Lyons, M., Akamatsu, S., Kamachi, M., & Gyoba, J. (1998, April). Coding facial expressions with gabor wavelets. In *Proceedings Third IEEE international conference on automatic face and gesture recognition* (pp. 200-205).
- [29] Valstar, M. F., Gunes, H., & Pantic, M. (2007, November). How to distinguish posed from spontaneous smiles using geometric features. In *Proceedings of the 9th international conference on Multimodal interfaces* (pp. 38-45). ACM.
- [30] Mlakar, U., Fister, I., Brest, J., & Potočnik, B. (2017). Multi-objective differential evolution for feature selection in facial expression recognition systems. *Expert Systems with Applications*, 89, 129-137.
- [31] Cohen, I., Sebe, N., Garg, A., Chen, L. S., & Huang, T. S. (2003). Facial expression recognition from video sequences: temporal and static modeling. *Computer Vision and image understanding*, 91(1-2), 160-187.
- [32] Davoine, F., Abboud, B., & Dang, V. M. (2004). Analyse de visages et d'expressions faciales par modèle actif d'apparence. *Traitement du signal*, 1(3).
- [33] Pantic, M., & Rothkrantz, L. J. (2000). Automatic analysis of facial expressions: The state of the art. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (12), 1424-1445.
- [34] Maalej, A., Amor, B. B., & Daoudi, M. (2011, June). Analyse locale de la forme 3D pour la reconnaissance d'expressions faciales.
- [35] M. Quinn, G. Sivesind, G. Reis, "Real-time Emotion Recognition From Facial Expressions", Stanford University 2015.
- [36] Ding, H., Zhou, S. K., & Chellappa, R. (2017, May). Facenet2expnet: Regularizing a deep face recognition net for expression recognition. In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)* (pp. 118-126).
- [37] Deng, L., & Yu, D. (2014). Deep learning: methods and applications. *Foundations and Trends in Signal Processing*, 7(3-4), 197-387.
- [38] Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends in Machine Learning*, 2(1), 1-127.
- [39] Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review

Bibliographies

- and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8), 1798-1828.
- [40] Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117.
- [41] Le Cun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436.
- [42] Arel, I., Rose, D.C., & Karnowski, T.P. (2010). Deep machine learning - a new frontier in artificial intelligence research. *IEEE computational intelligence magazine*, 5(4), 13-18.
- [43] J. Schmidhuber, "Deep Learning,," Scholarpedia, vol.10, no.11, p.32832, 2015.
- [44] Moualek, D. Y. (2017). *Deep Learning pour la classification des images* (Thèse de doctorat), Université Abou Bakr Belkaid-Tlemcen.
- [45] Y. Bengio, A. Courville, and P. Vincent. Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 38:1798–1828, 2013
- [46] Hubara, I., Courbariaux, M., Soudry, D., El-Yaniv, R., & Bengio, Y. (2017). Quantized neural networks: Training neural networks with low precision weights and activations. *The Journal of Machine Learning Research*, 18(1), 6869-6898.
- [47] Van den Oord, A., Dieleman, S., & Schrauwen, B. (2013). Deep content-based music recommendation. In *Advances in neural information processing systems* (pp. 2643-2651).
- [48] Collobert, R., & Weston, J. (2008, July). A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning* (pp. 160-167).
- [49] Matsugu, M., Mori, K., Mitari, Y., & Kaneda, Y. (2003). Subject independent facial expression recognition with robust face detection using a convolutional neural network. *Neural Networks*, 16 (5-6), 555-559.
- [50] Mikolov, T., Karafiát, M., Burget, L., Černocký, J., & Khudanpur, S. (2010). Recurrent neural network based language model. In *Eleventh annual conference of the international speech communication association*.
- [51] Liu, S., Yang, N., Li, M., & Zhou, M. (2014, June). A recursive recurrent neural network for statistical machine translation. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1491-1500).
- [52] Graves, A., & Jaitly, N. (2014, January). Towards end-to-end speech recognition with recurrent neural networks. In *International conference on machine learning* (pp. 1764-

Bibliographies

1772).

- [53] Karpathy, A., & Fei-Fei, L. (2015). Deep visual-semantic alignments for generating image descriptions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3128-3137).
- [54] Ng, A. Y., & Jordan, M.I. (2002). On discriminative vs. generative classifiers: A comparison of logistic regression and naïve bayes. In *Advances in neural information processing systems* (pp. 841-848).
- [55] Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A learning algorithm for Boltzmann machines. *Cognitive science*, 9(1), 147-169.
- [56] Salakhutdinov, R., Mnih, A., & Hinton, G. (2007, June). Restricted Boltzmann machines for collaborative filtering. In *Proceedings of the 24th international conference on Machine learning* (pp. 791-798).
- [57] Hinton, G.E.(2012).A practical guide to training restricted Boltzmann machines.In *Neural networks: Tricks of the trade* (pp.599-619).Springer, Berlin, Heidelberg.
- [58] Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, 18(7), 1527-1554.
- [59] Salakhutdinov, R., & Hinton, G. (2009, April). Deep boltzmann machines. In *Artificial intelligence and statistics* (pp. 448-455).
- [60] Srivastava, N., & Salakhutdinov, R. R. (2012). Multimodal learning with deep boltzmann machines. In *Advances in neural information processing systems* (pp. 2222- 2230).
- [61] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672-2680).
- [62] Denton, E. L., Chintala, S., & Fergus, R. (2015). Deep generative image models using a Laplacian pyramid of adversarial networks. In *Advances in neural information processing systems* (pp. 1486-1494).
- [63] Zhang, R., Zhu, J. Y., Isola, P., Geng, X., Lin, A. S., Yu, T., & Efros, A. A. (2017). Real-time user guided image colorization with learned Deep priors. *ArXiv preprint arXiv:1705.02999*.
- [64] Owens, A., Isola, P., McDermott, J., Torralba, A., Adelson, E. H., & Freeman, W. T. (2016). Visually indicated sounds. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2405-2413).

Bibliographies

- [65] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

WEB

[W1] <https://www.kaggle.com/>

[W2] <https://www.python.org/>

[W3] <https://opencv.org/about/>

[W4] <http://www.numpy.org/>

[W5] <https://www.courscopython.com/apprendre-numpy.html> 05 mai 2019

[W6] <https://fr.wikipedia.org/wiki/Keras>

[W7] <https://www.journaldunet.fr/web-tech/guide-de-l-intelligence-artificielle>

[W8] <https://pandas.pydata.org/>

[W9] <http://mrl.cs.vsb.cz/eyedataset>

Conclusion générale

Ce travail de mémoire de master réalisé à l'université de Biskra s'inscrit dans le domaine de reconnaissance des expressions faciales plus précisément la détection de somnolence. Pour réaliser ce but, nous avons utilisé les méthodes de l'apprentissage profondi (Deeplearning).

En premier cas, nous avons présenté des généralités sur les expressions faciales. Ensuite, nous avons présenté les différentes méthodes de reconnaissance des expressions faciales (méthode globale, méthode hybride, méthode locale, méthode l'apprentissage profondi, CNN). Enfin, nous avons réalisé un système non intrusif qui surveille l'état des yeux pour détecter l'état du conducteur. Notre système de détection de somnolence est basé sur l'état des yeux. L'application de notre système sur une base de donne et très large a donnée des résultats encourageants.

En perspectives ; pour améliorer les performances de ce travail, nous proposons de :

- Appliquer d'autres bases de données comme YAWDD
- Utiliser ce travail comme une application en temps réel
- Appliquer ce travail sur le domaine de la douleur.

ملخص

يتعلق عملنا في هذه الأطروحة بالتعرف على تعابير الوجه بشكل أكثر دقة للكشف عن نعاس السائق. يعتمد عملنا على أساليب التعلم العميق حيث اخترنا بنية مناسبة مع قاعدة البيانات المستخدمة (MRL)

يعتمد نظام اكتشاف النعاس لدينا على حالة العين. أعطى تطبيق نظامنا على قاعدة بيانات واسعة للغاية نتائج مشجعة

الكلمات المفتاحية: تعبيرات الوجه، اكتشاف الوجه، كشف النعاس، التعلم العميق، سي إن إن

Résumé

Notre travail dans ce mémoire concerne la reconnaissance des expressions faciales plus précisément la détection de somnolence du conducteur. Notre travail est basé sur les méthodes de L'apprentissage en profondeur (Deep learning) où nous avons choisis une architecture convenable avec la base de données utilisée (MRL).

Notre système de détection de somnolence est basé sur l'état des yeux. L'application de notre système sur une base de données très large a donné des résultats encourageants.

Mots clés : Expression faciale, Détection de visage, Détection de la somnolence, Apprentissage en profondeur, CNN.

Abstract

Our work in this thesis concerns the recognition of facial expressions more precisely the detection of driver drowsiness. Our work is based on the methods of. Deep learning where we have chosen a suitable architecture with the database used (MRL).

Our drowsiness detection system is based on eye condition. The application of our system on a very broad database has given encouraging results.

Keywords : Facial Expression, Face Detection, Drowsiness Detection, Deep Learning, CNN.

