



ENSA  
ÉCOLE NATIONALE DES SCIENCES  
APPLIQUÉES  
KHOURIBGA



## Projet de Fin de Module

*Master Big Data et Aide à la Décision*

---

# Reconnaissance des émotions dans les visages

---

*Présenté par*

**EL HAMID Youssef  
KOUALIL Mohammed**

Encadré par : **Pr. EL HADFI Youssef**

# Table des matières

<b>Introduction générale</b>	<b>5</b>
<b>1 Analyse des expressions faciales</b>	<b>6</b>
1.1 Intoduction . . . . .	6
1.2 Définitions . . . . .	6
1.2.1 Emotion . . . . .	6
1.2.2 Expression Faciale . . . . .	6
1.2.3 Six expressions faciales universelles . . . . .	7
1.3 Domaines d'applications . . . . .	7
1.4 Architecture de système de reconnaissance des expressions faciales . . . . .	7
1.4.1 Detection du visage . . . . .	8
1.4.2 Extraction des caracterestiques . . . . .	9
1.4.3 Reconnaissance de l'expression faciale . . . . .	9
1.5 Conclusion . . . . .	9
<b>2 Apprentissage profond</b>	<b>11</b>
2.1 Intoduction . . . . .	11
2.2 Deep learning . . . . .	11
2.3 Réseaux de neurones artificiels . . . . .	11
2.4 Réseaux de neurones convolutifs . . . . .	12
2.5 Evaluation . . . . .	15
2.6 Conclusion . . . . .	16
<b>3 Reconnaissance des émotions faciales par réseaux convolutionnels</b>	<b>17</b>
3.1 Introduction . . . . .	17
3.2 Architecture du réseau . . . . .	17
3.3 Collecte des Données . . . . .	18
3.4 Prétraitement des Données . . . . .	19
3.5 Création du modèle CNN . . . . .	19
3.6 Résultats . . . . .	20
3.7 Évaluation des Performances . . . . .	22
3.8 Conclusion . . . . .	23



# Table des figures

1.1	Émotions primaires exprimées sur le visage. . . . .	7
1.2	Structure de base des systèmes de reconnaissance des émotions faciales. . .	8
1.3	Détection des visages. . . . .	9
2.1	Schéma d'un réseau de neurones artificiel . . . . .	12
2.2	Exemple d'une opération de convolution . . . . .	13
2.3	Max pool 2 x 2 . . . . .	13
2.4	Fully connected Layer . . . . .	14
2.5	Fonction d'activation ReLU . . . . .	14
2.6	Matrice de confusion . . . . .	16
3.1	Architecture du réseau utilisé. . . . .	18
3.2	Exemples d'images de visage de l'ensemble de données . . . . .	18
3.3	l'architecture CNN utilisée. . . . .	20
3.4	Évolution Accuracy et Accuracy-validation sur l'apprentissage. . . . .	21
3.5	Évolution loss et loss-validation sur l'apprentissage. . . . .	21
3.6	Evaluation . . . . .	22
3.7	matrice de confusion . . . . .	23

# Introduction générale

Ces dernières années ont été marquées par une révolution scientifique et technologique, avec la vision par ordinateur émergeant comme un exemple notable. Cette discipline vise à créer des applications intelligentes capables de comprendre et extraire divers objets d'une base de données, ainsi que des informations à partir d'images, de vidéos, etc. Une évolution significative au cours de cette décennie a été le développement de l'informatique affective, introduisant la dimension émotionnelle à l'interaction avec les ordinateurs.

La reconnaissance des émotions faciales, un domaine de recherche actif depuis les travaux de Darwin en 1872, a connu des progrès majeurs. À la fin des années 1970, Suwa et Motoi ont pionnièrement proposé l'analyse automatique des expressions faciales, en suivant le mouvement de 20 points identifiés sur une séquence d'images. Cette activité a perduré jusqu'à aujourd'hui, avec le développement constant d'applications de reconnaissance émotionnelle automatique utilisant diverses méthodes d'apprentissage, parmi lesquelles l'apprentissage profond se distingue par sa capacité à traiter de vastes ensembles de données et à fournir des réponses de haute qualité.

C'est dans ce contexte que nous avons entrepris d'étudier et de créer un modèle capable de reconnaître les sept émotions faciales (colère, dégoût, peur, bonheur, tristesse, surprise, neutralité) à partir d'images, en utilisant un réseau de neurones convolutionnel.

# Chapitre 1

## Analyse des expressions faciales

### 1.1 Introduction

Dans ce chapitre, nous plongerons au cœur de la reconnaissance des expressions faciales, explorant les concepts fondamentaux liés aux émotions et l'architecture globale des systèmes de reconnaissance faciale.

### 1.2 Définitions

#### 1.2.1 Emotion

L'émotion est un état psychologique et physiologique intense qui résulte d'une réponse subjective à des stimuli internes ou externes. Elle se manifeste par des changements dans les pensées, les comportements, et le corps. Les émotions peuvent inclure une gamme de sentiments tels que la joie, la tristesse, la colère, la peur, le dégoût, la surprise, etc. Elles sont souvent accompagnées de réponses physiologiques, comme des changements dans le rythme cardiaque, la respiration, les expressions faciales, et peuvent également influencer le comportement et la prise de décision. Les émotions jouent un rôle crucial dans la vie quotidienne en influençant nos interactions sociales, nos expériences personnelles et notre bien-être émotionnel.

#### 1.2.2 Expression Faciale

L'expression faciale est une mimique faciale chargée de sens. Le sens peut être l'expression d'une émotion, un indice sémantique ou une intonation dans la Langue des Signes. L'interprétation d'un ensemble de mouvements musculaires en expression est dépendante du contexte d'application. Par exemple dans le cas d'une application en interaction Homme-Machine où l'on désire connaître une indication sur l'état émotionnel d'un individu, on cherchera à classifier les mesures en termes d'émotions.

### 1.2.3 Six expressions faciales universelles

Dès le vingtième siècle, six émotions de base ont été définies sur la base d'une étude culturelle et psychologique. Il a été expliqué que les humains perçoivent ces émotions de base de manière similaire, indépendamment de leur culture. Les tentatives visant à décrire les émotions humaines reposent sur l'hypothèse selon laquelle les êtres humains expriment universellement un ensemble d'émotions primaires distinctes, comprenant le bonheur, la tristesse, la peur, la colère, le dégoût et la surprise (voir figure 1.1). En raison de sa simplicité et de sa revendication d'universalité, cette hypothèse des émotions primaires universelles a largement été exploitée dans le domaine de l'informatique affective.



FIGURE 1.1 – Émotions primaires exprimées sur le visage.

## 1.3 Domaines d'applications

De nombreuses études ont été menées sur l'analyse automatique de l'expression faciale en raison de son importance pratique en robotique sociale, traitement médical, surveillance de la fatigue du conducteur et de nombreuses autres interactions homme-machine systèmes. Ainsi que Dans le domaine de la vision par ordinateur et de l'apprentissage automatique. Divers systèmes de reconnaissance de l'expression faciale ont été explorés pour coder les informations d'expression à partir de représentations faciales.

## 1.4 Architecture de système de reconnaissance des expressions faciales

Un système dédié à la reconnaissance automatique des expressions faciales est généralement constitué de trois modules principaux, comme illustré dans la figure 1.2. Le premier module a pour fonction de détecter et enregistrer la région du visage présente dans les images ou les séquences d'images en entrée. Cela peut impliquer l'utilisation d'un détecteur pour repérer le visage dans chaque image ou simplement détecter le visage dans la première image puis suivre son évolution dans le reste de la séquence vidéo. Le deuxième module se charge d'extraire et de représenter les changements faciaux induits par les expressions faciales. Enfin, le dernier module évalue la similarité entre l'ensemble des caractéristiques extraites et un ensemble de caractéristiques de référence. D'autres filtres ou modules de prétraitement des données peuvent également être intégrés entre ces modules

principaux afin d'améliorer les résultats de détection, d'extraction de caractéristiques ou de classification.

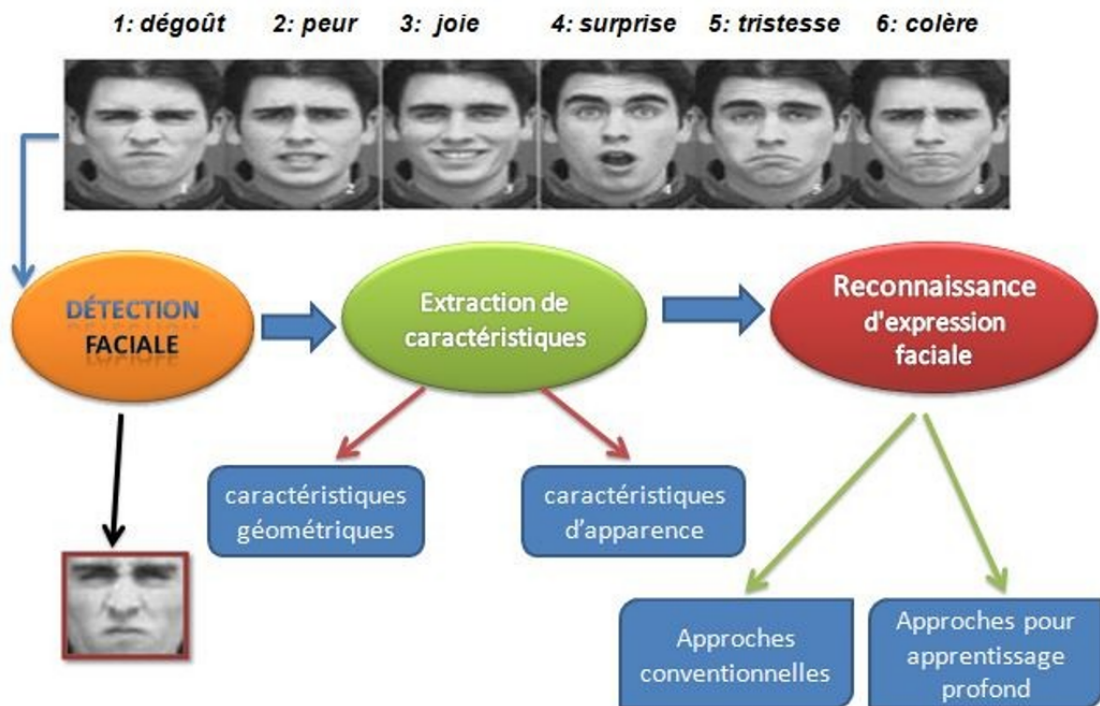


FIGURE 1.2 – Structure de base des systèmes de reconnaissance des émotions faciales.

### 1.4.1 Détection du visage

Le problème de détection des visages implique l'identification de la présence de visages dans une image et la détermination des emplacements et des échelles des visages. La précision de la détection du visage est particulièrement importante dans des conditions réalistes, où la présence du visage dans une scène et sa localisation globale ne sont pas connues a priori.



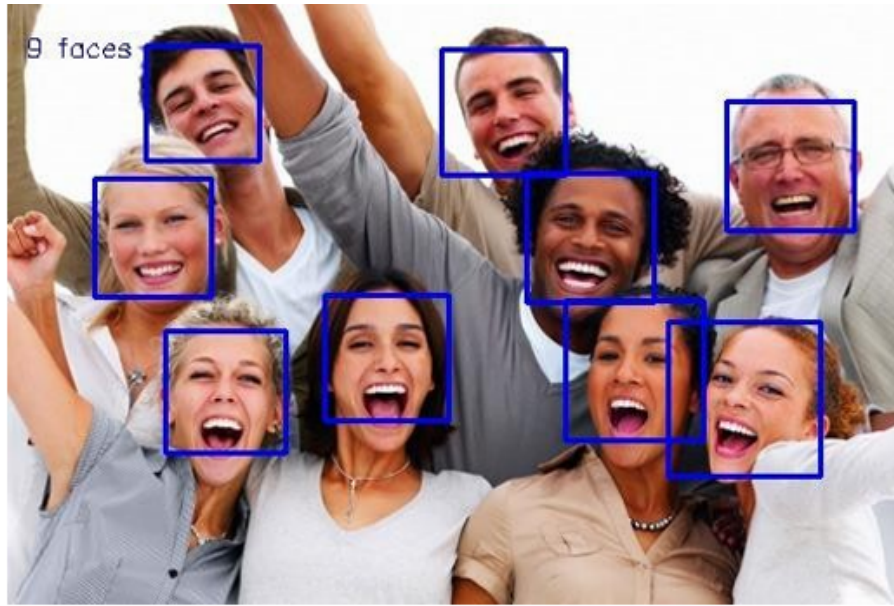


FIGURE 1.3 – Détection des visages.

### 1.4.2 Extraction des caractéristiques

Les points caractéristiques du visage sont principalement situés autour des composants faciaux tels que les yeux, la bouche, les sourcils, le nez et le menton. La détection des points caractéristiques du visage commence habituellement à partir d'une boîte englobante rectangulaire renvoyée par un détecteur de visage qui localise ce dernier. L'extraction de caractéristiques géométriques telles que les contours des composants faciaux, les distances faciales, etc. fournit les emplacements ou les caractéristiques d'apparence peuvent être calculées.

### 1.4.3 Reconnaissance de l'expression faciale

C'est la dernière étape d'un système de reconnaissance des expressions faciales. Cette phase cruciale vise à identifier et classifier l'ensemble des six expressions prototypes. Une fois que les caractéristiques faciales ont été extraites et représentées, le système utilise des techniques de comparaison pour évaluer la correspondance entre ces caractéristiques et les modèles de référence associés à chaque expression émotionnelle. La précision de cette reconnaissance finale est essentielle pour attribuer avec succès une étiquette émotionnelle à la personne analysée, enrichissant ainsi la compréhension globale des états émotionnels dans le contexte de l'interaction homme-machine.

## 1.5 Conclusion

Ce chapitre nous a initiés aux concepts clés de la reconnaissance des expressions faciales, abordant les émotions, et l'architecture des systèmes de reconnaissance. Dans le

chapitre suivant nous présenterons l'apprentissage profond en citant les différents types de réseaux de neurones et en détaillant les réseaux de neurones convolutifs .

# Chapitre 2

## Apprentissage profond

### 2.1 Introduction

Dans ce deuxième chapitre, nous plongerons dans l'univers de l'apprentissage profond, explorant ses concepts fondamentaux et ses différentes branches, notamment les réseaux de neurones artificiels et les réseaux de neurones convolutifs. Ce chapitre vise à établir une base solide pour comprendre les mécanismes sous-jacents à la reconnaissance des expressions faciales.

### 2.2 Deep learning

Le deep learning est une branche du machine learning (ML), qui est lui-même un sous-ensemble de l'intelligence artificielle (IA). Le concept d'IA existe depuis les années 1950, avec pour objectif de rendre les ordinateurs capables de penser et de raisonner d'une manière similaire à celle des humains. Pour rendre les machines capables de penser, le deep learning se concentre sur leur apprentissage sans une programmation explicite. Le deep learning va plus loin que le machine learning en créant des modèles hiérarchiques plus complexes destinés à imiter la façon dont les humains assimilent de nouvelles informations.

### 2.3 Réseaux de neurones artificiels

Les techniques et les calculs numériques le plus complexe permettent aux ordinateurs de faire le traitement durant un temps très court et on les comparant avec le cerveau humain, on notera que, théoriquement, l'ordinateur devrait être plus puissant que notre cerveau : S'il comprend 10<sup>9</sup> transistors alors le temps de commutation est de 10<sup>-9</sup> secondes. Le cerveau humain contient 10<sup>11</sup> neurones, mais ceux-ci ont seulement un temps de commutation d'environ 10<sup>-3</sup> secondes. La plus grande partie du cerveau fonctionne en permanence, par contre la plus grande partie de l'ordinateur est un espace de sauvegarde de données passives. Ainsi, l'ordinateur est statique, le cerveau comme un réseau neuronal

biologique peut se réorganiser au cours de sa durée de vie, et est donc en mesure d'apprendre, pour compenser les erreurs et ainsi de suite et c'est la capacité d'apprentissage qu'on ne trouve pas dans les ordinateurs. Le but de l'utilisation des réseaux de neurones est de s'inspirer de la biologie pour établir un système artificiel capable d'apprendre et de compenser ses erreurs donc c'est un modèle de calcul dont la conception est inspirée du fonctionnement des neurones biologiques. On trouve beaucoup de types de réseaux neurones, chaque type a pour un objectif particulier.

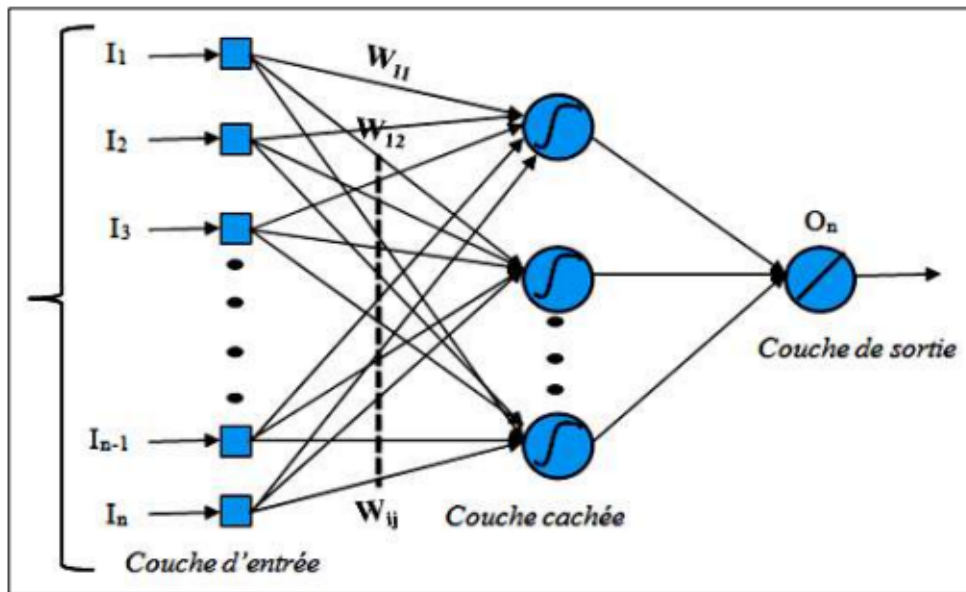


FIGURE 2.1 – Schéma d'un réseau de neurones artificiel

## 2.4 Réseaux de neurones convolutifs

Un réseau neuronal convolutif (CNN ou convnet) est un sous-ensemble de l'apprentissage automatique. Il s'agit de l'un des différents types de réseaux de neurones artificiels utilisés pour différentes applications et types de données. Un CNN est une sorte d'architecture réseau pour les algorithmes d'apprentissage en profondeur et est spécifiquement utilisé pour la reconnaissance d'images et les tâches impliquant le traitement des données de pixels.

Il existe d'autres types de réseaux de neurones dans l'apprentissage profond, mais pour identifier et reconnaître des objets, les CNN constituent l'architecture réseau de choix. Cela les rend parfaitement adaptés aux tâches de vision par ordinateur et aux applications où la reconnaissance d'objets est vitale, comme les voitures autonomes et la reconnaissance faciale. Les CNN sont construits en empilant plusieurs couches, dont les principales sont les suivantes :

- **Conv 2d** : La couche de convolution est la partie essentielle du CNN. Cette

couche est utilisée pour la détection de descripteurs ou motifs (*features*) à l'aide d'un ensemble de filtres composés de pondérations. Ces filtres sont 'glissés' sur la totalité de l'image dans le but de calculer les produits de convolution comme le montre la figure 2.2. Les pondérations des différents filtres sont déterminées lors de la phase d'apprentissage.

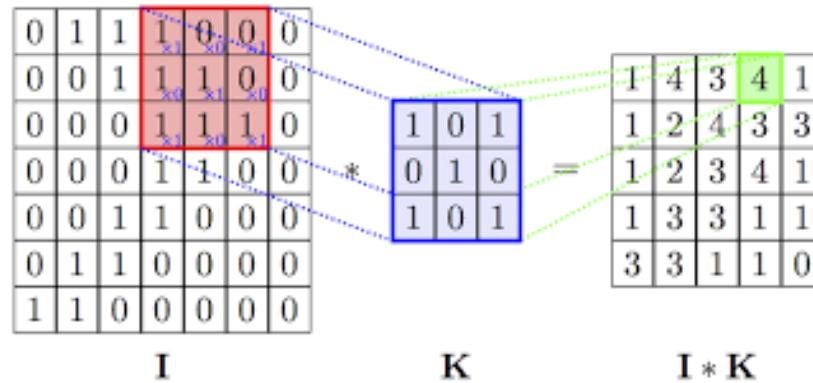


FIGURE 2.2 – Exemple d'une opération de convolution

Le rôle de cette première couche est d'analyser les images fournies en entrée et de détecter la présence d'un ensemble de *features*. En sortie de cette couche, on obtient un ensemble de *features maps*. Cette couche effectue la convolution des filtres sur l'image en entrée pour extraire des caractéristiques locales, telles que les bords, les textures, et les motifs.

- **Max pooling** : La couche Max-pool prend en entrée des listes de caractéristiques et, en fonction de la taille souhaitée, retourne une matrice de taille choisie. Pour ce faire, l'image est découpée.

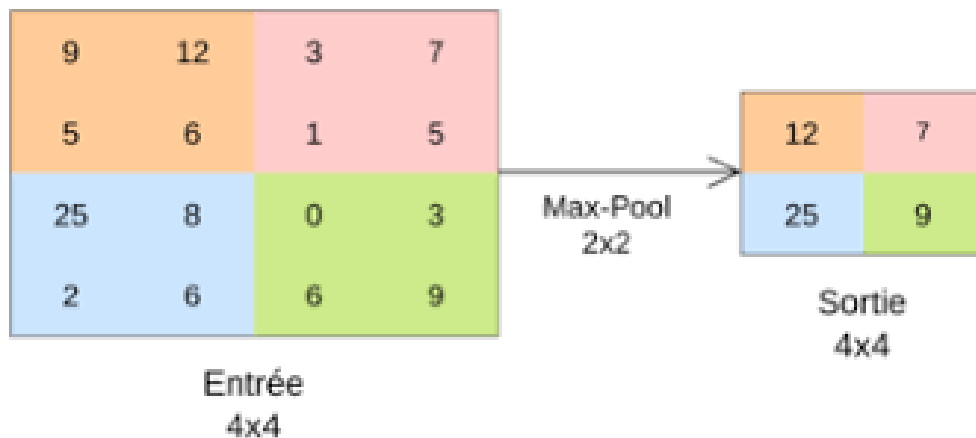


FIGURE 2.3 – Max pool 2 x 2

- **Fully connected** : Ces couches sont positionnées en fin d'architecture du CNN et sont entièrement connectées à tous les neurones de sortie, d'où le terme *fully-*

*connected*. Après avoir reçu un vecteur en entrée, la couche FC applique successivement une combinaison linéaire suivie d'une fonction d'activation dans le but final de classifier l'image d'entrée (voir le schéma suivant). En sortie, elle renvoie un vecteur de taille  $d$ , correspondant au nombre de classes, où chaque composante représente la probabilité pour l'image d'entrée d'appartenir à une classe.

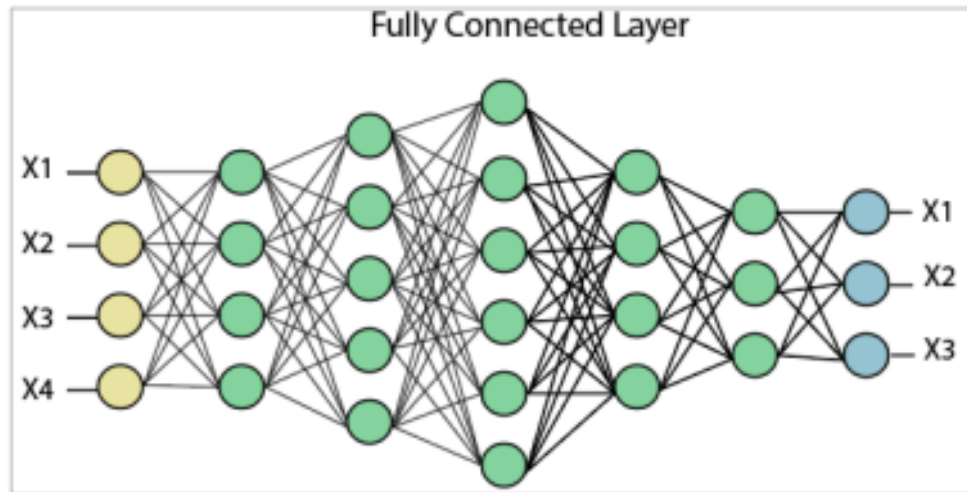


FIGURE 2.4 – Fully connected Layer

- **Relu** : La couche d'activation ReLU (Rectified Linear Units) : Cette couche remplace toutes les valeurs négatives reçues en entrée par des zéros. L'intérêt de ces couches d'activation est de rendre le modèle non linéaire, et par conséquent, plus complexe.

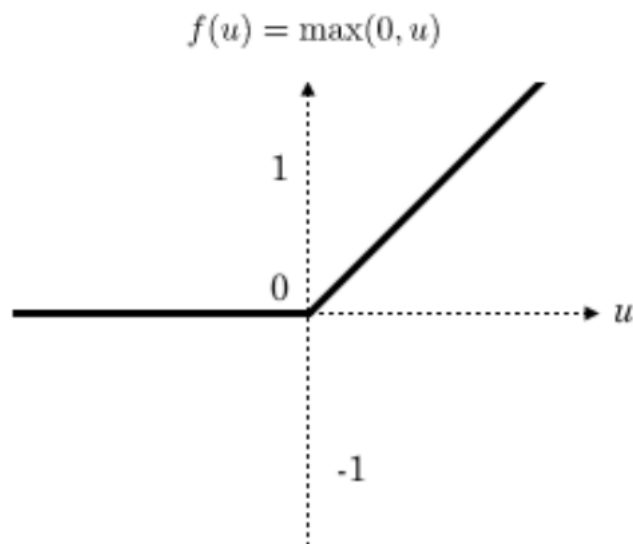


FIGURE 2.5 – Fonction d'activation ReLU

## 2.5 Evaluation

L'évaluation d'un réseau neuronal convolutif (CNN) se fait généralement en utilisant des mesures de performance spécifiques, telles que l'exactitude (accuracy), la perte (loss), la précision (precision), le rappel (recall), le F1-score, et la matrice de confusion.

- **Accuracy** : Il s'agit du pourcentage de prédictions correctes par rapport au nombre total d'échantillons.

$$\text{accuracy} = \frac{\text{Nombre de prédictions correctes}}{\text{Nombre total d'échantillons}} \quad (2.1)$$

- **Loss** : La perte (Loss) est généralement calculée à l'aide d'une fonction de perte spécifique, comme la perte d'entropie croisée (cross-entropy loss) pour la classification :

$$\text{Loss} = -\frac{1}{N} \sum_{i=1}^N (y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)) \quad (2.2)$$

$N$  est le nombre d'échantillons.

$y_i$  est la vraie étiquette de l'échantillon  $i$  (0 ou 1).

$\hat{y}_i$  est la prédiction du modèle pour l'échantillon  $i$  (probabilité de classe).

- **Precision** : Le nombre de vrais positifs divisé par la somme des vrais positifs et des faux positifs. Il mesure la précision des prédictions positives.

$$\text{Precision} = \frac{\text{Vrais positifs}}{\text{Vrais positifs} + \text{Faux positifs}} \quad (2.3)$$

- **Recall** : Le nombre de vrais positifs divisé par la somme des vrais positifs et des faux négatifs. Il mesure la capacité du modèle à identifier tous les exemples positifs.

$$\text{Recall} = \frac{\text{Vrais positifs}}{\text{Vrais positifs} + \text{Faux négatifs}} \quad (2.4)$$

- **F1-Score** : La moyenne harmonique de la précision et du rappel, ce qui donne une mesure équilibrée entre les deux.

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2.5)$$

- **Matrice de Confusion** : Est une table qui montre les vrais positifs (TP), les faux positifs (FP), les vrais négatifs (TN) et les faux négatifs (FN). Elle est souvent utilisée pour calculer les métriques de précision et de rappel. La matrice de confusion peut être représentée comme suit :

Confusion matrix		Reality	
		Negative : 0	Positive : 1
Prediction	Negative : 0	True Negative : TN	False Negative : FN
	Positive : 1	False Positive : FP	True Positive : TP

FIGURE 2.6 – Matrice de confusion

- **True Positives (TP)** : les cas dans lesquels nous avons prédit OUI et le résultat réel était également OUI.
- **True Negatives (TN)** : les cas dans lesquels nous avons prédit NON et le résultat réel était NON.
- **False Positives (FP)** : les cas dans lesquels nous avons prédit OUI et le résultat réel était NON.
- **False Negatives (FN)** : les cas dans lesquels nous avons prédit NON et le résultat réel était OUI.

## 2.6 Conclusion

Nous avons parcouru les fondements de l'apprentissage profond, en mettant l'accent sur les réseaux de neurones artificiels et les réseaux de neurones convolutifs. Ces concepts joueront un rôle crucial dans le développement de notre système de reconnaissance des expressions faciales. La prochaine étape consistera à appliquer ces connaissances dans le cadre spécifique de notre projet, ouvrant ainsi la voie à des expérimentations et à des analyses plus approfondies.



# Chapitre 3

## Reconnaissance des émotions faciales par réseaux convolutionnels

### 3.1 Introduction

Dans ce chapitre, nous parlons sur l'architecture retenue dans notre travail qui basée sur le réseaux de neurones convolutionnels. Quelle base de données on a exploitée ? et quelles sont les couches du réseau de neurones utilisée ? quelles sont les entrées et les sorties de réseau et quel langage de programmation nous avons choisi ?. Toutes les questions qui concernant le réseau CNN utilisé dans notre travail seront débattues en détail dans ce chapitre.

### 3.2 Architecture du réseau

Les réseaux des neurones convolutionnels sont devenu un sujet de recherche très actif dans le domaine de la vision par ordinateur. Leur usage s'est élargit à divers systèmes tels que la classification des images, la reconnaissance des objets et plus particulièrement dans la reconnaissance des émotions faciales qui est le sujet de notre fin d'étude.

Nous avons que l'architecture d'un système varie en fonction du champ d'application. A cette fin, nous avons proposé une architecture d'un réseau convolutionnel adapté à notre problème qui est de classer les sept émotions à partir d'une image.

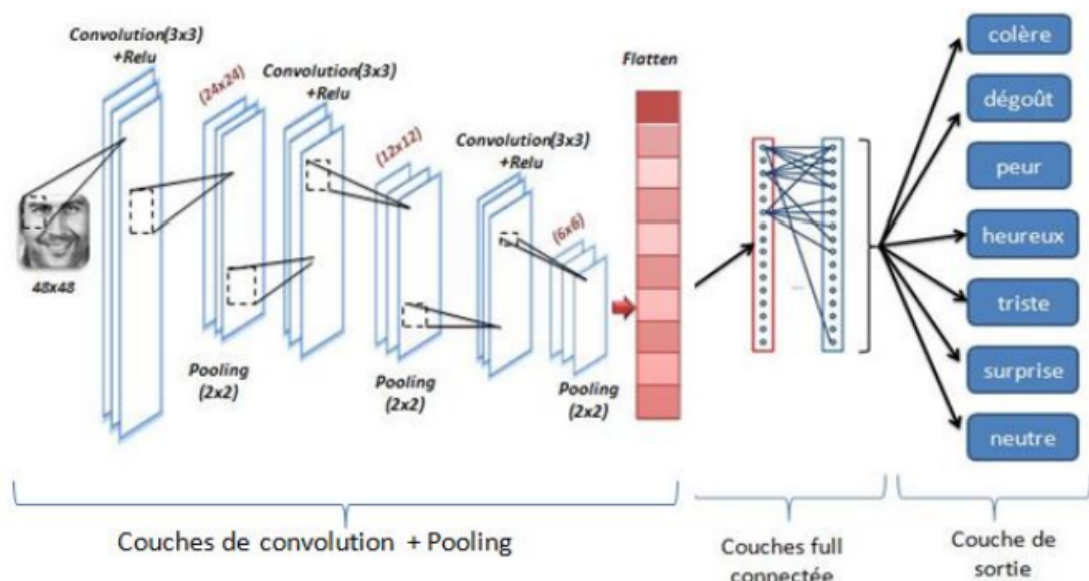


FIGURE 3.1 – Architecture du réseau utilisé.

### 3.3 Collecte des Données

Nous avons procédé à la collecte de données depuis la plateforme Kaggle, acquérant un ensemble de 35 887 images pour notre projet. Cette étape revêt une importance cruciale, car la qualité et la diversité des données influent directement sur les performances du modèle d'apprentissage automatique.

Pour garantir une évaluation pertinente et la robustesse du modèle, nous avons choisi de diviser l'ensemble de données selon un ratio de 80% pour l'entraînement, 10% pour les tests et 10% pour la validation. Cette stratégie permet d'optimiser le processus d'apprentissage en fournissant une quantité suffisante de données d'entraînement, tout en réservant des portions distinctes pour évaluer la généralisation du modèle et ajuster ses hyperparamètres.



FIGURE 3.2 – Exemples d'images de visage de l'ensemble de données

## 3.4 Prétraitement des Données

Le prétraitement des données revêt une importance cruciale dans le développement de notre modèle de détection des émotions faciales. Cette section se penchera sur les différentes étapes de prétraitement des données que nous avons effectuées pour garantir la qualité et la cohérence de notre ensemble de données.

- **Organisation des Données par Classes**

L'organisation des données par classes revêt une importance cruciale dans le développement de notre modèle de détection des émotions faciales, qui cible spécifiquement sept catégories émotionnelles distinctes. Ces classes comprennent la neutralité, le bonheur, la surprise, la tristesse, la colère, le dégoût et la peur. Chacune de ces classes représente une gamme unique d'expressions faciales, permettant ainsi au modèle d'apprendre à discerner avec précision les différentes émotions manifestées.

- **Redimensionnement des Images**

Toutes les images de notre ensemble de données ont été redimensionnées à une taille de 48x48 pixels.

- **Création de Lots (Batches)**

Nous avons créé des lots de 32 images à la fois pour l'entraînement de notre modèle. Cette approche par lots a amélioré l'efficacité du processus d'entraînement en permettant au modèle de s'entraîner sur chaque lot successivement.

- **augmentation des données**

L'augmentation des données est une étape essentielle dans le processus de prétraitement de nos données pour l'entraînement de notre modèle. Cette approche vise à diversifier les données d'entraînement en appliquant une gamme de transformations aux images d'origine. Son objectif fondamental consiste à renforcer la capacité de notre modèle à apprendre des caractéristiques plus robustes tout en améliorant sa capacité à généraliser de manière plus efficace aux données de l'entraînement, tout en minimisant le risque de surapprentissage.

Pour cette tâche, nous avons utilisé l'objet `ImageDataGenerator` de la bibliothèque Keras, en spécifiant plusieurs paramètres de transformation. Ces paramètres comprennent la rotation, le décalage en largeur et en hauteur, la normalisation, la distorsion, le zoom, la retournement horizontal, et la méthode de remplissage des pixels manquants.

## 3.5 Création du modèle CNN

Pour la création de notre modèle CNN, nous avons utilisé 14 couches de convolution 2D, ainsi que trois couches de max-pooling, avec un taux de dropout de 0.5 pour réguler le modèle en désactivant aléatoirement des neurones. Ensuite, nous avons configuré notre réseau neuronal artificiel (ANN) avec une couche cachée composée de 1024 neurones. Cette

couche cachée prend en entrée les sorties de la dernière couche de convolution du CNN.

<b>couche</b>	<b>fonction d'activation</b>	<b>valeur</b>
conv2d_1	Relu	filtres=64 size_filtres=(3,3) img_input=(48, 48)
conv2d_2	Relu	filtres=64 size_filtres=(3,3)
dropout	None	dropout = 0.5
conv2d_3	Relu	filtres=64 size_filtres=(3,3)
conv2d_4	Relu	filtres=64 size_filtres=(3,3)
max_pooling2d	None	size = (2,2)
conv2d_5		filtres=128 size_filtres=(3,3)
..... .....	..... .....	..... .....
flatten (Flatten)		unit = 1024
Dense	softmax	unit = 7

FIGURE 3.3 – l'architecture CNN utilisée.

## 3.6 Résultats

Les hyperparamètres choisis pour notre modèle sont un batch de taille 32, 100 époques, et un taux d'apprentissage de 0.001, avec une taille d'image fixée à (63,52) pixels. Cette partie se concentrera sur la comparaison des résultats obtenus avec ces hyperparamètres. Bien qu'il aurait été intéressant de les varier et de chercher à les optimiser pour explorer les performances potentielles de notre modèle, notre principal objectif était de réussir à le faire fonctionner et de comprendre son fonctionnement.

### -Discussion

pour fig 3.4 :

Les valeurs d'accuracy et de validation accuracy montrent une amélioration progressive du modèle au fil de l'entraînement. Initialement, l'accuracy était relativement basse, mais

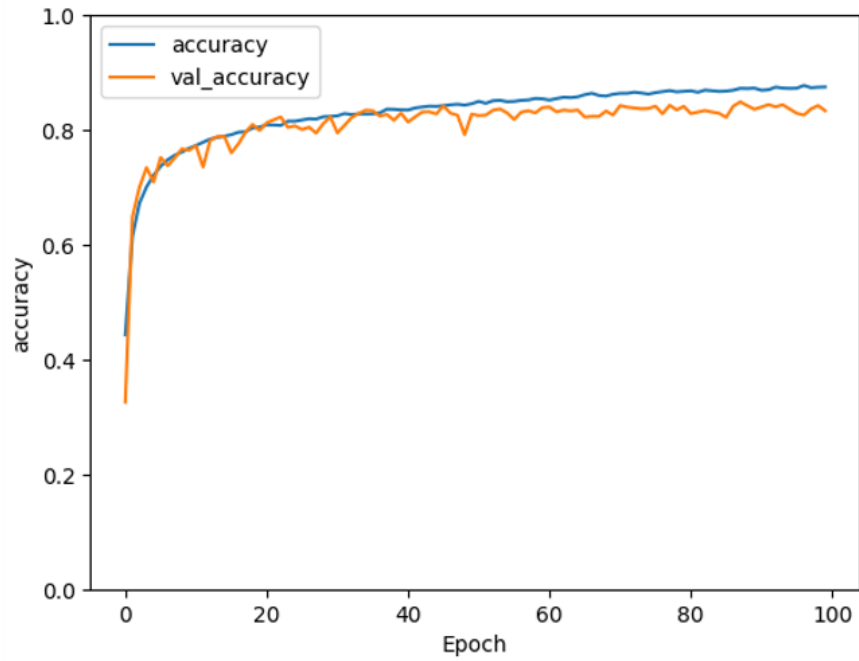


FIGURE 3.4 – Évolution Accuracy et Accuracy-validation sur l'apprentissage.

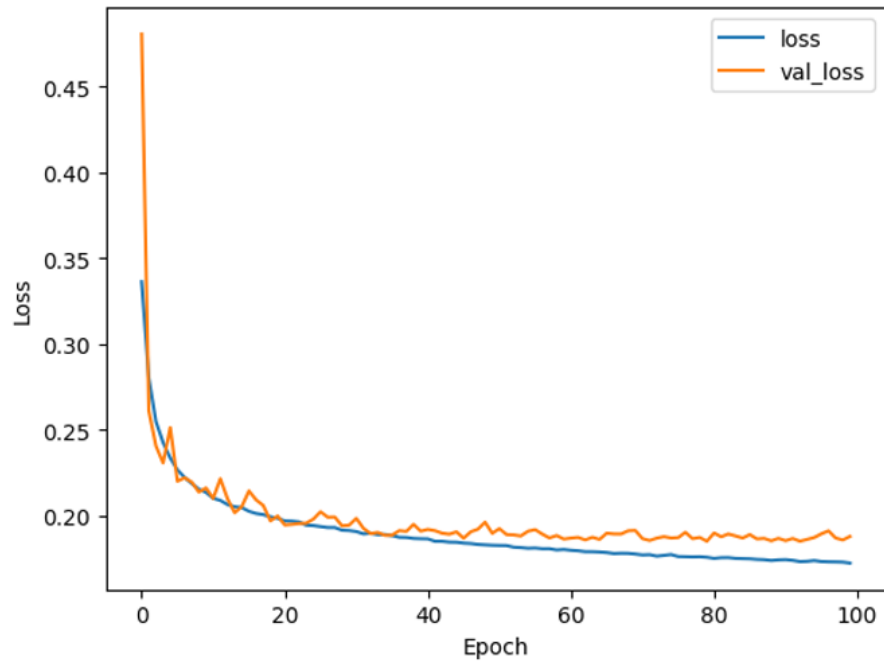


FIGURE 3.5 – Évolution loss et loss-validation sur l'apprentissage.

elle a augmenté au fur et à mesure des époques, atteignant finalement un niveau satisfaisant. De même, la validation accuracy a montré une tendance à l'augmentation, indiquant que le modèle est capable de bien généraliser les données. Cependant, il est à noter que ces valeurs semblent rester relativement constantes après un certain nombre d'époques, suggérant que le modèle a atteint un niveau de performance stable.

pour fig 3.5 :

Au départ, les valeurs de loss étaient élevées, mais elles ont rapidement convergé vers des niveaux plus bas, indiquant un apprentissage efficace. Nous avons également constaté que le modèle a bien généralisé, évitant ainsi le surajustement. Cependant, il est important de noter que les courbes semblent rester constantes après un certain nombre d'époques, ce qui suggère qu'ajouter davantage d'époques d'entraînement pourrait ne pas améliorer considérablement la performance du modèle.

## 3.7 Évaluation des Performances

Pour évaluer notre modèle, nous avons utilisé un jeu de données composé de 35 887 images. Il convient de noter que 10% jeu de données a été utilisé à des fins de validation, et de même pour le test . Après l'exécution du modèle, nous avons obtenu les résultats suivants :

Classification Report:				
	precision	recall	f1-score	support
neutre	0.77	0.87	0.82	1357
bonheur	0.94	0.90	0.92	899
surprise	0.87	0.81	0.84	422
tristesse	0.72	0.60	0.65	449
colère	0.77	0.76	0.77	326
dégoût	0.82	0.30	0.44	30
peur	0.64	0.53	0.58	66
accuracy			0.82	3549
macro avg	0.79	0.68	0.72	3549
weighted avg	0.82	0.82	0.81	3549

FIGURE 3.6 – Evaluation

En analysant la matrice de confusion de notre modèle, nous pouvons mieux comprendre comment il se comporte lors de la classification des images. Les valeurs obtenues se répartissent en 7 catégories clés :

les True Positives, les True Negatives, les False Positives et les False Negatives. Ces résultats nous permettent d'évaluer la performance du modèle en termes de détection des Expliquons ces termes en détail :

True Positives (TP) : Il s'agit du nombre d'images qui étaient réellement de la classe positive (par exemple, des vraies images) et que le modèle a correctement classées comme telles. En d'autres termes, le modèle a correctement détecté ces images comme étant positives.

True Negatives (TN) : Il s'agit du nombre d'images qui étaient réellement de la classe négative (par exemple, des fausses images) et que le modèle a correctement classées comme telles. Le modèle a bien fait de ne pas détecter ces images comme étant positives.

False Positives (FP) : Il s'agit du nombre d'images qui étaient réellement de la classe négative (fausses images) mais que le modèle a incorrectement classées comme positives. Cela signifie que le modèle a fait une erreur en identifiant ces images comme étant positives alors qu'elles ne l'étaient pas.

False Negatives (FN) : Il s'agit du nombre d'images qui étaient réellement de la classe positive (vraies images) mais que le modèle a incorrectement classées comme négatives. Cela signifie que le modèle a fait une erreur en ne parvenant pas à détecter ces images comme étant positives.

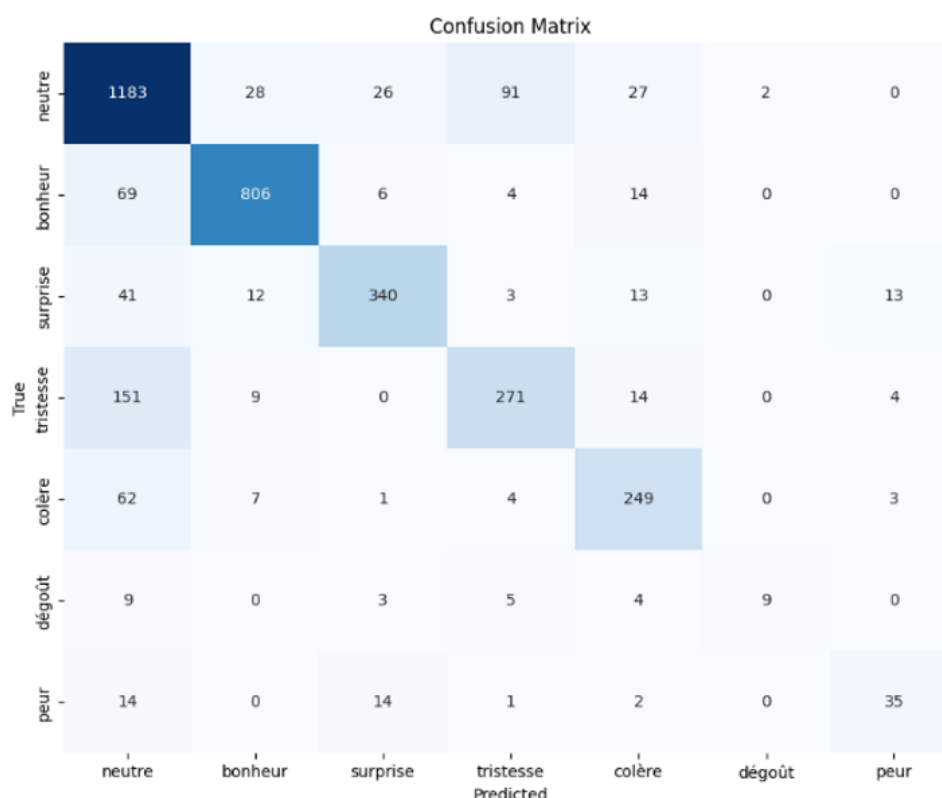


FIGURE 3.7 – matrice de confusion

## 3.8 Conclusion

En conclusion, ce chapitre nous a permis d'explorer en profondeur les résultats obtenus dans notre étude sur la détection des émotions faciales. Notre travail de collecte et de prétraitement des données a jeté les bases essentielles pour l'analyse ultérieure, tandis que la création de notre modèle CNN a donné lieu à des résultats prometteurs.

L'évaluation des performances de notre modèle a révélé son efficacité dans la détection des émotions faciales, renforçant ainsi la sécurité des systèmes de reconnaissance faciale. De plus, en comparant notre approche aux méthodes existantes, nous avons pu démontrer son avantage compétitif dans la prévention des attaques basées sur la contrefaçon faciale.

# Conclusion générale

Dans ce projet, nous avons exploré les bases de l'apprentissage automatique, mettant l'accent sur les méthodes avancées telles que l'apprentissage profond. Plus spécifiquement, nous avons mis en lumière le rôle crucial des réseaux de neurones convolutionnels dans le traitement des images, en examinant de près leurs couches constitutives, à savoir les couches de convolution, de pooling et les couches entièrement connectées.

L'objectif central de notre démarche est la conception et l'implémentation d'une application de reconnaissance des émotions. En tirant parti des avancées en matière d'apprentissage profond, nous cherchons à élaborer un système robuste capable de déchiffrer et d'interpréter les expressions faciales, contribuant ainsi à l'évolution des technologies de compréhension des émotions.

À travers cette exploration, nous avons jeté les bases nécessaires pour la création d'une application qui va au-delà de la simple identification d'images, en se plongeant dans le domaine complexe de la reconnaissance des émotions humaines. Les défis rencontrés au cours de ce processus ont enrichi notre compréhension des nuances associées à cette tâche et nous ont incités à exploiter pleinement les capacités des réseaux de neurones convolutionnels.