



République Algérienne Démocratique Et Populaire

Ministère De L'enseignement Supérieure Et De La Recherche Scientifique

Université 20 Août 1955 –Skikda-

Faculté des sciences  
Département de l'informatique  
Intelligence Artificielle

## **Mémoire de fin d'études en vue de l'obtention du diplôme de Master en Informatique**

**THÈME**

**Systeme de détection faciale des émotions avec  
CNN**

**Présenté par :**

KRAIM Radji.

BEN HADDAD Aghiles.

BOUACHA Rami Ahmed Mokhtar.

**Encadré par :**

Dr. Mazouzi Smaine.

Lala Bouali Yasmine.

**Juillet 2024**

---

# Remerciements

Premièrement, nous remercions Dieu tout puissant pour tous les bienfaits qu'il nous a accordés et pour nous avoir donné la force et la volonté nécessaires pour accomplir ce travail.

Nous remercions par la suite, nos encadrants, Docteur smaine Mazouzi Et Lala Bouali Yasmine pour la liberté qu'ils nous ont donné et pour leur précieux conseils.

Nous exprimons notre respect aux membres du jury qui ont accepté d'y participer et de juger notre travail.

Nous remercions également tous les enseignants qui ont contribué à notre instruction durant nos années d'étude.

Sans oublier tous ceux qui nous ont aidés de près ou de loin à l'élaboration de ce modeste travail.

---

## **Dédicaces**

Avec tout l'amour que nous portons dans nos cœurs, nous dédions ce mémoire de fin d'études universitaires à nos deux êtres respectifs qui nous sont les plus chers dans ce monde : Nos parents respectifs. Nous dédions ce modeste travail aussi à nos familles respectives, nos amis ainsi qu'à tous nos enseignants.

---

# Sommaire

Table des matières	
Liste de figures .....	1
Liste des tableaux.....	3
Résumé.....	4
Introduction Générale.....	7
Introduction générale .....	8
Chapitre I :.....	<b>10</b>
Les Emotions .....	10
1 Introduction .....	11
2 Emotion et expression faciale .....	11
2.1 Les émotions .....	11
2.1.1 Définition.....	11
2.1.2 Types d'émotions .....	11
a. Emotion primaires.....	12
b. Emotion secondaire.....	13
c. Emotion sociale.....	13
2.2 Les expressions faciales .....	14
2.2.1 Définition.....	14
2.2.2 Description des six expressions faciales.....	14
2.2.3 Reconnaissance des expressions faciales .....	15
2.3 Différence entre expression faciales et émotions .....	16
3 Détection des Emotion .....	17
3.1 Méthodes de détection des émotions.....	17
3.1.1 Méthodes directes : expression faciale.....	17
3.1.2 Méthodes indirectes.....	18
3.1.3 Analyse comparative des méthodes directes et indirectes.....	18
3.2 Difficultés et défis de détection des émotions .....	19
3.3 Intérêts de détections des émotions .....	19
4 Conclusion.....	20

---

Chapitre II : .....	21
Apprentissage Automatique et Emotions .....	21
1 Introduction : .....	22
2 Partie 1 : Apprentissage Automatique (Machine Learning) : .....	22
2.1 Définition.....	22
2.2 Méthodes d'apprentissage automatiques.....	22
2.2.1 Méthodes supervisées et non supervisées.....	22
a. Apprentissage non-supervisé (ou classification automatique) : .....	22
b. Apprentissage supervisé.....	23
2.2.2 Méthodes classiques en ML .....	24
2.2.3 Méthodes modernes en ML.....	25
2.2.4 Méthodes ensemblistes.....	25
2.2.5 Méthodes profondes .....	25
2.2.6 Méthodes par renforcement.....	26
3 Partie 2 : Systèmes de détection d'émotion .....	27
3.1 Taxonomie des systèmes .....	27
3.1.1 Système à base d'apprentissage automatique.....	27
3.1.2 Systèmes à base d'apprentissage profond .....	28
3.1.3 Systèmes basé apprentissage par renforcement.....	29
4 Conclusion.....	30
Chapitre III : .....	31
Conception d'un système de détection d'émotion basé sur CNNs.....	31
1. Introduction.....	32
2. Aperçu du système .....	32
3. Conception .....	33
4. Diagrammes UML .....	33
4.1 Qu'est-ce qu'un diagramme UML ?.....	33
4.2 Diagramme de cas d'utilisation.....	33
4.3 Diagramme de séquence.....	37
4.4 Diagramme de classes .....	39
5. Modèle ML proposé .....	44
6. Présentation du modèle CNN.....	46

---

6.1	Différents modules d'un réseau de neurones convolutif.....	47
6.1.1	La convolution.....	48
6.1.2	Les différentes convolutions.....	49
6.1.3	Le pooling.....	50
6.1.4	Les fonctions d'activations.....	52
6.1.5	Couche entièrement connectée (Fully Conneceted Layer (FC)).....	53
6.2	Outils d'optimisation des réseaux convolutifs .....	54
6.2.1	Le batch normalisation .....	54
6.2.2	Les fonctions de perte.....	55
6.2.3	Méthodes de régularisation.....	56
6.3	Les architectures neuronales convolutifs.....	57
7.	Conclusion.....	59
Chapitre IV :.....		<b>60</b>
Développement et tests.....		60
1	Introduction .....	61
2	Implémentation de système :.....	61
3	La comparaison.....	80
4	Discussion.....	81
5	Le prototype .....	81
6	Conclusion.....	88
Conclusion Générale.....		89
Conclusion générale .....		90
Bibliographie.....		91

---

## Liste de figures

Figure 1: Les six expressions émotionnelles universelles dans l'ordre suivant : Bonheur, Peur, Dégoût, colère, tristesse et surprise [W2].	12
Figure 2: Les émotions et leurs expressions faciales [W4].	15
Figure 3 : Un mini architecture pour le processus de reconnaissance des émotions faciales.	15
Figure 4: Expression faciale de la joie [W6].	17
Figure 5: Modèle d'Apprentissage supervisé.	24
Figure 6 : Apprentissage par renforcement.	30
Figure 7 : Diagramme cas d'utilisation.	34
Figure 8 : Diagramme de séquence.	37
Figure 9 : Diagramme de classe.	40
Figure 10 : Architecture générale d'un CNN.	48
Figure 11: Illustration d'une opération de convolution entre deux couches.	48
Figure 12 : Les différentes couches d'un réseau de neurones convolutif standard.	50
Figure 13 : Illustration du Max Pooling.	50
Figure 14 : Les différents types de Pooling avec un filtre 2x2 et un pas de 2.	51
Figure 15 : Fonction d'activation de RELU.	52
Figure 16 : Illustration du fonctionnement d'une couche ReLU. (Dans la case de gauche, tous les nombres négatifs ont été convertis en zéro après l'application de la fonction d'activation, tandis que toutes les autres valeurs sont restées inchangées).	53
Figure 17 : Architecture standard d'un réseau convolutifs.	54
Figure 18 : La fonction de perte Softmax.	55
Figure 19 : Fonctions d'activation SoftMax.	56
Figure 20 : LeNet et AlexNet Architectures.	58
Figure 21 : Logo de python.	62
Figure 22 : Logo de Google Colab.	62
Figure 23: Logo Spyder.	63
Figure 24 : Répartition de la base de données FER2013 par émotion.	65
Figure 25 : Data set FER2013.	66
Figure 26 : code python.	67
Figure 27 : Architectures de modèle séquentiel.	68

Figure 28 : Les courbes de précision et de perte pour le modèle Séquentiel.....	69
Figure 29 : Matrice de confusion pour le modèle Séquentiel. ....	70
Figure 30 : code python.....	70
Figure 31: code python .....	71
Figure 32 : Architecture de Fer-modèle.....	71
Figure 33 : Les courbes de précision et de perte pour le FER-Modèle.....	72
Figure 34 : Matrice de confusion pour le FER-Modèle.....	73
Figure 35: Architecteur Global de Modèle Res-Net [W44]. ....	74
Figure 36 : Les courbes de précision et de perte pour le modèle RES-NET. ....	75
Figure 37 : Matrice de confusion pour le modèle RES-NET.....	75
Figure 38 : Les composants de L'architecture VGG Modèle.....	76
Figure 39 : Les courbes de précision et de perte pour le modèle VGG. ....	77
Figure 40 : Matrice de confusion pour le modèle VGG.....	77
Figure 41 : The accuracy and loss curves for the FER model with SVM. ....	78
Figure 42: Matrice de confusion pour le modèle FER CNN avec SVM.....	79
Figure 43 : Page d'accueil.....	81
Figure 44 : interface offline. ....	82
Figure 45 : Sélectionner l'émotion.....	83
Figure 46 : interface online.....	84
Figure 47 : interface sélection cameras.....	84
Figure 48 : interface sélectionner l'émotion.....	85
Figure 49 : What to detect. ....	85
Figure 50 : paramètres des alertes. ....	86

## Liste des tableaux

Tableau 1 : Tableau comparatives des méthodes. ....	18
Tableau 2 : FER2013 Dataset par émotion.....	65
Tableau 3: FER2013 dataset. ....	66
Tableau 4 : les paramètres et ses utilisations. ....	68
Tableau 5 : Taille des paramètres du modèle Séquentiel. ....	68
Tableau 6 : Résultats de Loss, Accuracy, Val_Loss et Val_Accuracy pour le modèle Séquentiel. ....	69
Tableau 7 : Taille des paramètres du modèle FER.....	72
Tableau 8 :Résultats de Loss, Accuracy, Val_Loss et Val_Accuracy pour le modèle FER .....	72
Tableau 9 : Taille des paramètres du modèle Res-Net.....	74
Tableau 10 : Résultats de Loss, Accuracy, Val_Loss et Val_Accuracy pour le modèle Res-Net. ....	75
Tableau 11 : Taille des paramètres du modèle VGG.....	77
Tableau 12 : Résultats de Loss, Accuracy, Val_Loss et Val_Accuracy pour le modèle VGG.....	77
Tableau 13 : Résultats de Loss, Accuracy, Val_Loss et Val_Accuracy pour le modèle CNNSVM.....	78
Tableau 14 : Résultats de précision, de rappel et de score F1 pour le modèle FER avec SVM. ....	79
Tableau 15 : La différence de précision et les résultats de perte des modèles proposés. ....	80

## Résumé

Au cours des dernières décennies, la reconnaissance des émotions et l'informatique affective ont suscité un vif intérêt dans divers domaines de recherche. Les expressions faciales sont reconnues comme un moyen efficace pour identifier les caractéristiques distinctives du comportement humain et décrire les états émotionnels avec précision.

Dans ce travail de Master, nous nous concentrons sur la détection des émotions humaines à partir de visages. Nous utilisons des réseaux de neurones à convolution (CNN) pour analyser le dataset FER2013 et identifier des émotions comme le bonheur, la tristesse, la colère, la surprise, le dégoût, la neutralité et la peur. Nos expérimentations sont menées sur les plateformes Spyder et Google Colab.

Les CNN sont des architectures spécialisées dans la reconnaissance de formes visuelles. En plus de l'analyse théorique, nous avons développé un modèle innovant de détection d'émotions basé sur les CNN. Notre étude inclut une comparaison détaillée entre différentes architectures comme ResNet, VGG, FER-Model, SVM-CNN et Sequential-Model pour leur capacité à apprendre et reconnaître les émotions.

# Abstract

Over the past few decades, emotion recognition and affective computing have garnered significant interest across various research domains. Facial expressions are particularly recognized as an effective means to identify distinctive features of human behavior and accurately describe emotional states.

In this Master's thesis, our focus lies on detecting human emotions from faces. To achieve this, we employ Convolutional Neural Networks (CNNs) to analyze the FER2013 dataset and identify emotions such as happiness, sadness, anger, surprise, disgust, neutrality, and fear. Our experiments are conducted using Spyder and Google Colab platforms.

CNNs are multilayered neural networks specialized in pattern recognition. Beyond theoretical analysis, we have developed an innovative emotion detection model based on CNNs. Our study includes a comprehensive comparison of different architectures like ResNet, VGG, FER-Model, SVM-CNN, and Sequential-Model in learning and recognizing emotions.

على مر العقود القليلة الماضية، لفتت التعرف على العواطف والحوسبة العاطفية اهتمامًا كبيرًا عبر مختلف مجالات البحث. تُعتبر التعبيرات الوجهية بشكل خاص وسيلة فعالة لتحديد السمات المميزة لسلوك الإنسان ووصف الحالات العاطفية بدقة.

في رسالة الماجستير هذه، يكمن تركيزنا في كشف العواطف البشرية من الوجوه. لتحقيق ذلك، نستخدم شبكات الأعصاب التكرارية (CNNs) لتحليل مجموعة بيانات FER2013 وتحديد العواطف مثل السعادة، الحزن، الغضب، المفاجأة، الاشمئزاز، اللامبالاة، والخوف. تتم تجربتنا باستخدام منصتي Spyder و Google Colab.

تعتبر شبكات الأعصاب التكرارية (CNNs) شبكات عصبية متعددة الطبقات متخصصة في التعرف على الأنماط. بالإضافة إلى التحليل النظري، قمنا بتطوير نموذج مبتكر للكشف عن العواطف استنادًا إلى شبكات الأعصاب التكرارية. تتضمن دراستنا مقارنة شاملة بين هياكل مختلفة مثل ResNet ، VGG ، FER-Model ، SVM-CNN ، و Sequential-Model في تعلم والتعرف على العواطف.

# **Introduction**

## **Générale**

## **Introduction générale**

Dans un monde de plus en plus connecté et technologiquement avancé, les interactions entre les humains et les machines revêtent une importance croissante. Les systèmes intelligents capables de comprendre et de répondre aux émotions humaines jouent un rôle clé dans cette évolution. L'émotion est un aspect fondamental de notre expérience quotidienne, influençant nos décisions, nos interactions sociales et notre bien-être général. L'intégration de la reconnaissance des émotions dans les interfaces utilisateur représente ainsi un domaine de recherche et de développement stratégique, visant à créer des expériences utilisateur plus intuitives, personnalisées et efficaces.

La reconnaissance des émotions à partir des expressions faciales constitue l'une des approches les plus directes et accessibles pour comprendre les états affectifs des individus. Les expressions faciales sont des manifestations visibles et souvent universelles des émotions humaines, rendant la détection automatisée des émotions à partir des images faciales une tâche prometteuse et pertinente. Cette capacité technologique trouve des applications diversifiées, allant des soins de santé aux interfaces utilisateur adaptatives, en passant par l'éducation et les loisirs interactifs.

Au cours du développement de ce travail, des algorithmes d'apprentissage automatique et des techniques d'apprentissage en profondeur ont été utilisés sur des images montrant les émotions faciales suivantes : bonheur, tristesse, colère, surprise, dégoût, neutralité et peur pour extraire des caractéristiques sémantiques à partir d'images faciales.

L'objectif principal de ce mémoire est d'explorer et de concevoir un système avancé de détection d'émotions basé sur les réseaux de neurones convolutifs (CNNs). Ce système sera conçu pour analyser en temps réel les expressions faciales des utilisateurs et identifier les émotions sous-jacentes avec une précision et une fiabilité accrue. L'application de cette technologie vise à enrichir les interactions homme-machine en permettant aux systèmes de s'adapter et de répondre de manière empathique aux états émotionnels des utilisateurs.

Pour atteindre cet objectif, cette recherche est structurée en quatre chapitres, organisés comme suit :

Le premier chapitre établit les bases théoriques en définissant les émotions humaines, en explorant les théories de la classification émotionnelle et en analysant les expressions faciales

associées à chaque état affectif. Nous examinerons également les aspects psychologiques et neurologiques sous-jacents à la manifestation des émotions à travers les expressions faciales.

Le second chapitre se concentrera sur les méthodes et les techniques d'apprentissage automatique appliquées à la reconnaissance des émotions à partir des images faciales. Nous explorerons les approches traditionnelles ainsi que les récentes avancées en matière de réseaux de neurones convolutifs (CNNs), en mettant l'accent sur leurs capacités à extraire des caractéristiques pertinentes des données visuelles complexes.

Le troisième chapitre décrira la conception détaillée de notre système proposé, en incluant l'architecture spécifique des CNNs choisie, les techniques de prétraitement des images, et les stratégies d'évaluation de la performance du système. Nous aborderons également les considérations de mise en œuvre et les défis techniques rencontrés lors du développement du système.

Quant au dernier chapitre, il est consacré au développement concret du système de reconnaissance des émotions basé sur les CNNs. Nous discutons des étapes de développement, dévoilant les résultats obtenus et orchestrant une comparaison éclairante entre ces diverses architectures de CNN dédiées à l'exploration des émotions. Nous présentons après notre interface du programme et nous avec l'explication de chaque page et son fonctionnement.

Enfin, on termine par une conclusion générale et quelques perspectives.

# **Chapitre I :**

# **Les Emotions**

# **1 Introduction**

Les émotions sont essentielles à notre vie quotidienne, influençant nos comportements et interactions sociales. Ce chapitre explore les différents types d'émotions, leur définition et leur classification. Nous examinons comment les émotions se manifestent à travers les expressions faciales et discutons des méthodes de détection des émotions, en soulignant les défis et opportunités. Cette analyse fournira une base solide pour comprendre les techniques de reconnaissance des émotions et leur application dans les technologies modernes.

## **2 Emotion et expression faciale**

### **2.1 Les émotions**

#### **2.1.1 Définition**

Les émotions représentent un état subjectif que nous associons souvent à nos sentiments. En général, le terme émotion fait référence à un état subjectif et affectif relativement intense qui émerge en réponse à une expérience vécue. On a tendance à penser que les émotions sont exprimées de manière délibérée et intentionnelle. Elles sont souvent le moteur de la motivation, qu'elle soit positive ou négative, et se manifestent à travers une variété de comportements, tels que la modulation de la voix, les expressions faciales et le langage corporel [W1].

#### **2.1.2 Types d'émotions**

Il existe différents types d'émotions qui ponctuent nos journées et influencent nos humeurs. Ces catégories comprennent les émotions primaires, les émotions secondaires et les émotions sociales. Les émotions primaires, ou de base, sont instinctives et universelles, tandis que les émotions secondaires sont souvent plus complexes et peuvent découler des primaires. Les émotions sociales sont influencées par notre interaction avec les autres et notre environnement social.

### a. Emotion primaires

Les émotions primaires sont considérées comme universelles et innées, et on pense qu'elles sont ressenties par les personnes de toutes cultures. Ces émotions sont considérées comme les plus fondamentales et sont généralement exprimées à travers des expressions faciales. Les principales émotions primaires sont :

**Joie** : Cette émotion se caractérise par des sentiments de bonheur, de contentement et de satisfaction. Elle est souvent exprimée par un sourire et des expressions faciales lumineuses et ouvertes.

**Tristesse** : La tristesse implique des sentiments de chagrin, de déception ou de peine. Elle est généralement exprimée par une bouche baissée, des yeux larmoyants et une posture affaissée.

**Peur** : La peur est une réponse à une menace ou un danger perçu. Elle est souvent caractérisée par des yeux écarquillés, une posture tendue et un désir d'échapper ou d'éviter la source de la peur. **Colère** : La colère implique des sentiments de frustration, d'irritation ou d'indignation. Elle est souvent exprimée par des sourcils froncés, une mâchoire serrée et une voix forte.

**Surprise** : La surprise survient en réponse à des événements ou stimuli inattendus. Elle est exprimée par des yeux écarquillés, une bouche ouverte et des sourcils levés.

**Dégoût** : Le dégoût est une réaction à des choses considérées comme sales, offensantes ou répugnantes. Il est généralement exprimé par un nez retroussé, une lèvre supérieure relevée et un détournement de la source du dégoût [1].



Figure 1: Les six expressions émotionnelles universelles dans l'ordre suivant : Bonheur, Peur, Dégoût, colère, tristesse et surprise [W2].

## **b. Émotion secondaire**

Les émotions secondaires sont des réponses émotionnelles qui se produisent en réaction à d'autres émotions, souvent appelées émotions primaires. Elles sont généralement plus complexes et peuvent nécessiter une exploration approfondie pour les comprendre pleinement.

Les émotions secondaires sont construites à partir de mélanges d'émotions primaires et de cognitions et d'appréhensions particulières. Par exemple, la honte peut être une émotion secondaire qui réunit de la peur et de la colère (envers soi-même). D'autres émotions secondaires peuvent inclure l'amour, la haine, la méfiance, la culpabilité, etc [W3].

Selon le psychologue Paul Ekman, les émotions secondaires pourraient inclure l'amusement, la satisfaction, la gêne, l'excitation, la culpabilité, la fierté dans la réussite, le soulagement, le plaisir sensoriel, la honte et le mépris [2].

## **c. Émotion sociale**

Les émotions sociales jouent un rôle crucial dans le façonnement du comportement humain et des interactions sociales [3].

Elles peuvent être catégorisées en émotions sociales douloureuses telles que la culpabilité, la timidité, la peur sociale, et en émotions sociales agréables comme l'amour, l'amitié, l'empathie, la solidarité, et les sentiments moraux [4]. Ces émotions influencent la conduite sociale et ont un impact significatif sur les interactions individuelles au sein de la société [5].

Les émotions sociales sont étroitement liées au développement des compétences sociales et ont un impact significatif sur les interactions individuelles avec autrui [6].

## **2.2 Les expressions faciales**

### **2.2.1 Définition**

Une expression faciale est un ou plusieurs mouvements ou positions des muscles sous la peau du visage, qui transmettent l'état émotionnel d'un individu aux observateurs, formant ainsi une forme de communication non verbale [7].

### **2.2.2 Description des six expressions faciales**

Voici les six expressions faciales de base décrites :

**La joie** : La joie se manifeste par un sourire avec les coins des lèvres levés, souvent accompagné de plis autour des yeux, appelés communément "pattes d'oie".

**La colère** : La colère se caractérise par un froncement des sourcils, des yeux largement ouverts, une bouche serrée ou tordue, parfois avec des dents serrées ou exposées.

**La tristesse** : La tristesse se traduit par une bouche tombante, des coins des lèvres baissés, des paupières tombantes et des yeux tristes avec des ridules sous les yeux.

**La peur** : La peur se manifeste par des sourcils levés, des yeux écarquillés avec les pupilles dilatées, des lèvres étirées vers l'extérieur ou vers le bas.

**La surprise** : La surprise se caractérise par des sourcils levés, des yeux écarquillés, souvent avec des paupières supérieures soulevées, et une bouche légèrement ouverte.

**Le dégoût** : Le dégoût se manifeste par une lèvre supérieure levée, parfois avec une ride nasolabiale, et parfois un froncement de nez.[8]



Figure 2: Les émotions et leurs expressions faciales [W4].

### 2.2.3 Reconnaissance des expressions faciales

Le processus fondamental d'un système automatique d'analyse des expressions faciales comprend trois phases principales : la détection du visage à analyser, l'extraction des caractéristiques du visage pertinentes, et enfin, la classification de l'expression faciale. Dans ce chapitre, nous offrons une vue d'ensemble détaillée de ces phases clés du système de reconnaissance automatique des expressions faciales, tout en présentant diverses méthodes existantes pour chaque étape du processus [9].

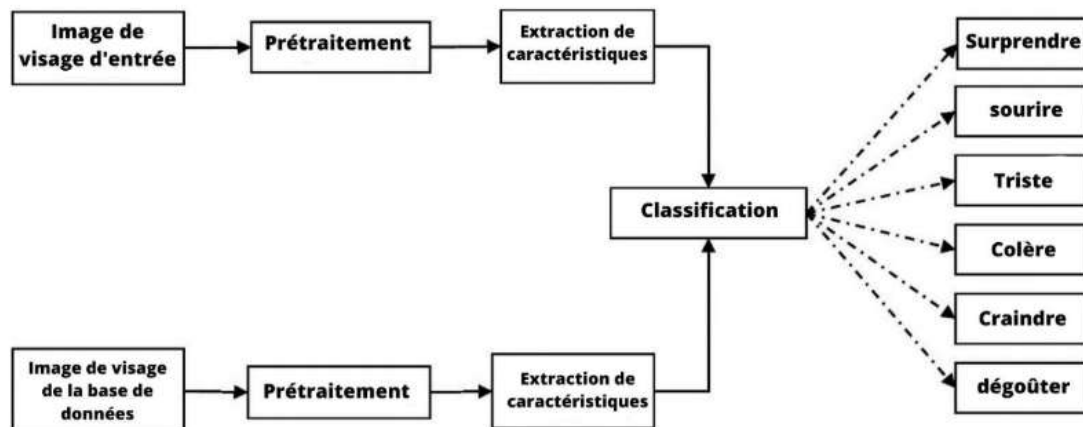


Figure 3 : Un mini architecture pour le processus de reconnaissance des émotions faciales.

### **2.3 Différence entre expression faciales et émotions**

Les expressions faciales et les émotions sont souvent confondues, mais elles ont des fonctions distinctes dans la communication et l'expérience humaine. Les expressions faciales se réfèrent aux mouvements ou aux configurations des muscles du visage qui peuvent transmettre des émotions à un observateur. Ces mouvements incluent le sourire, le froncement des sourcils et le plissement des yeux, et peuvent être délibérés ou involontaires.[10]

D'un autre côté, les émotions sont des états psychologiques complexes qui impliquent une expérience personnelle, des réactions physiologiques et des comportements expressifs. Elles comprennent des sentiments tels que la joie, la tristesse, la peur, la colère, la surprise et le dégoût, et jouent un rôle crucial dans la motivation et la régulation du comportement humain.[11]

Les expressions faciales sont principalement utilisées pour exprimer les émotions de manière non verbale, ce qui facilite les interactions sociales et la compréhension mutuelle.[12] En ce qui concerne les émotions, elles sont essentielles pour la survie et l'adaptation, impactant la prise de décision et la mémoire.[13] Les émotions et les expressions faciales interagissent souvent ; cependant, les expressions faciales peuvent être modifiées indépendamment des émotions réellement ressenties, comme lors de la simulation ou de la dissimulation d'émotions pour des raisons sociales.[14]

### 3 Détection des Emotions

#### 3.1 Méthodes de détection des émotions

La détection des émotions est un domaine de recherche interdisciplinaire impliquant la psychologie, l'informatique et les neurosciences. Voici un aperçu des méthodes directes et indirectes utilisées pour détecter les émotions :

##### 3.1.1 Méthodes directes : expression faciale

La détection des émotions à travers l'expression faciale est une méthode largement étudiée et utilisée. Les expressions faciales servent à communiquer de manière non-verbale, transmettant des signaux conscients et inconscients concernant nos émotions, nos intentions et notre réflexion. Elles jouent un rôle clé dans nos interactions sociales, qu'elles soient privées ou professionnelles [W5].



Figure 4: Expression faciale de la joie [W6].

### 3.1.2 Méthodes indirectes

La méthode indirecte de détection des émotions repose sur l'utilisation de mesures physiologiques, verbales et comportementales pour évaluer les réponses émotionnelles des individus. Les mesures physiologiques, telles que la fréquence cardiaque et la conductance de la peau, fournissent des indices indirects sur les états émotionnels [15].

### 3.1.3 Analyse comparative des méthodes directes et indirectes

Méthodes directes		Méthodes indirectes	
Avantages	Inconvénients	Avantages	Inconvénients
<ul style="list-style-type: none"> <li>- Observation directe des expressions faciales et des réponses émotionnelles.</li> <li>- Interprétation intuitive et facile à comprendre.</li> <li>- Données plus fiables et représentatives des émotions vécues au quotidien.</li> <li>- Permettent d'obtenir des informations subjectives détaillées.</li> <li>- Interaction dynamique qui peut révéler des informations riches.</li> </ul>	<ul style="list-style-type: none"> <li>- Sensibilité aux biais culturels et individuels dans l'interprétation des expressions faciales.</li> <li>- Possibilité de modification du comportement par conscience de l'observation.</li> <li>- Consomment beaucoup de temps et de ressources.</li> </ul>	<ul style="list-style-type: none"> <li>- Mesures physiologiques, verbales et comportementales pour obtenir des indices indirects sur les états émotionnels.</li> <li>- Possibilité d'évaluation dans des contextes où les individus ne sont pas conscients d'être observés.</li> </ul>	<ul style="list-style-type: none"> <li>- Interprétation complexe des mesures physiologiques pouvant être influencées par d'autres facteurs.</li> <li>- Nécessitent des équipements spécialisés et coûteux.</li> <li>- Peuvent être intrusives pour les participants.</li> <li>- Nécessite des compétences techniques avancées pour l'analyse.</li> <li>- Les données textuelles peuvent manquer de nuances.</li> <li>- Complexes à concevoir et à interpréter.</li> </ul>

Tableau 1 : Tableau comparatives des méthodes.

Dans notre projet, on a choisi d'utiliser la reconnaissance faciale des émotions.

### 3.2 Difficultés et défis de détection des émotions

La détection des émotions présente plusieurs défis :

**Variabilité interindividuelle** : Les expressions faciales peuvent varier considérablement d'une personne à l'autre. Les études montrent une sensibilité aux changements d'expression faciale ainsi qu'une attirance pour les visages joyeux, probablement dès les premiers jours après la naissance et sûrement lors des premiers mois. De plus, il existe une capacité de distinguer les visages joyeux d'autres expressions après les premiers mois, ainsi qu'une attirance plus tardive, vers 6-7 mois, pour les visages de peur due à une modulation de l'attention [w7].

**Contexte culturel** : Les expressions et l'interprétation des émotions peuvent différer selon les cultures. Selon certains neuroscientifiques, il y aurait une forte variabilité des patterns pour chaque émotion entre les études, laquelle pourrait être attribuée aux différences de contexte [w8].

**Conditions environnementales** : L'éclairage, les angles de vue et les occlusions peuvent affecter la précision de la détection. Des travaux ont permis de prendre en compte certaines difficultés, comme l'éclairage ou l'arrière-plan. Cependant, des problèmes tels que l'occlusion et le mouvement restent à résoudre, notamment dans des contextes où l'on ne dispose que d'une image statique, mais surtout lorsque la source est un flux vidéo [w9].

**Expressions faciales subtiles** : Les micro-expressions peuvent être difficiles à détecter en temps réel. Une micro-expression est une expression faciale très brève et involontaire que l'homme fait lorsqu'il ressent une émotion. Elle se caractérise par sa courte durée de 0,5 à moins de 0,25 secondes et ne peut être simulée. Il s'agit donc d'une vue honnête sur une émotion spontanée [w10].

### 3.3 Intérêts de détections des émotions

La détection des émotions est un domaine en plein essor, notamment grâce aux avancées de l'intelligence artificielle et des technologies de reconnaissance. Les émotions sont le produit conjoint de l'évaluation affective et cognitive d'un objet ou d'un événement, en relation avec les intérêts du sujet [w11].

Grâce à l'IA, les machines peuvent désormais détecter, interpréter et réagir aux émotions humaines, ce qui ouvre de nouvelles perspectives dans divers domaines tels que le :

- **Santé mentale** : Aider à diagnostiquer et à traiter les troubles émotionnels.
- **Marketing** : Analyser les réactions des consommateurs aux produits et publicités.
- **Sécurité** : Détecter des comportements suspects ou menaçants.
- **Interactions homme-machine** : Améliorer l'interaction entre les utilisateurs et les systèmes intelligents comme la vidéoconférence et les jeux... [w12].

Dans l'ensemble, la reconnaissance des expressions faciales est un domaine prometteur avec le potentiel de révolutionner diverses industries et d'améliorer notre vie quotidienne de multiples façons [16].

## 4 Conclusion

Dans ce chapitre, nous avons discuté des émotions (définition, type...). Nous avons également évoqué les expressions faciales (Les six expressions), puis on a discuté les différences entre eux. Nous avons ensuite évoqué les méthodes de détections des émotions. Nous avons terminé par les défis et intérêts de la détection des émotions.

Dans le prochain chapitre nous allons nous approfondir dans l'apprentissage automatique.

# **Chapitre II :**

## Apprentissage Automatique et Emotions

## **1 Introduction :**

Ce chapitre explore le rôle de l'apprentissage automatique dans la reconnaissance des émotions à partir des expressions faciales. Nous commencerons par définir les concepts clés de l'apprentissage automatique, puis nous examinerons comment ces techniques peuvent être appliquées à la détection des émotions.

## **2 Partie 1 : Apprentissage Automatique (Machine Learning) :**

Ce chapitre explore le rôle de l'apprentissage automatique dans la reconnaissance des émotions à partir des expressions faciales. Nous commencerons par définir les concepts clés de l'apprentissage automatique, puis nous examinerons comment ces techniques peuvent être appliquées à la détection des émotions.

### **2.1 Définition**

L'apprentissage automatique, ou Machine Learning, est un domaine de l'intelligence artificielle où les machines apprennent à résoudre des problèmes sans être explicitement programmées pour chaque tâche. Cela inclut l'analyse de données complexes, qu'elles soient symboliques (comme des probabilités), structurées sous forme de graphes ou encore présentées sous forme de courbes temporelles [W13].

### **2.2 Méthodes d'apprentissage automatiques**

#### **2.2.1 Méthodes supervisées et non supervisées**

##### **a. Apprentissage non-supervisé (ou classification automatique) :**

L'apprentissage non supervisé, dans le domaine de l'informatique et de l'intelligence artificielle, désigne la situation d'apprentissage automatique où les données ne sont pas étiquetées. Cela signifie qu'il s'agit de découvrir les structures sous-jacentes à ces données non étiquetées. Les méthodes non supervisées présentent une auto-organisation qui capture les modèles comme des densités de probabilité ou, dans le cas des réseaux de neurones, comme combinaison de préférences de caractéristiques neuronales encodées dans les poids et les activations de la machine [w14].

En comparaison avec l'apprentissage supervisé, l'apprentissage non supervisé est celui où

l'algorithme doit opérer à partir d'exemples non annotés. Par exemple, en fournissant à l'algorithme de nombreuses images de chats et d'accessoires pour chats, l'algorithme d'apprentissage non supervisé ne regrouperait pas tous les chats ensemble mais par exemple une pelote de laine avec un chat. L'apprentissage non supervisé est souvent utilisé dans le domaine de la reconnaissance vocale, comme pour l'utilisation de Siri ou Alexa [w15].

Dans le cadre de l'apprentissage non-supervisé, il est souvent difficile d'évaluer la performance d'un modèle vu que les vrais labels ne sont pas connus. Le but de l'apprentissage non-supervisé est de trouver des formes cachées dans un jeu de données non-labélisées.

En machine learning, la technique de l'apprentissage non supervisé consiste à entraîner des modèles, sans réaliser d'étiquetage manuel ou automatique des données au préalable. Les algorithmes regroupent les données en fonction de leur similitude, sans aucune intervention humaine. L'apprentissage non supervisé détecte des données ou individus présentant des caractéristiques ou des structures communes [w16].

### **b. Apprentissage supervisé**

L'apprentissage supervisé est une méthode d'apprentissage automatique dans laquelle un système reçoit des données d'entrée ainsi que les données de sortie attendues, et utilise ces informations pour apprendre à produire la sortie désirée. Les données d'entrée et de sortie sont étiquetées en vue de leur classification, établissant ainsi une base d'apprentissage pour le traitement ultérieur des données [w17].

Dans le contexte de l'intelligence artificielle et de l'apprentissage automatique, l'apprentissage supervisé est utilisé pour alimenter les algorithmes d'apprentissage avec des quantités connues qui soutiendront les futures décisions. Il se distingue de l'apprentissage non supervisé, où les données fournies au système d'IA ne sont ni étiquetées ni classées, et les algorithmes du système traitent les données sans aucun entraînement préalable [15].

L'apprentissage supervisé se distingue également de l'apprentissage par renforcement, dans lequel le résultat attendu n'est pas énoncé, et de l'apprentissage non supervisé, où l'agent formule des règles à partir de données brutes [w18].

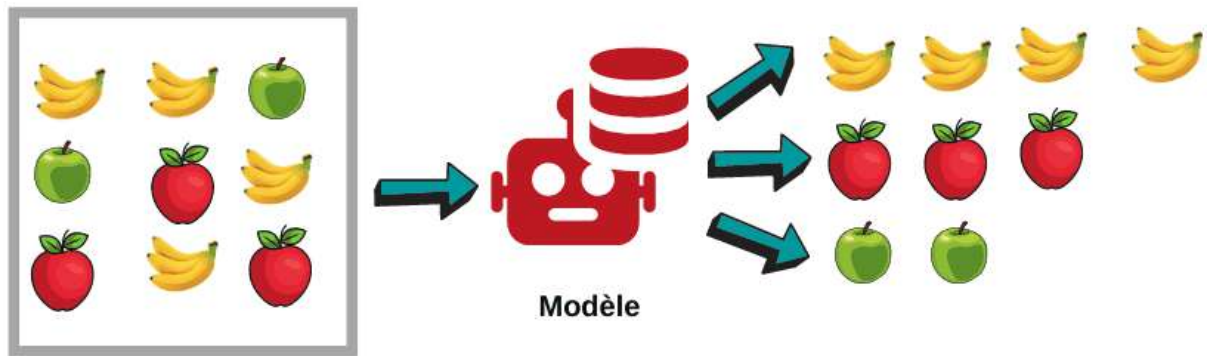


Figure 5: Modèle d'Apprentissage supervisé.

### 2.2.2 Méthodes classiques en ML

Les méthodes classiques en Machine Learning sont des approches traditionnelles utilisées pour résoudre des problèmes de classification, de prédiction et d'optimisation. Elles comprennent des techniques telles que les arbres de décision, les réseaux neuronaux, les méthodes ensemblistes, et les méthodes de clustering. Ces méthodes sont souvent caractérisées par une approche plus linéaire et une planification détaillée, et elles sont utilisées dans des contextes où les spécifications fonctionnelles et techniques sont précisément définies et détaillées pour lever tous les doutes et les risques possibles[w19].

Les méthodes classiques sont souvent utilisées pour des problèmes de petite taille ou pour des projets où la planification détaillée est essentielle. Elles sont également associées à des approches linéaires où les différentes phases de développement sont exécutées de manière séquentielle, avec une validation de chaque phase entraînant le début de la suivante. Cela limite les retours aux étapes précédentes et nécessite un cahier des charges exhaustif et détaillé, sans ambiguïté quant à l'interprétation des besoins, et qui n'évolue pas en cours de développement [w19].

Les méthodes classiques comprennent également des techniques de classification telles que les arbres de décision, qui peuvent être améliorées en utilisant des méthodes ensemblistes pour agréger des milliers de modèles ayant des avis divergents mais pouvant être chacun spécialisés sur des parties des données. Cette approche permet d'améliorer les performances des arbres de décision et des réseaux neuronaux pour des problèmes de classification où le nombre de classes

est important [w20].

En résumé, les méthodes classiques en Machine Learning sont des approches traditionnelles utilisées pour résoudre des problèmes de classification, de prédiction et d'optimisation, caractérisées par une approche plus linéaire, une planification détaillée et une utilisation dans des contextes où les spécifications fonctionnelles et techniques sont précisément définies et détaillées pour lever tous les doutes et les risques possibles. Elles comprennent des techniques telles que les arbres de décision, les réseaux neuronaux, les méthodes ensemblistes, et les méthodes de clustering [w20].

### **2.2.3 Méthodes modernes en ML**

Avec l'avènement de grandes quantités de données et de puissance de calcul, les méthodes modernes en ML ont gagné en popularité. Celles-ci incluent :

- Réseaux de neurones artificiels
- Réseaux de neurones convolutifs (CNN)
- Réseaux de neurones récurrents (RNN)
- Réseaux de neurones adversarial (GAN)
- Méthodes de traitement du langage naturel (NLP)

### **2.2.4 Méthodes ensemblistes**

Les méthodes ensemblistes sont une classe d'algorithmes d'apprentissage automatique qui combinent plusieurs modèles de base pour améliorer les performances prédictives. Elles sont devenues très populaires grâce à leur capacité à réduire le biais et la variance des modèles individuels [33].

Elles construisent un ensemble de modèles à partir du même algorithme d'apprentissage. Elles peuvent être divisées en trois catégories principales : le bagging, le Boosting et le Stacking [34].

### **2.2.5 Méthodes profondes**

Les "Méthodes Profondes" font référence à un ensemble de techniques et d'algorithmes avancés dans le domaine de l'apprentissage profond (deep learning), une branche de l'intelligence artificielle. Ces méthodes permettent d'apprendre à partir de grandes quantités de données brutes, sans avoir besoin d'extraire manuellement les caractéristiques.

Inspiré du fonctionnement des neurones biologiques, utilise des réseaux de neurones artificiels avec de multiples couches cachées. Ces réseaux apprennent des représentations de plus en plus abstraites des données d'entrée, permettant de résoudre des tâches complexes telles que la reconnaissance d'objets, la traduction automatique ou la génération de textes [w21].

### **2.2.6 Méthodes par renforcement**

L'apprentissage par renforcement est une méthode d'apprentissage automatique utilisée en intelligence artificielle. Il consiste à apprendre les actions à prendre par un agent autonome, tel qu'un robot, un agent conversationnel, ou un personnage dans un jeu vidéo, afin d'optimiser une récompense quantitative au fil du temps

Les algorithmes d'apprentissage par renforcement se basent sur un système de récompenses où à chaque action effectuée, l'agent reçoit une récompense ou une pénalité de la part de l'environnement, ce qui lui permet d'ajuster sa stratégie. Cette méthode se distingue de l'apprentissage supervisé, où la proposition est suivie du résultat attendu, et de l'apprentissage non supervisé, qui ne comporte aucune rétroaction [w22].

En résumé, l'apprentissage par renforcement est une méthode d'apprentissage statistique qui permet à un agent d'apprendre à prendre des décisions en fonction des récompenses reçues de son environnement.

### **3 Partie 2 : Systèmes de détection d'émotion**

#### **3.1 Taxonomie des systèmes**

Les systèmes de détection d'émotion peuvent être classés en différentes catégories en fonction de leur architecture et de leur approche algorithmique.

##### **3.1.1 Système à base d'apprentissage automatique**

Les systèmes à base d'apprentissage automatique (ou "machine learning") sont des systèmes informatiques capables d'apprendre et de s'améliorer grâce à l'expérience, sans être explicitement programmés. Ils utilisent des algorithmes et des techniques avancées pour construire des modèles analytiques à partir de données d'entraînement.

Ils fonctionnent en analysant d'énormes quantités de données pour détecter des modèles et des relations. Ils sont entraînés sur un ensemble de données d'entraînement, afin de construire un modèle capable de prendre des décisions précises ou de faire des prédictions sur de nouvelles données [w23].

Il existe différents types d'algorithmes d'apprentissage automatique, notamment :

- Apprentissage supervisé : Le système apprend à partir d'un ensemble de données étiquetées, en associant les entrées aux sorties souhaitées [w24].
- Apprentissage non supervisé : Le système découvre des modèles et des structures inhérentes dans les données d'entrée, sans aucune étiquette ou sortie prédéfinie[w23].
- Apprentissage par renforcement : Le système apprend par essais et erreurs, en interagissant avec un environnement et en recevant des récompenses ou des pénalités pour ses actions [w25].

Ils sont largement utilisés dans divers domaines, tels que :

- La reconnaissance de formes et d'images
- Le traitement du langage naturel

- La détection de fraudes et de spam
- Les systèmes de recommandation
- La robotique et l'automatisation
- Les jeux et la finance

Bien que puissants, ces systèmes nécessitent une grande quantité de données d'entraînement de haute qualité et peuvent présenter des biais ou des problèmes de sécurité s'ils ne sont pas conçus et déployés correctement.

**Exemple :** Un filtre anti-spam pour les emails.

- Explication : Ce système utilise des algorithmes d'apprentissage automatique comme la régression logistique ou les machines à vecteurs de support (SVM) pour classer les emails en "spam" ou "non spam". Il apprend à partir d'un ensemble de données étiquetées contenant des exemples d'emails précédemment identifiés comme spam ou non. Au fur et à mesure qu'il traite plus d'emails, il devient de plus en plus précis dans la détection des spams.

### 3.1.2 Systèmes à base d'apprentissage profond

L'apprentissage profond (Deep Learning en anglais) est un sous-ensemble des techniques d'apprentissage automatique (Machine Learning) basé sur les réseaux de neurones artificiels [w26].

Ces systèmes sont conçus pour apprendre de grandes quantités de données, en imitant la structure et le fonctionnement du cerveau humain [w27].

Les systèmes à base d'apprentissage profond exploitent plusieurs couches de réseaux de neurones artificiels pour extraire et transformer les données d'entrée en un niveau d'abstraction plus élevé. Ces niveaux d'abstraction permettent aux systèmes d'apprendre des représentations complexes des données de manière hiérarchique [w26].

Les systèmes d'apprentissage profond ont connu un essor ces dernières années, grâce à l'augmentation de la puissance de calcul et à la disponibilité de grandes quantités de données.

Ils sont désormais largement utilisés dans de nombreux domaines[w28] tels que :

- La reconnaissance d'images et d'objets.
- La reconnaissance vocale et le traitement du langage naturel.
- Les systèmes de recommandation (e-commerce, streaming).
- La détection de fraudes et d'anomalies.
- Les systèmes d'aide à la décision médicale Les véhicules autonomes.
  - Exemple : Un système de reconnaissance faciale.
  - Explication : Ce système utilise des réseaux de neurones profonds, comme les réseaux de neurones convolutifs (CNN), pour identifier et reconnaître des visages dans des images. Il est formé sur de vastes ensembles de données contenant des millions de photos étiquetées. Grâce à ses multiples couches de neurones, il peut extraire des caractéristiques complexes et apprendre des représentations hiérarchiques des images, ce qui lui permet de reconnaître des visages avec une grande précision.

### 3.1.3 Systèmes basé apprentissage par renforcement

L'apprentissage par renforcement (RL) est une branche de l'intelligence artificielle qui s'intéresse à la façon dont un agent apprend à prendre des décisions optimales dans un environnement incertain afin de maximiser une récompense définie [w29].

C'est une approche différente de l'apprentissage supervisé où les données d'entraînement sont étiquetées. En RL, l'agent découvre la meilleure stratégie par essais et erreurs [w30].

Son fonctionnement :

Un système basé sur l'apprentissage par renforcement comprend généralement un agent, un environnement et une fonction de récompense. L'agent observe l'état actuel de l'environnement, choisit une action, puis reçoit une récompense de l'environnement en fonction des conséquences de son action. Son objectif est d'apprendre une politique qui maximise la récompense cumulée à long terme [w31].

**Exemple :** AlphaGo de Google DeepMind.

• **Explication :** AlphaGo est un programme d'intelligence artificielle qui utilise l'apprentissage par renforcement pour jouer au jeu de Go. Il apprend en jouant des millions de parties contre lui-même et en recevant des récompenses basées sur ses performances. En combinant des réseaux de neurones profonds avec des techniques de recherche arborescente, AlphaGo a pu maîtriser le jeu et battre les meilleurs joueurs humains, démontrant la puissance de l'apprentissage par renforcement pour résoudre des problèmes complexes avec des environnements dynamiques.

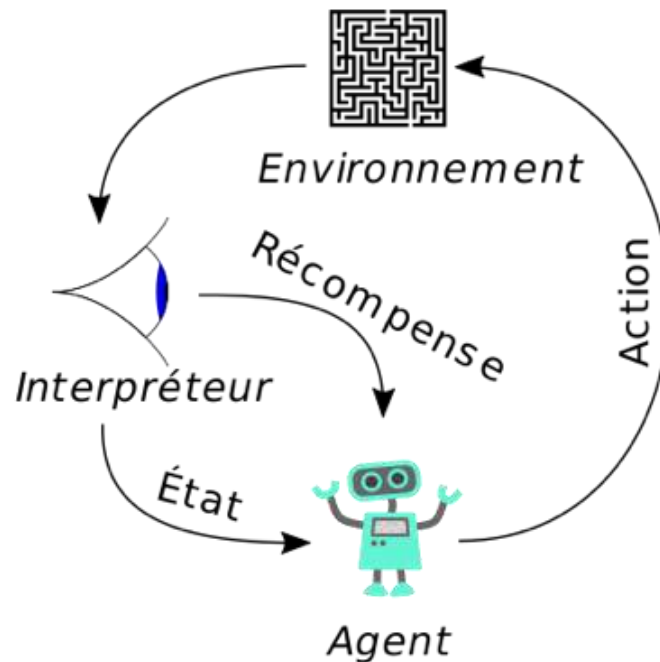


Figure 6 : Apprentissage par renforcement.

#### 4 Conclusion

Ce chapitre a examiné comment les techniques d'apprentissage automatique peuvent être appliquées à la détection des émotions.

**Chapitre III :**  
Conception d'un  
système de détection  
d'émotions basé sur  
CNNs

## **1. Introduction**

Dans ce chapitre, nous allons nous concentrer sur la conception et le développement d'un système de détection d'émotion destiné à améliorer le contrôle de satisfaction des utilisateurs.

Nous présenterons une vue d'ensemble du système, détaillerons sa conception et son architecture, et utiliserons des diagrammes UML pour illustrer les différents composants et leur interaction.

En particulier, nous mettrons en avant l'utilisation des réseaux de neurones convolutifs (CNN) pour l'analyse des expressions faciales, montrant comment cette technologie peut être appliquée pour catégoriser les émotions et optimiser l'expérience utilisateur.

## **2. Aperçu du système**

Il s'agit d'un système de reconnaissance faciale des émotions, notre programme commence par la capture d'une image déclenchée par l'utilisateur via une caméra. Une fois l'image capturée, elle est convertie en niveaux de gris pour simplifier les étapes de traitement ultérieures. Ensuite, le programme utilise des algorithmes de vision par ordinateur pour détecter les visages présents dans l'image.

Après la détection des visages, le programme initialise des trackers pour suivre les visages détectés dans les images successives, assurant ainsi une continuité et une précision dans la reconnaissance. Les expressions faciales des visages suivis sont alors analysées pour reconnaître les émotions telles que la joie, la tristesse, la colère, etc.

Les informations sur les visages détectés et les émotions reconnues sont ensuite mises à jour dans une base de données.

Enfin, les statistiques des émotions sont mises à jour pour permettre des analyses futures et ajuster le comportement du système si nécessaire.

### **3. Conception**

C'est un système de détection faciale des émotions en temps réel. Il commence par capturer en continu des images vidéo, qu'il convertit ensuite en niveaux de gris pour simplifier le traitement. Le système scanne ces images pour détecter et identifier les régions faciales, puis initialise des traqueurs pour suivre les visages détectés. À chaque image, il extrait la région d'intérêt (le visage), la prétraite et prédit l'émotion exprimée. Les visages connus sont mis à jour dans une base de données, et le système suit la dernière fois où chaque visage a été vu. Il met également à jour les statistiques des émotions détectées et calcule leurs pourcentages. Enfin, le système gère l'ajout de nouveaux visages et met à jour les positions des visages dans la base de données, assurant ainsi un suivi précis et une reconnaissance en temps réel des émotions.

### **4. Diagrammes UML**

#### **4.1 Qu'est-ce qu'un diagramme UML ?**

Le diagramme UML, basé sur le langage de modélisation unifié (UML), est un modèle dynamique facilitant la visualisation des processus et des séquences. Il capture les éléments essentiels d'un système, tels que les acteurs, les rôles, les actions et les artefacts. Les diagrammes UML sont indispensables à la conception et au développement des logiciels modernes, car ils améliorent la compréhension, facilitent les modifications et la maintenance, et assurent une documentation complète des informations cruciales du système [w32].

#### **4.2 Diagramme de cas d'utilisation**

Le diagramme de cas d'utilisation est un diagramme comportemental qui représente les fonctionnalités d'un système du point de vue des utilisateurs (acteurs). Il montre les interactions entre les acteurs et les cas d'utilisation, décrivant ainsi les différentes actions que les utilisateurs peuvent effectuer et les résultats attendus. Le diagramme de cas d'utilisation est souvent utilisé pour définir les exigences fonctionnelles d'un système [17].

Voici notre diagramme de cas d'utilisateur :

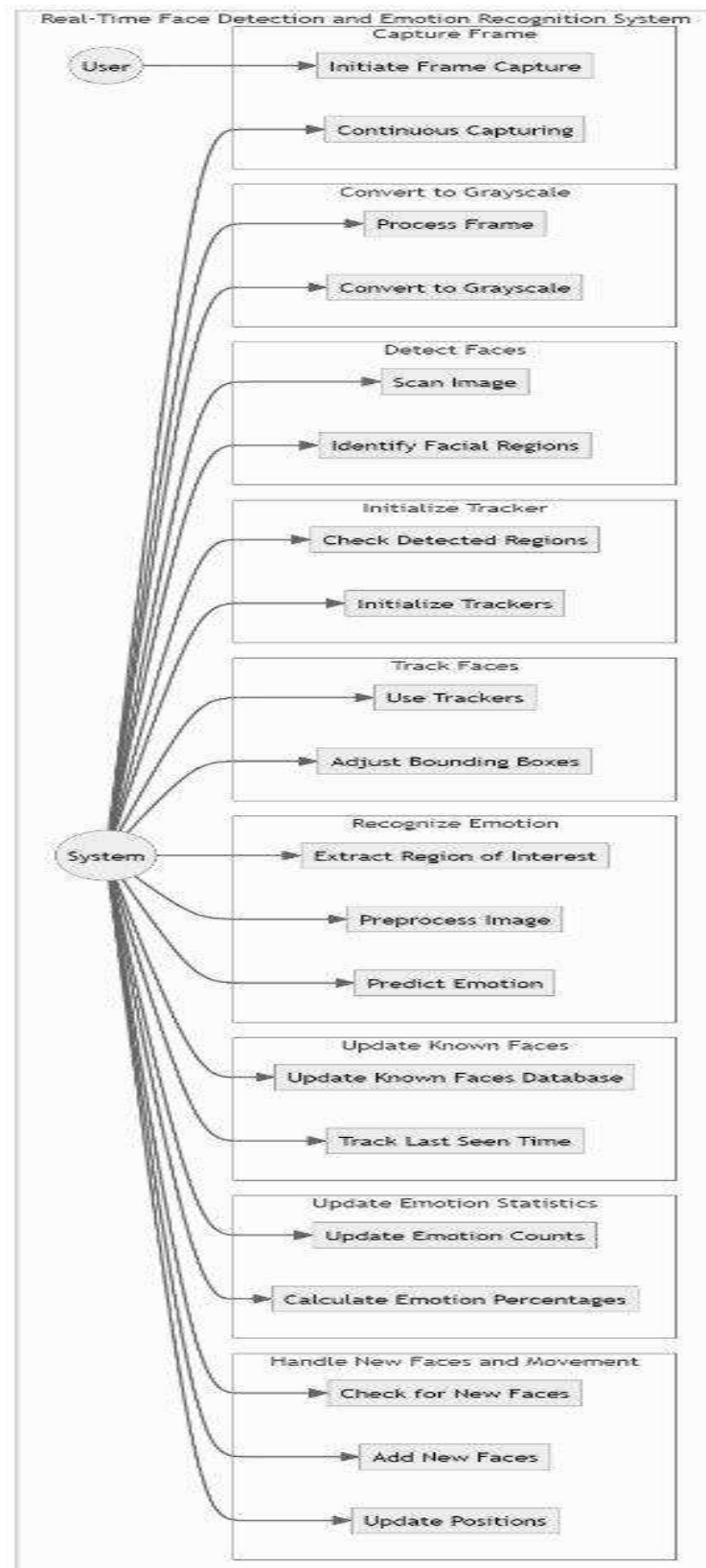


Figure 7 : Diagramme cas d'utilisation.

Ce diagramme représente un système de détection de visage en temps réel et de reconnaissance des émotions. Il se compose de plusieurs étapes regroupées en différents modules, chacun réalisant une fonction spécifique. Voici une explication détaillée des différentes étapes et de leur utilisation :

**1. Utilisateur :**

- **Initiate Frame Capture** : L'utilisateur déclenche la capture de la première image (cadre).
- **Continuous Capturing** : Une fois déclenchée, la capture d'images se poursuit de manière continue.

**2. Système :**

- Capture Frame :
  - **Process Frame** : L'image capturée est traitée.
  - **Convert to Grayscale** : L'image en couleur est convertie en niveaux de gris pour simplifier les traitements suivants.
- Detect Faces :
  - **Scan Image** : L'image en niveaux de gris est scannée pour détecter les visages.
  - **Identify Facial Regions** : Les régions faciales potentielles sont identifiées dans l'image.
- Initialize Tracker :
  - **Check Detected Regions** : Les régions faciales détectées sont vérifiées.
  - **Initialize Trackers** : Des trackers sont initialisés pour suivre les visages

déTECTÉS.

- Track Faces :
  - **Use Trackers** : Les trackers sont utilisés pour suivre les visages dans les images suivantes.
  - **Adjust Bounding Boxes** : Les boîtes de délimitation autour des visages sont ajustées pour correspondre aux mouvements.
- Recognize Emotion :
  - **Extract Region of Interest** : La région d'intérêt (le visage) est extraite pour l'analyse.
  - **Preprocess Image** : L'image extraite est prétraitée pour l'analyse des émotions.
  - **Predict Emotion** : L'émotion est prédite à partir de l'image prétraitée.
- Update Known Faces :
  - **Update Known Faces Database** : La base de données des visages connus est mise à jour avec les nouvelles informations.
  - **Track Last Seen Time** : Le temps de la dernière apparition de chaque visage est suivi.
- Update Emotion Statistics :
  - **Update Emotion Counts** : Les statistiques des émotions détectées sont mises à jour.
  - **Calculate Emotion Percentages** : Les pourcentages des différentes émotions sont calculés.
- Handle New Faces and Movement:
  - **Check for New Faces** : Le système vérifie si de nouveaux visages apparaissent.

- **Add New Faces** : Les nouveaux visages sont ajoutés à la base de données.
- **Update Positions** : Les positions des visages sont mises à jour pour le suivi continu.

### 4.3 Diagramme de séquence

Le diagramme de séquence est un diagramme comportemental qui montre l'ordre chronologique des messages échangés entre les objets du système. Il met l'accent sur les interactions dynamiques entre les objets et permet de modéliser le flux d'exécution d'un scénario particulier. Le diagramme de séquence est souvent utilisé pour décrire le comportement des cas d'utilisation et l'enchaînement des actions entre les objets [18].

Voici notre diagramme de classe :

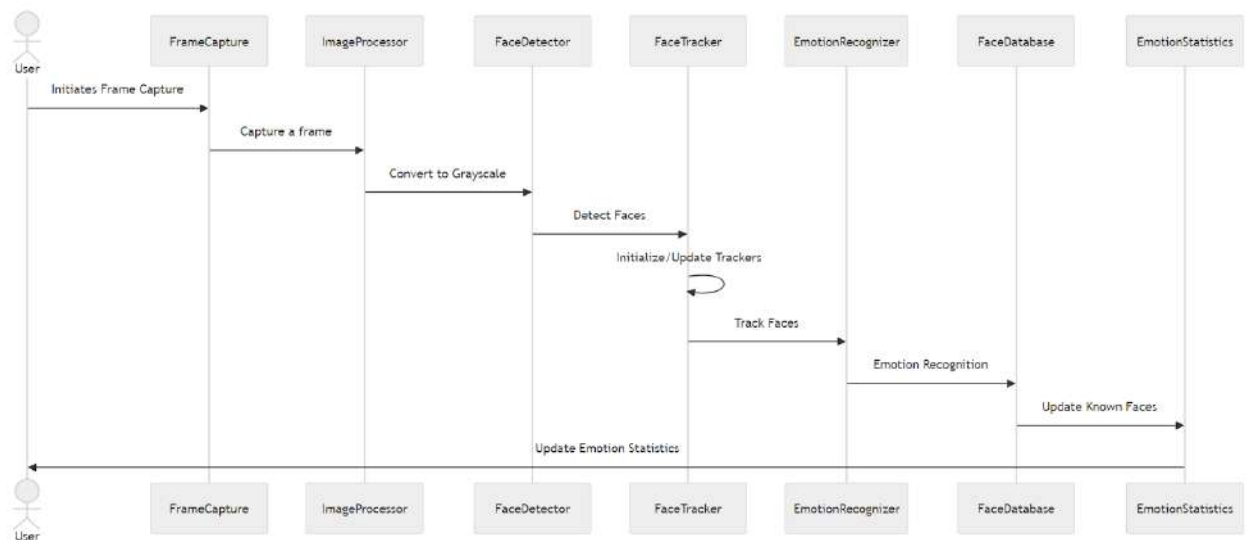


Figure 8 : Diagramme de séquence.

Le diagramme illustre le processus de reconnaissance des émotions en détail, montrant les interactions entre les différents composants du système

. Voici les différentes étapes et interactions :

1. User (Utilisateur) :

L'utilisateur initie la capture d'une image.

2. FrameCapture (Capture d'image) :

Le composant FrameCapture capture une image en réponse à la demande de l'utilisateur.

3. ImageProcessor (Processeur d'image) :

Le composant ImageProcessor prend l'image capturée et la convertit en niveaux de gris pour faciliter la détection des visages.

4. FaceDetector (Détecteur de visages) :

Le composant FaceDetector analyse l'image en niveaux de gris pour détecter les visages présents.

5. FaceTracker (Suivi des visages) :

Après la détection des visages, le composant FaceTracker initialise ou met à jour les trackers pour suivre les visages détectés.

Il continue de suivre les visages dans les images successives.

6. EmotionRecognizer (Reconnaissance des émotions) :

Le composant EmotionRecognizer analyse les expressions faciales des visages suivis pour reconnaître les émotions.

7. FaceDatabase (Base de données des visages) :

Le composant FaceDatabase met à jour les informations sur les visages connus avec les nouvelles données obtenues.

8. EmotionStatistics (Statistiques des émotions) :

Le composant EmotionStatistics met à jour les statistiques des émotions basées sur les nouvelles reconnaissances effectuées.

Résumé des Interactions :

- **L'utilisateur** initie la capture d'image.
- **FrameCapture** capture l'image et la passe à **ImageProcessor**.
- **ImageProcessor** convertit l'image en niveaux de gris et la transmet à **FaceDetector**.
- **FaceDetector** détecte les visages et envoie cette information à **FaceTracker**.
- **FaceTracker** initialise ou met à jour les trackers et continue à suivre les visages.
- **EmotionRecognizer** reconnaît les émotions des visages suivis et transmet les données à FaceDatabase.
- **FaceDatabase** met à jour les informations sur les visages connus et **EmotionStatistics** met à jour les statistiques des émotions.

**User** reçoit finalement la mise à jour ou le résultat du processus.

#### 4.4 Diagramme de classes

Le diagramme de classe est un outil de modélisation qui représente la structure statique d'un système en mettant l'accent sur les classes, les attributs, les méthodes et les relations entre les différentes classes. Il permet de visualiser les entités du système, leurs propriétés et leurs interactions. Le diagramme de classe est souvent utilisé pour concevoir et documenter les modèles de données et les structures d'un système [19].

Voici notre diagramme de classe :

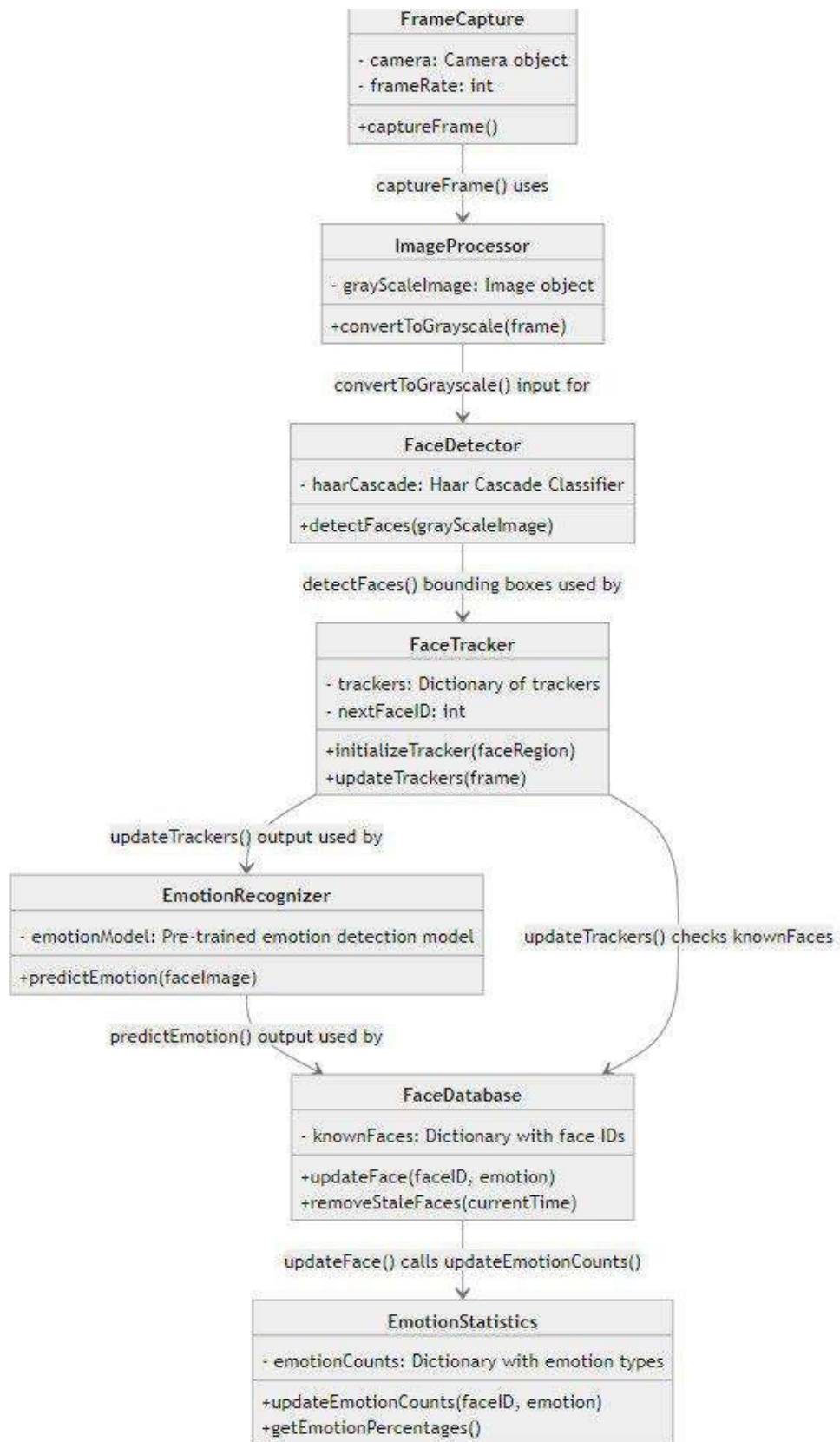


Figure 9 : Diagramme de classe.

## 1. FrameCapture :

### ○ Attributs :

- Camera : Objet de caméra.
- FrameRate: Taux de capture des images (images par seconde).

### ○ Méthode :

- CaptureFrame(): Capture une image de la caméra.

### ○ Interaction :

- Utilise la méthode captureFrame() pour capturer des images, lesquelles sont ensuite traitées par la classe ImageProcessor.

## 2. ImageProcessor :

### ○ Attributs :

- grayScaleImage: Objet Image en niveaux de gris.

### ○ Méthode :

- convertToGrayscale(frame): Convertit une image couleur en niveaux de gris.

### ○ Interaction :

- La méthode convertToGrayscale() prend l'image capturée par FrameCapture et la convertit en niveaux de gris, qui est ensuite utilisée comme entrée pour FaceDetector.

## 3. FaceDetector :

### ○ Attributs :

- haarCascade: Classificateur de cascade de Haar pour la détection des visages.

- Méthode :
  - `detectFaces(grayScaleImage)`: Détecte les visages dans une image en niveaux de gris.
- Interaction :
  - La méthode `detectFaces()` utilise l'image en niveaux de gris pour détecter les visages, les régions de visage détectées sont utilisées par `FaceTracker`.

#### 4. `FaceTracker` :

- Attributs :
  - `trackers`: Dictionnaire des suivis de visage.
  - `nextFaceID`: Identifiant pour le prochain visage.
- Méthodes :
  - `initializeTracker(faceRegion)`: Initialise le suivi pour une nouvelle région de visage.
  - `updateTrackers(frame)`: Met à jour les suivis de visage dans le cadre actuel.
- Interaction :
  - Utilise les régions de visage détectées pour initialiser et mettre à jour les suivis, les résultats sont ensuite utilisés par `EmotionRecognizer` et `FaceDatabase`.

#### 5. `EmotionRecognizer` :

- Attributs :
  - `emotionModel`: Modèle pré-entraîné de détection des émotions.
- Méthode :
  - `predictEmotion(faceImage)`: Prédit l'émotion à partir d'une image de visage.

- Interaction :
  - Utilise les images de visage provenant de FaceTracker pour prédire les émotions, les résultats sont utilisés par FaceDatabase.

## 6. FaceDatabase :

- Attributs :
  - knownFaces: Dictionnaire avec les identifiants de visage.
- Méthodes :
  - updateFace(faceID, emotion): Met à jour l'émotion pour un visage donné.
  - removeStaleFaces(currentTime): Supprime les visages qui ne sont plus suivis.
- Interaction :
  - Utilise les prédictions d'émotion de EmotionRecognizer pour mettre à jour les visages connus et leur émotion, les résultats sont utilisés par EmotionStatistics.

## 7. EmotionStatistics :

- Attributs :
  - emotionCounts: Dictionnaire des types d'émotions.
- Méthodes :
  - updateEmotionCounts(faceID, emotion): Met à jour le comptage des émotions.
  - getEmotionPercentages(): Calcule les pourcentages des différentes émotions.
- Interaction :
  - Utilise les mises à jour des visages provenant de FaceDatabase pour tenir à jour les statistiques des émotions.

## 5. Modèle ML proposé

Nous avons choisi le modèle CNN (Convolutional Neural Network) séquentiel pour notre système de détection d'émotions pour plusieurs raisons, particulièrement dans le contexte de l'analyse d'images ou de vidéos, qui sont couramment utilisées pour détecter des expressions faciales et en inférer les émotions. Voici quelques raisons principales pour lesquelles ce choix est pertinent :

- Efficacité dans le traitement d'images

Les CNN sont spécifiquement conçus pour traiter les données structurées en grilles, comme les images. Les convolutions permettent de détecter automatiquement les caractéristiques importantes (features) telles que les contours, les textures et les motifs, ce qui est crucial pour analyser les expressions faciales.

- Extraction de caractéristiques

Les couches convolutives extraient des caractéristiques hiérarchiques des images. Les premières couches détectent des traits simples (bords, textures), tandis que les couches plus profondes combinent ces traits pour détecter des formes plus complexes (parties du visage, expressions).

- Réduction de la complexité

Les opérations de convolution et de pooling (sous-échantillonnage) réduisent la dimensionnalité des données tout en conservant les informations essentielles, ce qui améliore l'efficacité computationnelle et la performance du modèle.

- Généralisation et robustesse

Grâce à la Translational invariance (capacité à reconnaître des objets peu importe leur position dans l'image) et à la capacité de reconnaître les motifs locaux indépendamment de leur contexte global, les CNN offrent une robustesse accrue dans la reconnaissance des expressions faciales variées et dans des conditions différentes.

### ➤ Modèles séquentiels

Un modèle CNN séquentiel permet une construction simple et intuitive des couches du réseau. En empilant les couches séquentiellement, on peut facilement expérimenter et ajuster l'architecture pour optimiser les performances du modèle sur un jeu de données spécifique de détection d'émotions.

### ➤ Disponibilité et implémentation

Les frameworks de deep learning comme TensorFlow et Keras facilitent l'implémentation des CNN séquentiels avec des bibliothèques préexistantes et des outils de visualisation, ce qui accélère le développement et le déploiement des modèles.

### ➤ Expérience et littérature

Il existe une abondante littérature scientifique et de nombreuses études de cas démontrant l'efficacité des CNN pour la reconnaissance faciale et la détection d'émotions. Tirer parti de cette connaissance et des modèles pré-entraînés (comme VGG, ResNet) peut grandement améliorer la performance et accélérer le développement.

## 6. Présentation du modèle CNN

Les réseaux de neurones convolutionnels (CNN) sont actuellement les modèles les plus performants dans des domaines tels que la reconnaissance et la classification d'images. Ils ont démontré leur capacité à identifier des éléments tels que des visages, des objets, des panneaux de signalisation, et même à être utilisés dans la conduite autonome. Durant cette phase de formation, nous nous concentrerons exclusivement sur les CNN (Convolutional Neural Networks)[20].

Le terme "réseau de neurones convolutif" fait référence à l'utilisation d'une opération mathématique appelée convolution, qui est une opération linéaire spécifique. Les réseaux convolutifs sont des réseaux de neurones qui utilisent la convolution à la place de la multiplication matricielle dans au moins une de leurs couches. Ils se composent de deux parties principales.

Lorsque nous fournissons une image en entrée, celle-ci est représentée sous forme de matrice de pixels à deux dimensions pour les images en niveaux de gris et à trois dimensions pour les images en couleur (avec les canaux Rouge, Vert et Bleu). Cette image traverse la première partie d'un CNN, appelée la partie convolutive. Cette section agit comme un extracteur de caractéristiques d'images.

L'image passe successivement à travers une série de filtres ou de noyaux de convolution, qui la transforment en de nouvelles images appelées cartes de caractéristiques ou "feature maps". Certains de ces filtres intermédiaires peuvent réduire la résolution de l'image en utilisant une opération de maximum local, souvent appelée "pooling". Finalement, les cartes de caractéristiques sont aplaties et combinées en un vecteur de caractéristiques appelé le "code CNN".

Le résultat en sortie de la partie convolutive est ensuite utilisé comme entrée pour une deuxième partie du réseau, composée de couches entièrement connectées, similaires à un perceptron multicouche. Cette partie a pour rôle de combiner les caractéristiques extraites par l'ensemble du réseau pour classifier l'image. La couche de sortie contient généralement un neurone par classe, fournissant des valeurs numériques normalisées entre 0 et 1, qui représentent la distribution de probabilité sur les différentes classes [21].

Les CNN sont particulièrement adaptés au traitement de grandes quantités de données telles que les images pour plusieurs raisons :

Les images présentent une corrélation spatiale, ce qui signifie que les valeurs des pixels voisins sont généralement très similaires. Les couches de convolution sont capables de capturer ces liens spatiaux entre les données, permettant ainsi de détecter efficacement les motifs locaux dans l'image.

Les couches de convolution sont souvent suivies de couches de sous-échantillonnage (pooling), ce qui permet de réduire significativement la taille des données. Ces couches de pooling aident à diminuer la complexité computationnelle tout en conservant les caractéristiques essentielles de l'image [22].

## **6.1 Différents modules d'un réseau de neurones convolutif**

Les CNN sont composés de trois types de couches distinctes : les couches de convolution, les couches de pooling et les couches entièrement connectées. Lorsqu'elles sont disposées en série, elles constituent l'architecture d'un CNN. Chaque type de couche a une fonction particulière :

La couche de convolution (CONV) analyse les données à l'intérieur d'une fenêtre spécifique appelée champ récepteur.

La couche de pooling (POOL) a pour rôle de réduire la taille des données en compressant généralement l'image intermédiaire grâce à une opération de sous-échantillonnage.

La couche d'activation (ReLU), souvent abrégée en "ReLU" pour sa fonction d'activation linéaire rectifiée, est responsable d'introduire une non-linéarité dans le réseau neuronal.

La couche FC, qui présente une structure similaire à celle d'un perceptron traditionnel.

Pour finir, la couche de perte (LOSS) est chargée d'évaluer la disparité entre les prédictions du modèle et les valeurs réelles [23].

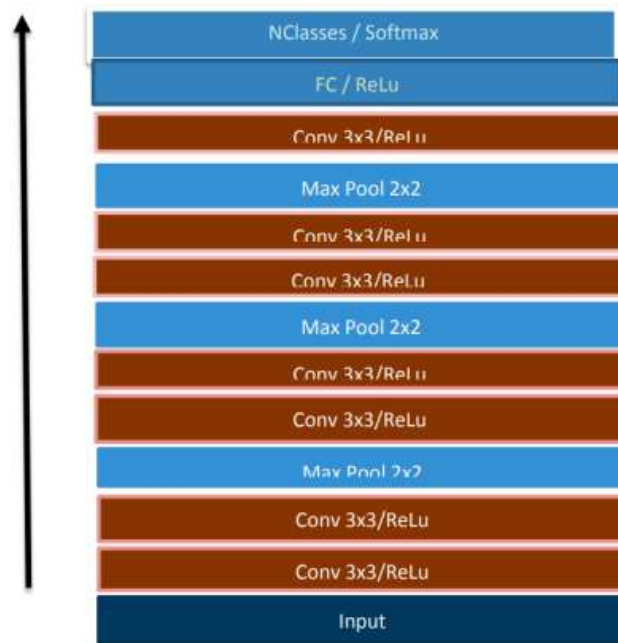


Figure 10 : Architecture générale d'un CNN.

### 6.1.1 La convolution

La convolution est le pilier des réseaux de neurones convolutionnels. Au départ, c'est un concept mathématique majeur en traitement d'images, car il permet de distiller des caractéristiques cruciales à partir des images, facilitant l'application de filtres pertinents. En pratique, la convolution opère en prenant une image en entrée ainsi qu'un filtre (qui est en substance une autre image), effectue un calcul spécifique, et génère une nouvelle image en sortie (habituellement de dimensions plus réduites).

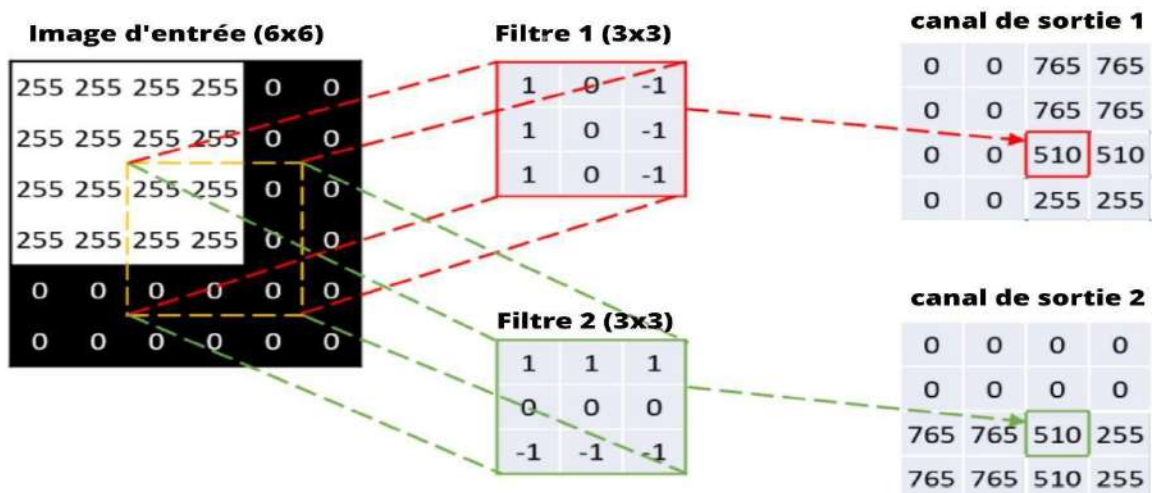


Figure 11: Illustration d'une opération de convolution entre deux couches.

La couche de convolution, également connue sous le nom de volume de sortie, est dimensionnée en fonction de trois hyperparamètres principaux : la profondeur, le pas et la marge [22].

- **La profondeur de la couche** représente le nombre de noyaux de convolution, ou le nombre de neurones associés à un même champ récepteur.
- **Le pas (stride)** contrôle le chevauchement des champs récepteurs. Si le pas est plus petit, les champs récepteurs se chevauchent davantage, ce qui augmente la taille du volume de sortie.
- **La marge (zero-padding)**, souvent réglé à 0, implique d'ajouter des zéros autour des bords du volume d'entrée. La taille de ce "zero-padding" constitue le troisième hyperparamètre.

L'utilisation de cette marge permet de contrôler la dimension spatiale du volume de sortie. Dans certains cas, il est souhaitable de conserver la même surface que celle du volume d'entrée [22].

À chaque image utilisée pour l'apprentissage, des filtres sont appliqués à différentes résolutions, et la sortie de chaque image convoluée est utilisée comme entrée de la couche suivante.

### 6.1.2 Les différentes convolutions

Effectivement, il existe diverses méthodes de convolution, bien que la convolution standard soit souvent privilégiée. Toutefois, il est bénéfique de se familiariser avec les différentes options disponibles [20].

- **La convolution classique** : Elle consiste en un décalage du noyau entre chaque calcul, et le padding détermine la manière dont on peut « déborder » de l'image pour appliquer la convolution.
- **La dilated convolution** : Elle est similaire à la convolution classique, mais avec la particularité que le noyau est étendu (par exemple, en sautant un pixel sur deux pour calculer la convolution). Un paramètre supplémentaire, le taux de dilation, indique le nombre de pixels à ignorer.

- La transposed convolution : Elle construit la sortie comme si on inversait une convolution sur l'image.
- La séparable convolution : Elle représente une convolution décomposable en convolutions plus simples.

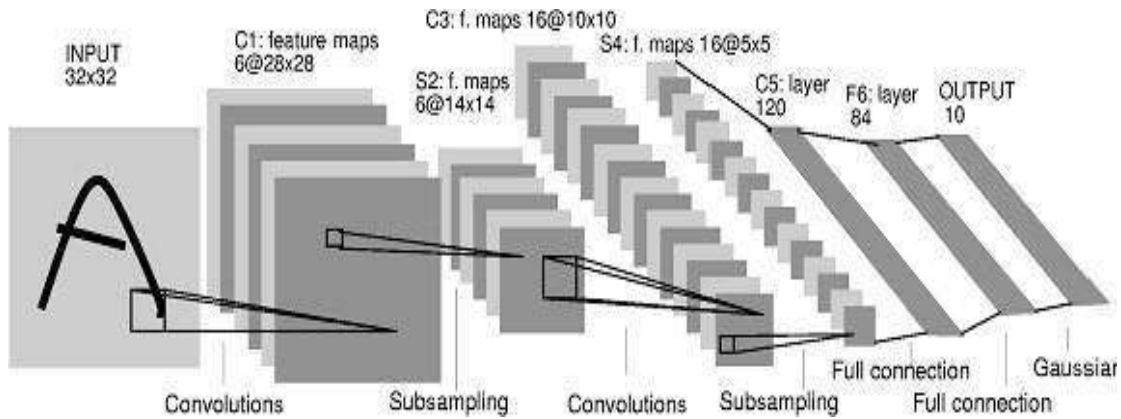


Figure 12 : Les différentes couches d'un réseau de neurones convolutif standard.

### 6.1.3 Le pooling

Cette couche est habituellement placée entre deux couches de convolution. Elle prend en charge plusieurs cartes de caractéristiques en entrée et applique l'opération de Pooling à chacune d'elles. L'objectif du Pooling est de diminuer la taille des images tout en maintenant intactes leurs caractéristiques fondamentales.

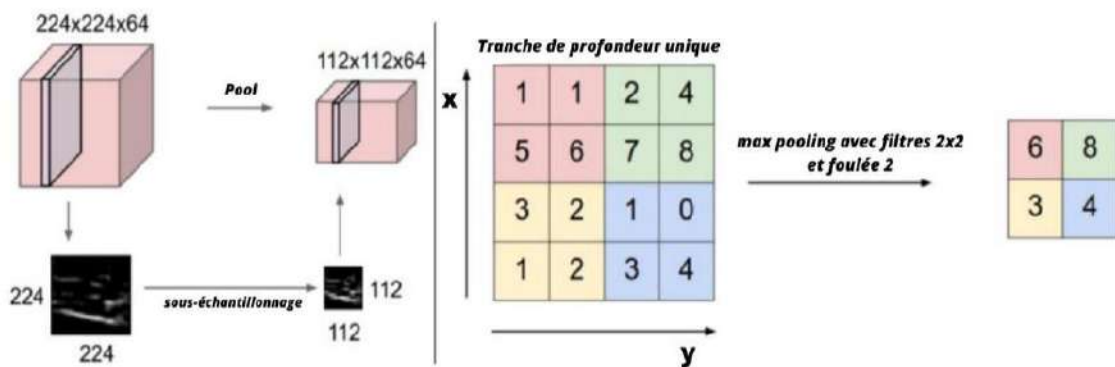


Figure 13 : Illustration du Max Pooling.

Pour ce faire, l'image est découpée en cellules régulières, et à l'intérieur de chaque cellule, seule la valeur maximale est retenue. Généralement, des cellules carrées de petite taille

sont utilisées pour limiter la perte d'informations. Les choix typiques incluent des cellules carrées de 2 x 2 pixels sans chevauchement ou des cellules de 3 x 3 pixels avec un chevauchement de 2 pixels. En sortie, le nombre de cartes de caractéristiques reste inchangé, mais leur taille est significativement réduite. Cette couche de pooling aide à réduire le nombre de paramètres et de calculs dans le réseau, ce qui améliore son efficacité et prévient le surapprentissage...

Il existe plusieurs types de pooling, dont les principaux sont les suivants :

- **Max pooling** : Ce type de pooling implique la sélection de la valeur maximale dans une région donnée. Il est largement privilégié pour sa rapidité de calcul et son efficacité dans la simplification de l'image.
- **Mean pooling** : Ce type de pooling calcule la moyenne des valeurs dans une région donnée. Il consiste à faire la somme de toutes les valeurs et à diviser par le nombre de valeurs, créant ainsi une valeur intermédiaire qui représente ce groupe de pixels.
- **Sum pooling** : Ce type de pooling est similaire au mean pooling, mais il ne divise pas la somme totale des valeurs par le nombre de valeurs, se contentant simplement de calculer cette somme.

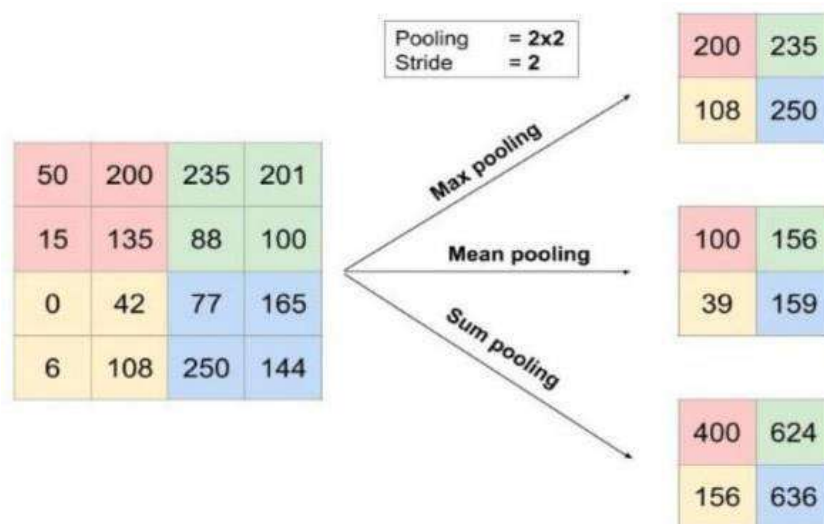


Figure 14 : Les différents types de Pooling avec un filtre 2x2 et un pas de 2.

### 6.1.4 Les fonctions d'activations

Pour augmenter l'efficacité d'un réseau de neurones convolutifs (CNN), il est courant d'ajouter des fonctions agissant comme des couches de correction entre les couches de traitement. Ces fonctions introduisent des non-linéarités qui permettent au réseau d'apprendre des systèmes complexes non linéaires. Plusieurs fonctions d'activation sont utilisées pour introduire cette non-linéarité dans différentes couches des CNN. Parmi les plus connues figurent le sigmoïde (logistique), la tangente hyperbolique et la fonction ReLU (Rectified Linear Unit).

- RELU :

Cette couche vise à améliorer l'efficacité du traitement en insérant une fonction mathématique, appelée fonction d'activation, entre les différentes couches de traitement. Parmi les fonctions d'activation couramment utilisées se trouve la fonction ReLU (Rectified Linear Unit), définie comme  $F(x) = \max(0, x)$ . Cette fonction contraint les neurones à produire des valeurs positives. En règle générale, les fonctions d'activation sont non linéaires, et leur rôle principal est de permettre aux réseaux de neurones d'apprendre des fonctions plus complexes que celles obtenues par une simple régression linéaire, car la multiplication des poids dans une couche cachée représente essentiellement une transformation linéaire [24].

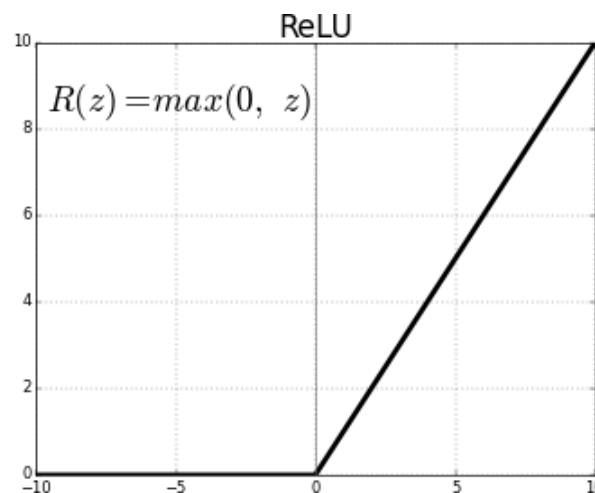


Figure 15 : Fonction d'activation de RELU.

La fonction Rectified Linear Unit (ReLU) est souvent appliquée après chaque opération de convolution dans un CNN, où toutes les valeurs de pixels négatives sont transformées en zéro. Son but est d'introduire de la non-linéarité dans le réseau, car la plupart des caractéristiques du monde réel, lorsqu'elles sont appliquées à l'une des cartes d'entrée, produisent une carte de sortie, également appelée carte de caractéristiques rectifiées [w33].

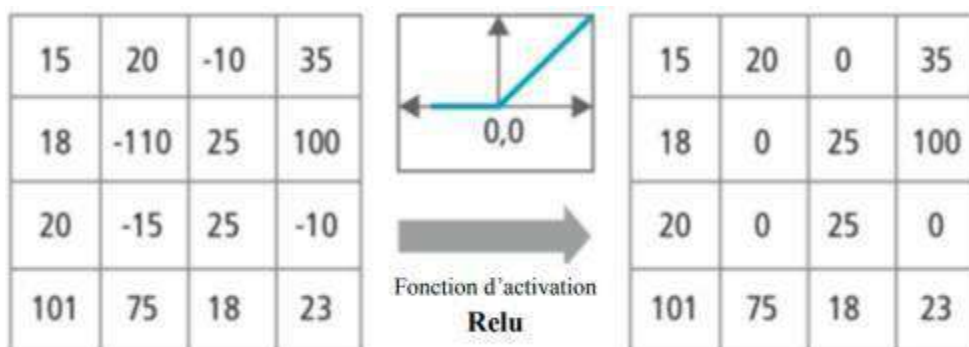


Figure 16 : Illustration du fonctionnement d'une couche ReLU. (Dans la case de gauche, tous les nombres négatifs ont été convertis en zéro après l'application de la fonction d'activation, tandis que toutes les autres valeurs sont restées inchangées).

### 6.1.5 Couche entièrement connectée (Fully Connected Layer (FC))

Après les étapes de convolution et de pooling, le raisonnement de haut niveau dans le réseau neuronal se réalise grâce à des couches entièrement connectées. Dans les réseaux de neurones convolutifs, chaque couche agit comme un filtre pour détecter la présence de caractéristiques spécifiques ou de motifs dans les données d'origine. Les premières couches identifient des caractéristiques facilement reconnaissables et interprétables, tandis que les couches ultérieures découvrent des caractéristiques de plus en plus abstraites. La dernière couche du réseau convolutif est capable d'effectuer une classification ultraprécise en combinant toutes les caractéristiques spécifiques détectées par les couches précédentes dans les données d'entrée.

Les couches totalement connectées accomplissent des tâches similaires à celles des réseaux de neurones artificiels standard (ANN) en produisant des scores de classe à partir des activations, qui sont ensuite utilisés pour la classification. Il est également recommandé d'utiliser la fonction ReLU entre ces couches pour améliorer les performances. L'objectif de la couche entièrement connectée est de classifier l'image d'entrée dans différentes catégories en fonction de l'ensemble de données d'apprentissage [20].

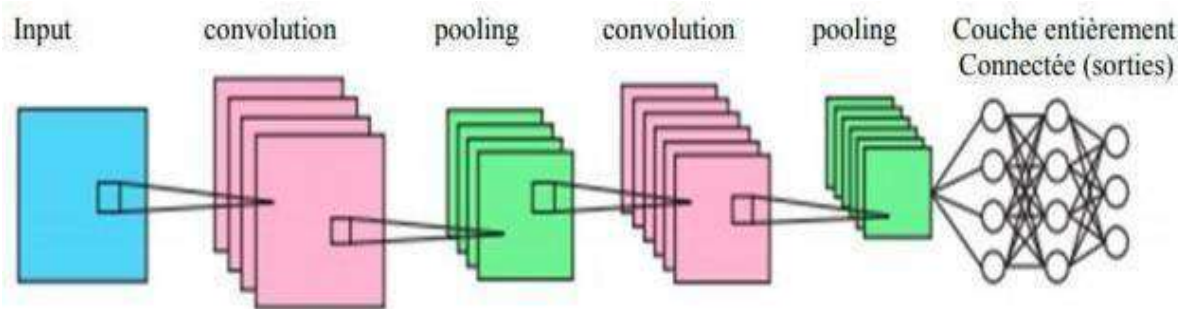


Figure 17 : Architecture standard d'un réseau convolutifs.

## 6.2 Outils d'optimisation des réseaux convolutifs

Les réseaux de neurones convolutifs (CNN) se distinguent des réseaux de neurones multicouches classiques en utilisant une variété de paramètres d'optimisation. Dans cette section, nous allons explorer différentes méthodes visant à améliorer l'efficacité de l'optimisation des CNN.

### 6.2.1 Le batch normalisation

Dans le contexte de l'apprentissage automatique, l'ordre dans lequel les échantillons sont présentés revêt une grande importance, car les paramètres du réseau sont mis à jour après chaque passage d'un échantillon d'apprentissage. Toutefois, il n'existe pas de méthodes efficaces pour prédire à l'avance quel échantillon apportera le plus d'informations sans effectuer une recherche exhaustive, ce qui est coûteux.[25] Une méthode simple consiste donc à mélanger l'ordre de présentation des échantillons après chaque passage de l'ensemble d'apprentissage. Ainsi, les échantillons successifs ne proviennent pas tous de la même classe, ce qui peut potentiellement apporter davantage d'informations [26].

Une technique plus récente, introduite en 2015[25], vise à accélérer et à améliorer l'apprentissage des CNN. Elle repose sur l'observation que, pendant l'apprentissage, la distribution des données d'entrée aux différentes couches du réseau change à chaque itération. Cette variation constante induit une adaptation continue des paramètres du CNN à ces différentes distributions, ce qui rallonge le temps nécessaire à l'apprentissage. La batch normalization (normalisation par lots) est une idée novatrice qui consiste à normaliser les données d'entrée de chaque couche de manière à ce que leurs distributions aient une moyenne nulle et une variance unitaire. Pendant l'apprentissage, les couches de batch normalization utilisent des paramètres (un facteur d'échelle et un biais) pour ajuster cette normalisation [27].

Ces paramètres permettent d'appliquer une transformation à la distribution normalisée. En d'autres termes, pendant l'apprentissage, si le réseau estime que la distribution normalisée n'est pas adaptée à une couche particulière, il apprend les paramètres nécessaires pour l'ajuster.

## 6.2.2 Les fonctions de perte

La fonction de perte, également appelée couche de perte, joue un rôle crucial en quantifiant l'écart entre les prédictions d'un réseau de neurones et les valeurs réelles. Elle guide le réseau en indiquant dans quelle mesure les prédictions s'éloignent de la vérité terrain. Diverses fonctions de perte existent, chacune étant adaptée à un type spécifique de tâche, comme la classification ou la régression. Le choix de la fonction de perte est déterminant pour l'entraînement efficace du réseau de neurones. Voici quelques-unes des fonctions de perte les plus fréquemment utilisées :

- **La fonction de perte Softmax** : Utilisée conjointement avec une couche Softmax, cette fonction transforme les scores logistiques générés par la couche dense en une distribution de probabilité en sortie. Contrairement à une distribution uniforme, la transformation est exponentielle, ce qui permet de mettre en évidence les différences entre les probabilités attribuées aux différentes classes. Par exemple, elle peut augmenter la probabilité d'une classe proche de 1 tout en réduisant celle d'une autre classe proche de 0.

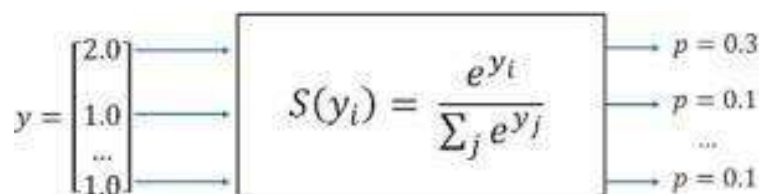


Figure 18 : La fonction de perte Softmax.

Ainsi, le modèle peut représenter la répartition de probabilité des expressions faciales...

- **La fonction de perte par entropie croisée (sigmoïde)** : Elle permet d'effectuer une régression en se basant sur des probabilités.
- **La perte quadratique (squared loss)** : Cette fonction mesure les écarts au carré entre

les valeurs prédites par un modèle pour un exemple donné et les valeurs réelles de l'étiquette. Une fonction connexe, l'erreur quadratique moyenne (Mean Squared Error, MSE), se calcule en divisant la somme des pertes quadratiques par le nombre total d'exemples.

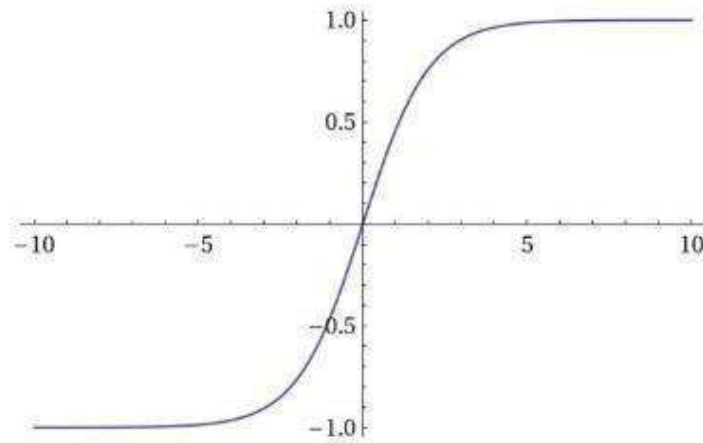


Figure 19 : Fonctions d'activation SoftMax.

### 6.2.3 Méthodes de régularisation

Le nombre d'entrées et de sorties d'un réseau est généralement déterminé par les données d'apprentissage. Cependant, le nombre total de neurones dans les couches intermédiaires, ainsi que le nombre de ces couches, sont des paramètres qui doivent être ajustés afin d'éviter à la fois le surapprentissage (overfitting) et le sous-apprentissage (underfitting) [28].

En Machine Learning, la régularisation est une stratégie qui vise à améliorer la capacité d'un algorithme à généraliser ses connaissances à de nouvelles données. En d'autres termes, elle cherche à réduire les erreurs lorsqu'on utilise le modèle avec des données qu'il n'a pas encore vues, tout en maintenant de bonnes performances sur les données d'entraînement. Pour ce faire, différentes techniques de régularisation peuvent être appliquées [28] [29] :

- **Dropout** : Pour éviter que le modèle ne surapprenne aux données d'entraînement et devienne trop spécialisé, le dropout est une méthode qui intervient pendant l'apprentissage. Elle consiste à aléatoirement désactiver certains neurones à chaque itération. Cela force le modèle à apprendre des caractéristiques plus générales au lieu de se focaliser uniquement sur les détails précis des données d'entraînement. Une fois l'apprentissage achevé, tous les neurones sont réactivés

pour permettre au modèle d'effectuer des prédictions précises sur de nouvelles données [w33].

- **Arrêt précoce (Early stopping) :** L'arrêt précoce désigne la méthode consistant à entraîner un réseau en utilisant à la fois des données d'entraînement et de test, mais en arrêtant l'entraînement dès que l'erreur sur l'ensemble de test recommence à augmenter.
- **Augmentation de données :** Cette méthode a pour objectif d'augmenter la taille de l'ensemble d'apprentissage en incorporant des données supplémentaires générées par des transformations telles que l'ajout de bruit, des modifications géométriques, etc., appliquées aux données initiales.
- **Régularisation L1 :** Cette forme de régularisation cible spécifiquement les entrées aléatoires et faibles en réduisant leur poids, tout en augmentant le poids des entrées jugées "importantes". Cela a pour effet de rendre le système moins sensible au bruit et aux variations mineures dans les données d'entrée [27].
- **Régularisation L2 (norme euclidienne) :** Cette technique de régularisation diminue l'impact des entrées les plus importantes, incitant ainsi le neurone à privilégier les entrées de moindre poids [26] [27].

### 6.3 Les architectures neuronales convolutifs

- **LeNet :** Les premières avancées des réseaux de neurones convolutifs ont été initiées dans les années 1990 par Yann LeCun. L'une des architectures les plus emblématiques de cette période est LeNet, qui s'est notamment illustrée dans la reconnaissance de codes postaux, de chiffres, et dans d'autres applications.
- **AlexNet :** AlexNet, élaboré par Alex Krizhevsky, Ilya Sutskever et Geoff Hinton, est le premier réseau convolutif à avoir popularisé l'utilisation de ces architectures dans le domaine de la vision par ordinateur. En 2012, lors du défi ImageNet, ce CNN a surpassé de loin ses concurrents. AlexNet présentait une architecture similaire à LeNet, mais était plus profond et plus large, avec des couches

convolutives empilées les unes sur les autres (contrairement à l'approche antérieure qui consistait à avoir une seule couche convolutive suivie immédiatement d'une couche de pooling).

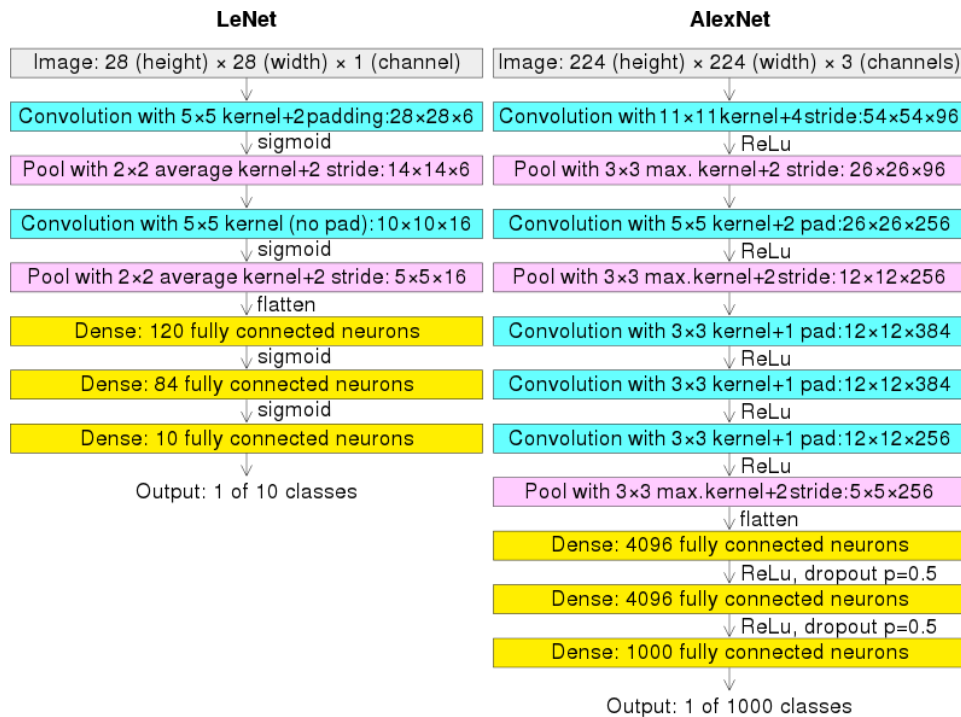


Figure 20 : LeNet et AlexNet Architectures.

- **ZFnet** : Cette architecture représente une évolution d'AlexNet, réalisée en ajustant les hyperparamètres, notamment en augmentant la taille des couches convolutionnelles tout en réduisant la taille du noyau pour la première couche.
- **GoogLeNet** : Ce modèle, créé par Google, a introduit le concept de module Inception, ce qui a considérablement réduit le nombre de paramètres dans le réseau (passant de 60 millions pour AlexNet à 4 millions). De plus, ce module utilise le "global average pooling", ce qui permet de supprimer de nombreux paramètres. Plusieurs versions de GoogLeNet existent, dont Inception-v4 et Xception.
- **ResNet** : Le réseau résiduel, créé par Kaiming He et son équipe, a remporté la compétition ILSVRC en 2015. Il se distingue par l'usage de sauts de connexion et une utilisation intensive de la normalisation par lots (batch normalization). De plus, il incorpore le "global average pooling" à la fin du réseau.
- **VGG Net** : Cette architecture a été développée en 2017 par le Visual Geometry

Group d'Oxford, sous la direction d'Andrea Vedaldi et Andrew Zisserman.

- **MobileNet** : Ces réseaux neuronaux convolutionnels profonds légers se démarquent par leur compacité et leurs performances élevées comparativement à de nombreux autres modèles populaires. Leur particularité réside dans l'utilisation de convolutions séparables en profondeur, où chaque canal de couleur est traité individuellement par une convolution distincte plutôt que d'être combiné en un seul [30].
- **MobileNetV2** : Une évolution de MobileNet, conserve les atouts de son prédécesseur tout en introduisant des blocs résiduels inversés dotés de fonctionnalités de goulot d'étranglement. Cette version réduite de MobileNetV2 comporte nettement moins de paramètres que l'originale. Ces modèles peuvent traiter des tailles d'image plus grandes que 32 x 32 avec des performances améliorées pour ces tailles d'images plus grandes [31].

## 7. Conclusion

Ce chapitre a exploré la conception d'un système de détection d'émotion pour améliorer le contrôle de satisfaction des utilisateurs. En utilisant des réseaux de neurones convolutifs pour analyser les expressions faciales, nous avons illustré comment cette technologie peut optimiser l'expérience utilisateur en fournissant des retours en temps réel. Cette approche promet de renforcer les interactions homme-machine en rendant les systèmes plus intuitifs et adaptatifs aux émotions des utilisateurs.

Dans le prochain chapitre nous allons présenter notre modèle basé sur les réseaux de neurones à convolutions Les (CNN), les outils, les langages et l'environnement de développement que nous avons exploités pour la réalisation de notre modèle, ainsi que la base de données, suivie bien sûr d'une comparaison et discussion avec notre application de détection et les explications de son interface.

# **Chapitre IV :**

# **Développement et**

# **tests**

## **1 Introduction**

Dans ce chapitre nous allons entamer la partie développement et tests du système de détection d'émotion utilisant la base de données FER 2013, téléchargée depuis Kaggle. Nous explorerons l'environnement de développement avec Python, Google Colab et Spyder, ainsi que les étapes de préparation des données et d'entraînement du modèle. Les résultats des tests seront présentés avec une discussion sur les performances du système, mettant en lumière ses forces et les pistes d'amélioration pour des applications futures. L'interface de notre programme est également présentée et expliquée.

## **2 Implémentation de système :**

### **2.1 Environnement de développement :**

#### **a) Python**

Python est un langage de programmation interprété, orienté objet et de haut niveau avec une sémantique dynamique développé par Guido van Rossum. Il a été initialement publié en 1991. Conçu pour être facile et amusant, le nom "Python" est un clin d'œil au groupe de comédie britannique Monty Python. Python a la réputation d'être un langage convivial pour les débutants, remplaçant Java en tant que langage d'introduction le plus largement utilisé, car il gère une grande partie de la complexité pour l'utilisateur, permettant aux débutants de se concentrer sur la compréhension complète des concepts de programmation plutôt que sur les moindres détails.

Python est utilisé pour le développement Web côté serveur, le développement de logiciels, les mathématiques et les scripts système, et est populaire pour le développement rapide d'applications et comme langage de script ou de collage pour lier les composants existants en raison de ses structures de données intégrées de haut niveau, typage dynamique et liaison dynamique. Les coûts de maintenance du programme sont réduits avec Python en raison de la syntaxe facile à apprendre et de l'accent mis sur la lisibilité. De plus, la prise en charge des modules et des packages par Python facilite les programmes modulaires et la réutilisation du code. Python est un langage communautaire open source, de sorte que de nombreux programmeurs indépendants créent continuellement des bibliothèques et des fonctionnalités pour celui-ci [W34].



Figure 21 : Logo de python.

**b) Google Colab :**

Google Colab est un outil gratuit de Google qui offre des ressources telles que des GPU, des TPU et des bibliothèques Python pour l'apprentissage automatique et l'IA. Il s'agit d'un environnement de notebooks Jupyter hébergé par Google, permettant un accès à des ressources informatiques puissantes sans nécessiter d'installation locale de logiciel. Les notebooks sont stockés dans Google Drive et peuvent être partagés avec d'autres utilisateurs. Google Colab est particulièrement populaire pour l'apprentissage automatique, l'analyse de données et l'éducation [w35].



Figure 22 : Logo de Google Colab.

**c) Spyder :**

Spyder est un environnement de développement intégré (IDE) puissant et gratuit pour le langage de programmation Python, spécialement conçu pour les scientifiques, les ingénieurs et les analystes de données. Il offre de nombreuses fonctionnalités utiles pour rendre le développement en Python plus efficace et agréable [w36].



Figure 23: Logo Spyder.

## 2.2 Bibliothèques utilisées :

- **Open CV** : Open CV est parmi les premières bibliothèques de traitement d'image avec une variété fonctions qui a beaucoup d'avantages sur la méthode de programmation [w37].
- **Scikit-learn** : est une bibliothèque libre pour Python destiné à l'apprentissage automatique (machine learning), elle comporte divers algorithmes de classification, de régression et de regroupement, comme les machines à vecteurs de support, des forêts aléatoires , elle est conçue pour interagir avec les bibliothèques numériques et scientifiques Python NumPy et SciPy [w38].
- **Numpy** : est une bibliothèque python utilisé pour le calcul numérique, la création des tableaux et les matrices, elle est conçue pour gérer les tableaux N-dimension. Numpy remplit les fonctions de manipulation des tableaux, nous pouvons facilement calculer leurs normes, les normalise, les concaténer et plusieurs d'autres fonctions [w39].
- **Keras** : Keras est un précieux outil open source, développé en Python, qui facilite l'interaction avec les puissants algorithmes de réseaux de neurones profonds et d'apprentissage automatique, notamment avec des frameworks tels que TensorFlow et Theano. À l'origine, cette bibliothèque a été créée par François Chollet pour simplifier ces tâches complexes [W40].

- **Pandas** : Pandas, une bibliothèque open source avec une licence BSD, offre des structures de données performantes et conviviales ainsi que des outils d'analyse de données pour le langage de programmation Python. Il est important de noter que le projet Pandas est soutenu par NumFOCUS, ce qui favorise son développement en tant que projet open source de renommée mondiale et permet d'apporter des contributions essentielles au projet [W41].
- **TensorFlow** : Le groupe Google Brain, composé de scientifiques et d'architectes, est à l'origine de TensorFlow, qui se distingue comme la bibliothèque de programmation la plus répandue dans le domaine de l'apprentissage profond. Ce qui a particulièrement contribué à la renommée de TensorFlow, c'est sa polyvalence en matière de création de modèles d'apprentissage profond, avec un support étendu pour différents langages tels que Python, C++ et R. À un niveau plus fondamental, TensorFlow peut être considéré comme une bibliothèque Python qui permet aux utilisateurs de définir des calculs complexes sous forme de graphique de flux de données [W42].

### 2.3 Implémentation de la Base de données FER2013 :

La base de données FER 2013[w43] qu'on a télécharger depuis Kaggle se compose d'images de visages en niveaux de gris de taille 48x48 pixels. Les visages de cette base de données ont été automatiquement enregistrés de manière à être plus ou moins centrés et à occuper à peu près la même quantité d'espace dans chaque image. La tâche associée à cette base de données est de catégoriser chaque visage en fonction de l'émotion exprimée dans l'expression faciale, en l'assignant à l'une des sept catégories suivantes : Colère, Dégoût, Peur, Joie, Tristesse, Surprise et Neutre.

L'ensemble d'apprentissage du jeu de données FER 2013 comprend 28 709 exemples, tandis que l'ensemble de test public en compte 3 589

La tâche consiste à classer chaque visage en fonction de l'émotion montrée dans l'expression faciale dans l'une des sept catégories (0 = en colère, 1 = dégoût, 2 = peur, 3 = heureux, 4 = triste, 5 = surprise, 6 = neutre).

Elle contient 35887 images expressions faciales.

L'ensemble d'apprentissage composé de 28 709 exemples et l'ensemble de test public se compose de 7 178 exemples.

<b>Emotion</b>	<b>Nombres</b>
<b>Colère</b>	<b>4953</b>
<b>Peur</b>	<b>5121</b>
<b>Tristesse</b>	<b>6077</b>
<b>Neutre</b>	<b>6198</b>
<b>Joie</b>	<b>8989</b>
<b>Surprise</b>	<b>4002</b>
<b>Dégoût</b>	<b>547</b>

Tableau 2 : FER2013 Dataset par émotion

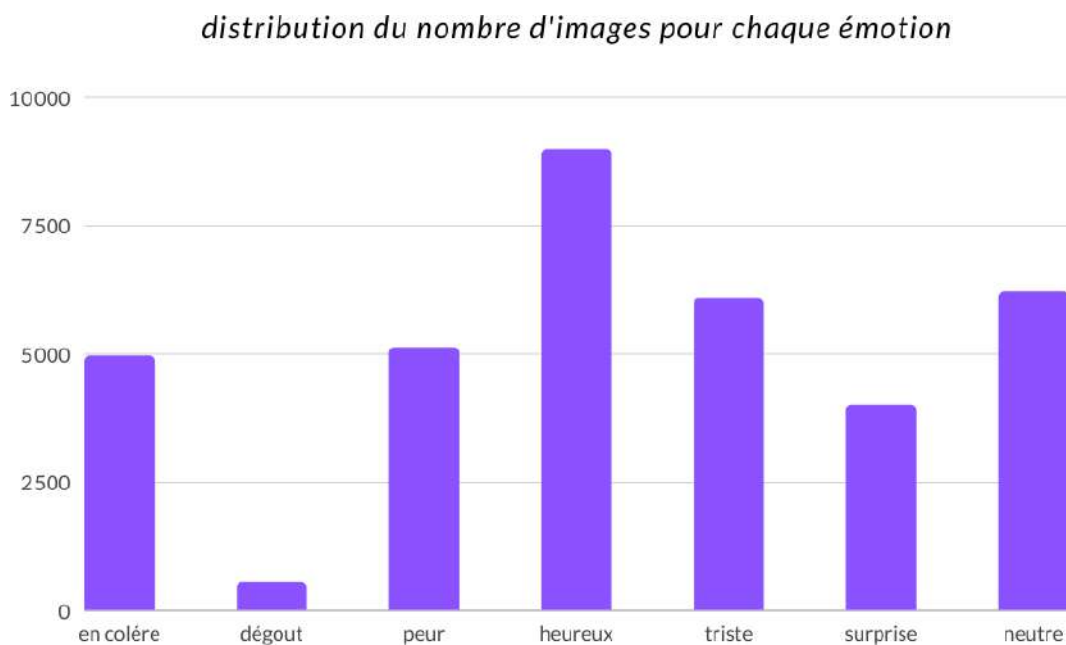


Figure 24 : Répartition de la base de données FER2013 par émotion.

Pour les valeurs d'usage on a 80% training, 10% validation et 10% pour le test.

Entrainement	Test Public	Test Privée
28709	3589	3589

Tableau 3: FER2013 dataset.

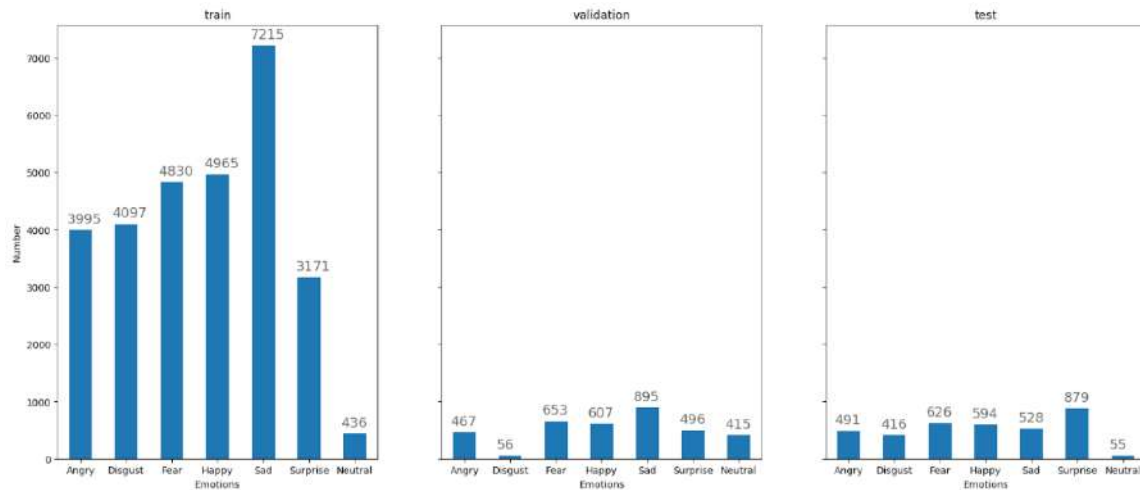


Figure 25 : Data set FER2013.

## 2.4 Implémentation et résultats des expériences

Cette partie représente les expériences effectuées sur notre système de reconnaissance des émotions faciales, en utilisant quatre modèles différents, Séquentiel, Fermodèle, Res-NET modèle et VGG modèle. L'objectif était d'évaluer la performance de l'information obtenue à partir de ces modèles.

- **Modèle Séquentiel**

Dans cette section, nous analyserons les résultats obtenus après avoir appliqué tous les modèles précédents sur le jeu de données FER2013. Pour notre modèle, nous avons atteint une perte de 0,9265 et une précision de 66-67 %. Voici le code :

```

model = Sequential()
#module 1
model.add(Conv2D(2*2*num_features, kernel_size=(3, 3), input_shape=(width, height, 1), data_format='channels_last'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(Conv2D(2*2*num_features, kernel_size=(3, 3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2, 2), strides=(2, 2)))
#module 2
model.add(Conv2D(2*num_features, kernel_size=(3, 3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(Conv2D(2*num_features, kernel_size=(3, 3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2, 2), strides=(2, 2)))
#module 3
model.add(Conv2D(num_features, kernel_size=(3, 3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(Conv2D(num_features, kernel_size=(3, 3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2, 2), strides=(2, 2)))
#flatten
model.add(Flatten())
#dense 1
model.add(Dense(2*2*2*num_features))
model.add(BatchNormalization())
model.add(Activation('relu'))
#dense 2
model.add(Dense(2*2*num_features))
model.add(BatchNormalization())
model.add(Activation('relu'))
#dense 3
model.add(Dense(2*num_features))
model.add(BatchNormalization())
model.add(Activation('relu'))
#output layer
model.add(Dense(num_classes, activation='softmax'))
model.compile(loss='categorical_crossentropy',
optimizer=Adam(lr=0.001, beta_1=0.9, beta_2=0.999, epsilon=1e-7),
metrics=['accuracy'])
import matplotlib.pyplot as plt
def plot_model_graph(model):
plt.figure(figsize=(10, 6))
tf.keras.utils.plot_model(model, to_file='model_graph.png', show_shapes=True)
plt.show()
plot_model_graph(model)
model.summary()

```

Figure 26 : code python

## ❖ Architectures de modèle séquentiel

Notre architecture finale (modèle séquentiel) affichait une précision sur les tests d'environ 66 à 67 %. Cette architecture est une combinaison des trois blocs suivants :

Total params	Trainable params	Non-trainable params
2,137,991	2,134,407	3,584

Tableau 5 : Taille des paramètres du modèle Séquentiel.

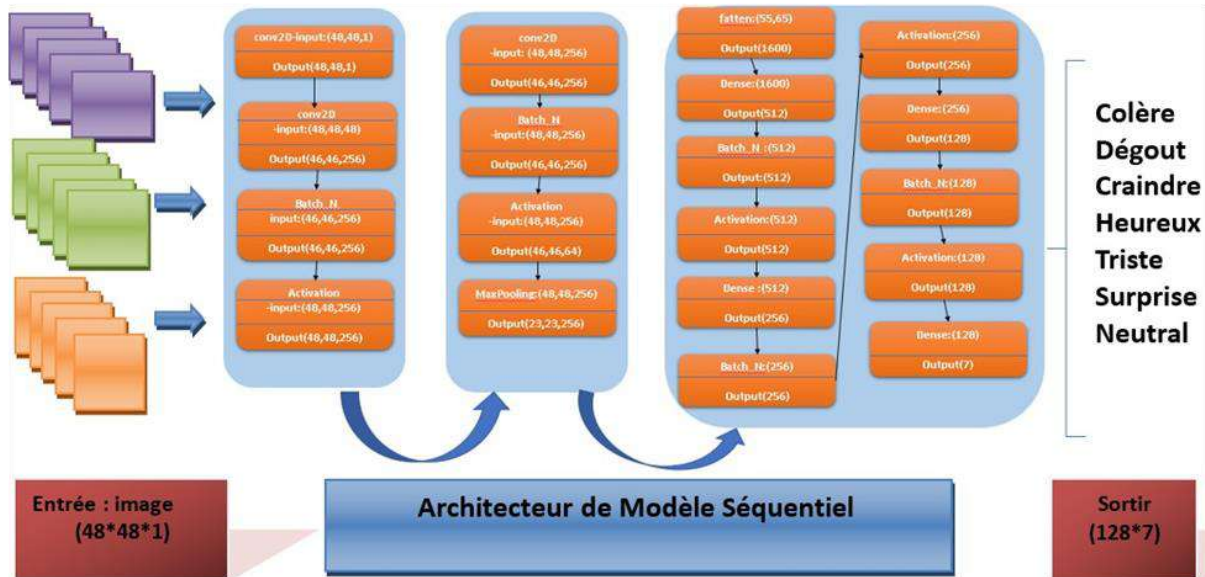


Figure 27 : Architectures de modèle séquentiel.

S'appuyer uniquement sur la précision et la perte d'un modèle entraîné ne fournit pas toujours une évaluation complète de ses performances.

❖ **Entraînement :**

Dans cette section, nous analysons les résultats obtenus avec notre approche développée. Une fois le modèle entraîné, nous allons examiner certains de ses paramètres :

Paramètres	Utilisation
<b>Epoch</b>	Désigne le nombre d'itération dans notre base de données.
<b>Loss</b>	Désigne le taux d'erreur.
<b>Accuracy</b>	Désigne le taux de précision.

Tableau 4 : les paramètres et ses utilisations.

❖ **Résultats expérimentaux et analyse de performance :**

- Résultats expérimentaux et analyse de performance :

Après 210/210 Epoch :

Epoch	Loss	Accuracy	Val_loss	Val_accuracy
210/210	1298	0.5024	0.8450	0.6601

Dans cette section, nous présentons les résultats de nos expérimentations sur la Tableau 6 : Résultats de Loss, Accuracy, Val\_Loss et Val\_Accuracy pour le modèle Séquentiel.

reconnaissance des émotions faciales. Nous allons visualiser la précision (accuracy) et le taux d'erreur (loss) du modèle séquentiel :

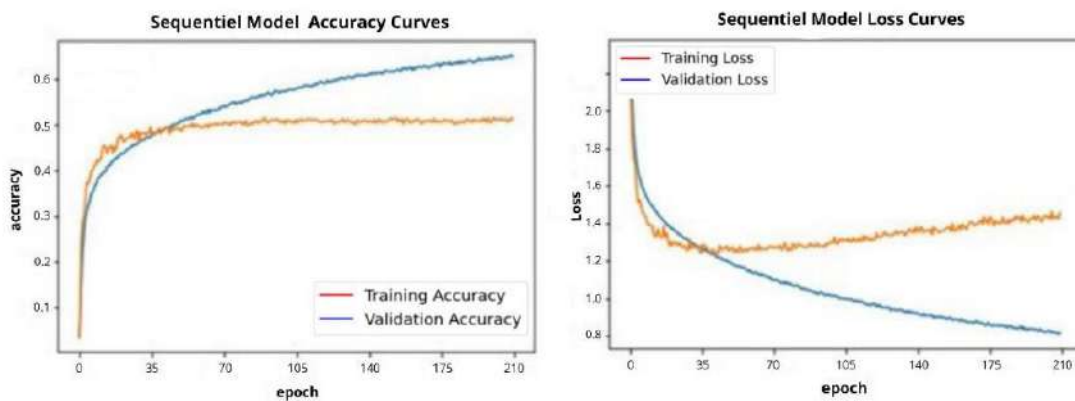


Figure 28 : Les courbes de précision et de perte pour le modèle Séquentiel.

Des métriques plus avancées sont à notre disposition, comme le score F1 que nous avons choisi d'adopter.

Le score F1 repose sur deux métriques préalablement calculées : la précision et le rappel. Ces mesures se basent sur les prédictions de vrais positifs, de faux positifs et de faux négatifs, ce qui est facilité par l'utilisation de la matrice de confusion.

Comme notre modèle est conçu pour reconnaître les 7 émotions faciales universelles, et que l'ensemble de données FER2013 incluait une 8e classe pour les émotions de « mépris », nous avons opté pour l'incorporation de tous les exemples de la classe « mépris » à la classe « joie » au lieu de les exclure.

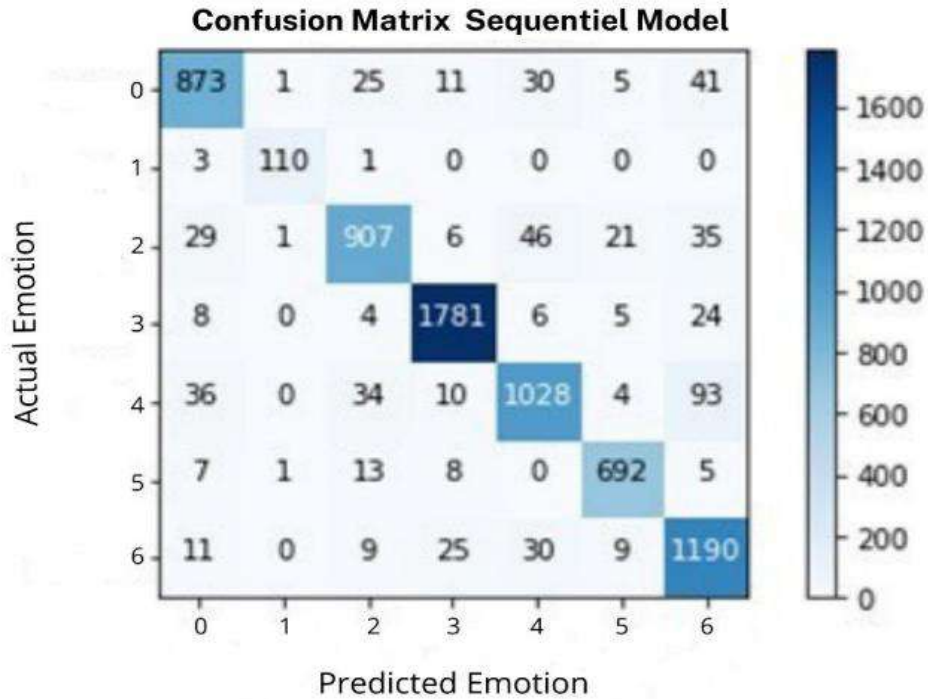


Figure 29 : Matrice de confusion pour le modèle Séquentiel.

❖ **Fer-modèle :**

Utilisons la même base de données “Fer 2013” mais avec le Modèle Fer-modèle.

Code source en python :

```
def FER_Model(input_shape=(48,48,1)):
    # first input model
    visible = Input(shape=input_shape, name='input')
    num_classes = 7
    #the 1-st block
    conv1_1 = Conv2D(64, kernel_size=3, activation='relu', padding='same', name = 'conv1_1')(visible)
    conv1_1 = BatchNormalization()(conv1_1)
    conv1_2 = Conv2D(64, kernel_size=3, activation='relu', padding='same', name = 'conv1_2')(conv1_1)
    conv1_2 = BatchNormalization()(conv1_2)
    pool1_1 = MaxPooling2D(pool_size=(2,2), name = 'pool1_1')(conv1_2)
    drop1_1 = Dropout(0.3, name = 'drop1_1')(pool1_1)
    #the 2-nd block
    conv2_1 = Conv2D(128, kernel_size=3, activation='relu', padding='same', name = 'conv2_1')(drop1_1)
    conv2_1 = BatchNormalization()(conv2_1)
    conv2_2 = Conv2D(128, kernel_size=3, activation='relu', padding='same', name = 'conv2_2')(conv2_1)
    conv2_2 = BatchNormalization()(conv2_2)
    conv2_3 = Conv2D(128, kernel_size=3, activation='relu', padding='same', name = 'conv2_3')(conv2_2)
    conv2_3 = BatchNormalization()(conv2_3)
    pool2_1 = MaxPooling2D(pool_size=(2,2), name = 'pool2_1')(conv2_3)
    drop2_1 = Dropout(0.3, name = 'drop2_1')(pool2_1)
    #the 3-rd block
    conv3_1 = Conv2D(256, kernel_size=3, activation='relu', padding='same', name = 'conv3_1')(drop2_1)
    conv3_1 = BatchNormalization()(conv3_1)
    conv3_2 = Conv2D(256, kernel_size=3, activation='relu', padding='same', name = 'conv3_2')(conv3_1)
    conv3_2 = BatchNormalization()(conv3_2)
    conv3_3 = Conv2D(256, kernel_size=3, activation='relu', padding='same', name = 'conv3_3')(conv3_2)
    conv3_3 = BatchNormalization()(conv3_3)
    conv3_4 = Conv2D(256, kernel_size=3, activation='relu', padding='same', name = 'conv3_4')(conv3_3)
    conv3_4 = BatchNormalization()(conv3_4)
    pool3_1 = MaxPooling2D(pool_size=(2,2), name = 'pool3_1')(conv3_4)
    drop3_1 = Dropout(0.3, name = 'drop3_1')(pool3_1)
```

Figure 30 : code python

```

#the 4-th block
conv4_1 = Conv2D(256, kernel_size=3, activation='relu', padding='same', name = 'conv4_1')(drop3_1)
conv4_1 = BatchNormalization()(conv4_1)
conv4_2 = Conv2D(256, kernel_size=3, activation='relu', padding='same', name = 'conv4_2')(conv4_1)
conv4_2 = BatchNormalization()(conv4_2)
conv4_3 = Conv2D(256, kernel_size=3, activation='relu', padding='same', name = 'conv4_3')(conv4_2)
conv4_3 = BatchNormalization()(conv4_3)
conv4_4 = Conv2D(256, kernel_size=3, activation='relu', padding='same', name = 'conv4_4')(conv4_3)
conv4_4 = BatchNormalization()(conv4_4)
pool4_1 = MaxPooling2D(pool_size=(2,2), name = 'pool4_1')(conv4_4)
drop4_1 = Dropout(0.3, name = 'drop4_1')(pool4_1)
#the 5-th block
conv5_1 = Conv2D(512, kernel_size=3, activation='relu', padding='same', name = 'conv5_1')(drop4_1)
conv5_1 = BatchNormalization()(conv5_1)
conv5_2 = Conv2D(512, kernel_size=3, activation='relu', padding='same', name = 'conv5_2')(conv5_1)
conv5_2 = BatchNormalization()(conv5_2)
conv5_3 = Conv2D(512, kernel_size=3, activation='relu', padding='same', name = 'conv5_3')(conv5_2)
conv5_3 = BatchNormalization()(conv5_3)
conv5_4 = Conv2D(512, kernel_size=3, activation='relu', padding='same', name = 'conv5_4')(conv5_3)
conv5_4 = BatchNormalization()(conv5_4)
pool5_1 = MaxPooling2D(pool_size=(2,2), name = 'pool5_1')(conv5_4)
drop5_1 = Dropout(0.3, name = 'drop5_1')(pool5_1)#Flatten and output
flatten = Flatten(name = 'flatten')(drop5_1)
ouput = Dense(num_classes, activation='softmax', name = 'output')(flatten)# create model
model = Model(inputs = visible, outputs = ouput)
# summary layers
print(model.summary())
return model
model = FER_Model()

```

Figure 31: code python

❖ Architecture du Fer-Modèle

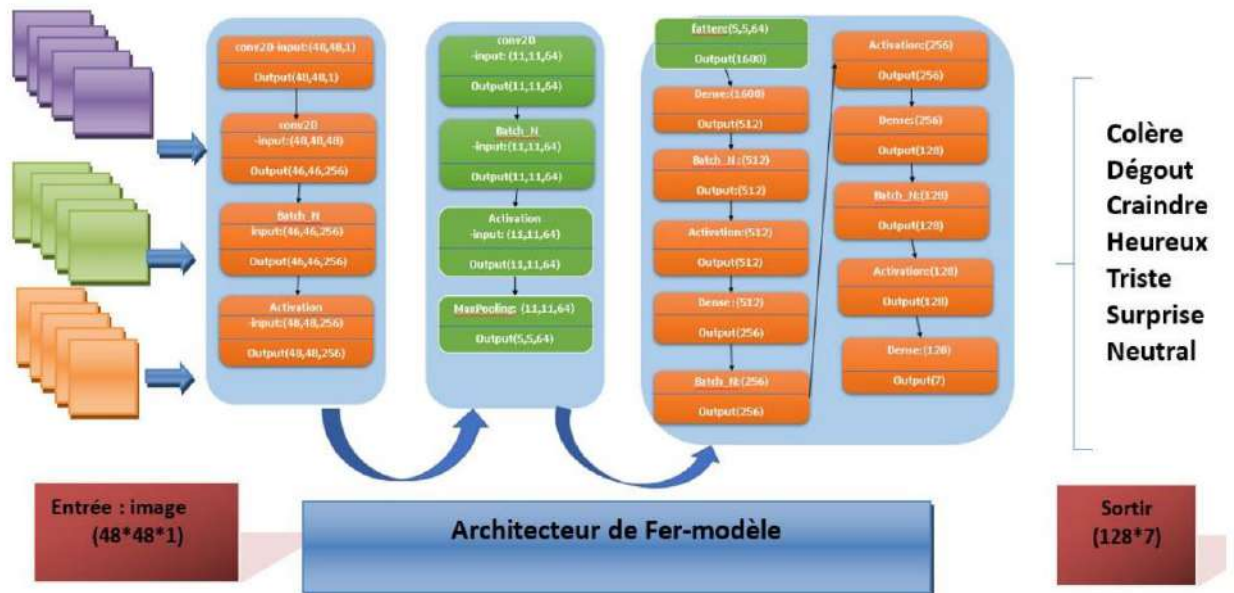


Figure 32 : Architecture de Fer-modèle.

❖ Résultats expérimentaux et analyse de performance

Cependant, dépendre uniquement de la précision (Accuracy) et de la perte (Loss) du

modèle entraîné ne donne pas toujours une compréhension complète des performances du modèle.

Total params	Trainable params	Non-Trainable params
13,111,367	13,103,431	7,936

Tableau 7 : Taille des paramètres du modèle FER.

Après avoir compilé le modèle, nous ajustons ensuite les données pour l'entraînement et la validation. Ici, nous utilisons un batch size de 64 pour 210 époques (epochs), soit un total de 210/210 epochs :

Epoch	Loss	Accuracy	Val_loss	Val_accuracy
210/210	0.8281	0.7296	0.8450	0.6980

Tableau 8 : Résultats de Loss, Accuracy, Val\_Loss et Val\_Accuracy pour le modèle FER

Dans cette section, nous présentons les résultats expérimentaux de notre système pour la reconnaissance des émotions faciales. Tracer la précision (accuracy) et représenter le taux d'erreur (loss) du modèle Fer-modèle :

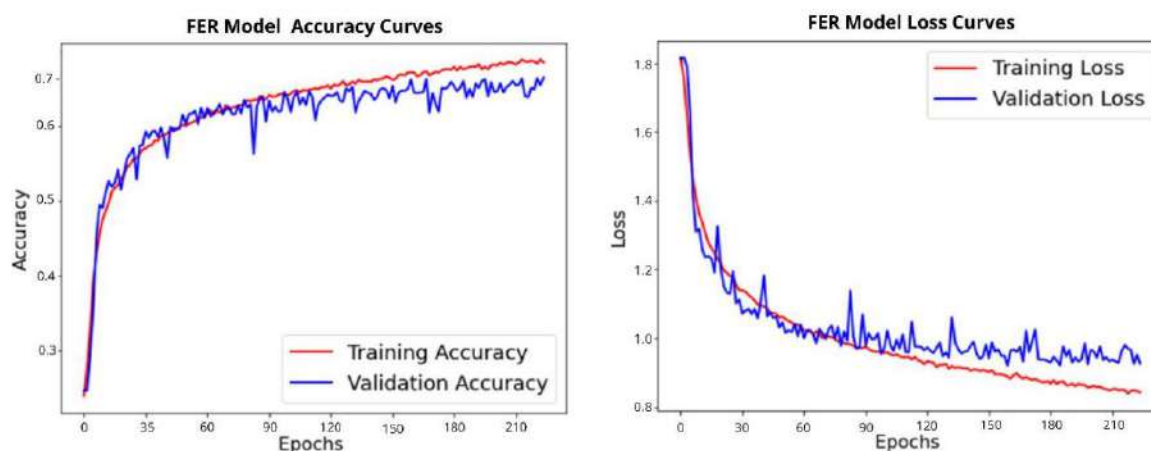


Figure 33 : Les courbes de précision et de perte pour le FER-Modèle.

Il existe des mesures plus avancées telles que le score F1, que nous avons choisi d'utiliser. Le score F1 est calculé à partir de deux métriques préalablement calculées :

La précision et le rappel. Ce sont deux mesures importantes qui utilisent les exemples prédits vrais positifs, faux positifs et faux négatifs. Ces données sont mieux visualisées à l'aide de la matrice de confusion.

Comme nous avons développé notre modèle pour identifier les 7 émotions faciales universelles et que le jeu de données FER2013 incluait une 8ème classe pour les émotions de "contempt", nous avons opté pour l'agrégation de tous les exemples de la classe "contempt" à la classe "Happy", plutôt que de supprimer ces données.

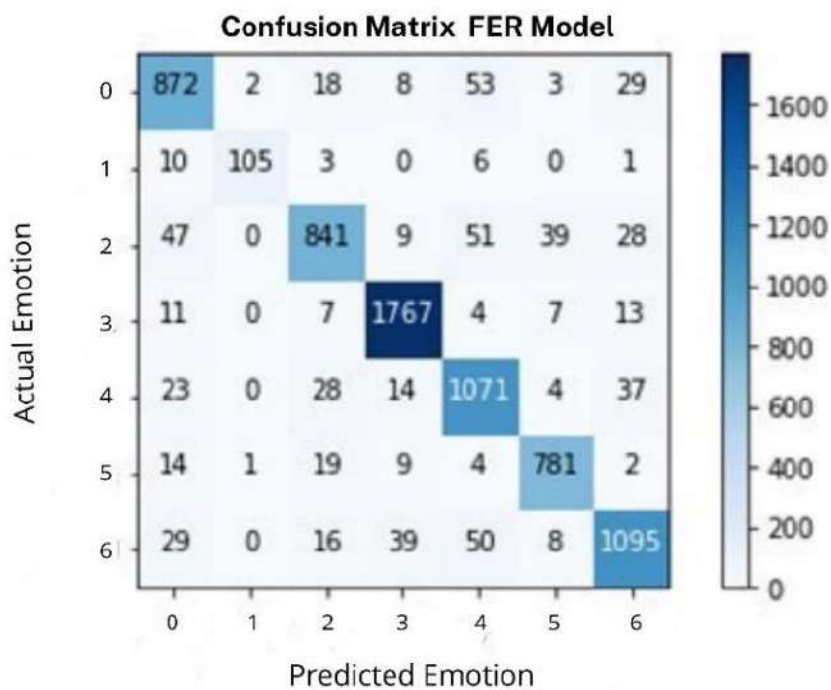


Figure 34 : Matrice de confusion pour le FER-Modèle.

- **RES-Net : (Residual Network)**

ResNet est un type d'architecture de réseau neuronal convolutif qui a été développé par Microsoft Research et présenté lors de la compétition ImageNet en 2015.

Son but était de surmonter le défi de la baisse des performances lorsque la profondeur du réseau augmentait.

L'innovation centrale des modèles ResNet réside dans l'introduction de blocs résiduels, une approche qui permet au réseau d'apprendre des fonctions résiduelles plutôt que des fonctions directes.

Les blocs résiduels incorporent des connexions "skip" pour créer des raccourcis qui facilitent la propagation du gradient à travers le réseau.

Les modèles ResNet sont célèbres pour leur grande profondeur, atteignant par exemple ResNet-152 avec 152 couches.

Le modèle ResNet a démontré des performances exceptionnelles sur divers ensembles de données et a été largement adopté dans le domaine de la vision par ordinateur.

❖ **Architecture global RES-NET**

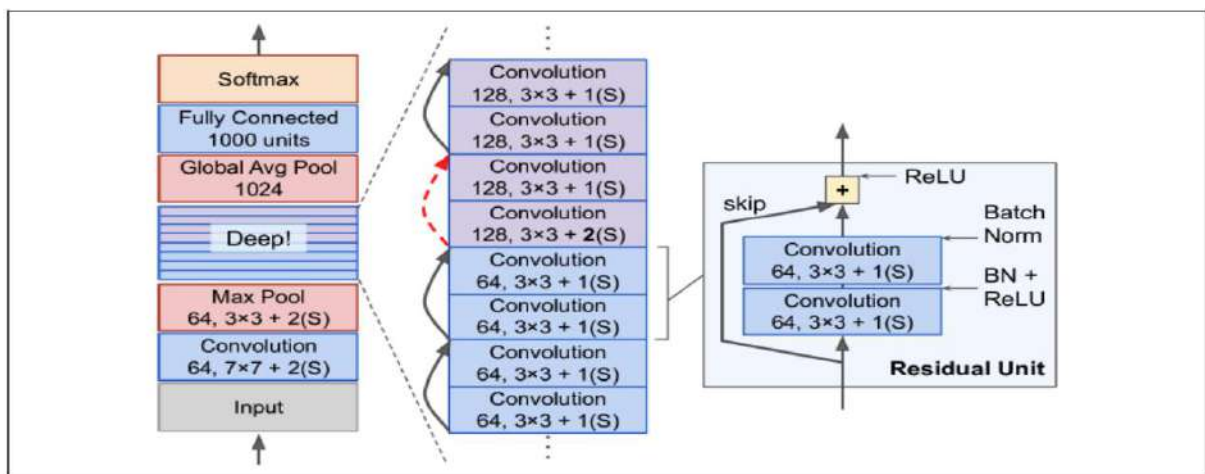


Figure 35: Architecteur Global de Modèle Res-Net [W44].

❖ **Résultats expérimentaux et analyse de performance**

**Total de « params »**

Total params	Trainable params	Non-Trainable params
21,303,943	21,288,71	15,232

Tableau 9 : Taille des paramètres du modèle Res-Net.

Après 210/210 Epoch :

Epoch	Loss	Accuracy	Val_loss	Val_accuracy
210/210	1.2103	0.4877	0.9265	0.6311

Tableau 10 : Résultats de Loss, Accuracy, Val\_Loss et Val\_Accuracy pour le modèle Res-Net.

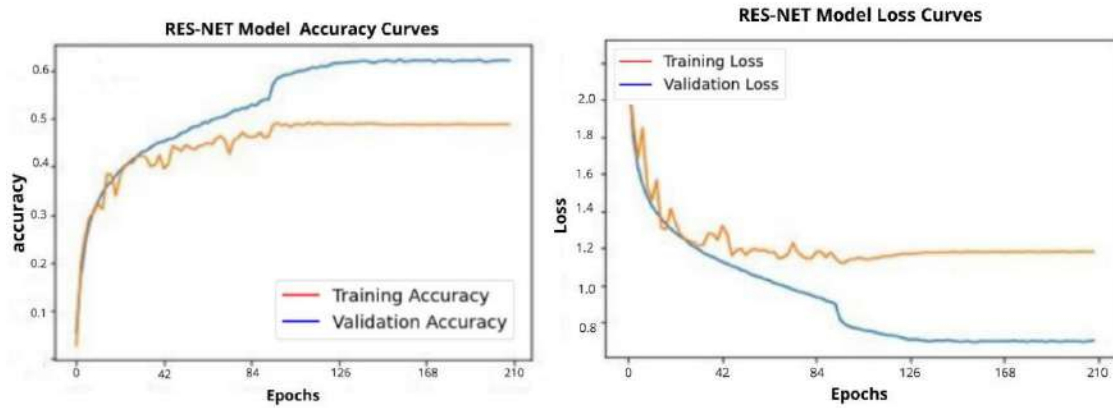


Figure 36 : Les courbes de précision et de perte pour le modèle RES-NET.

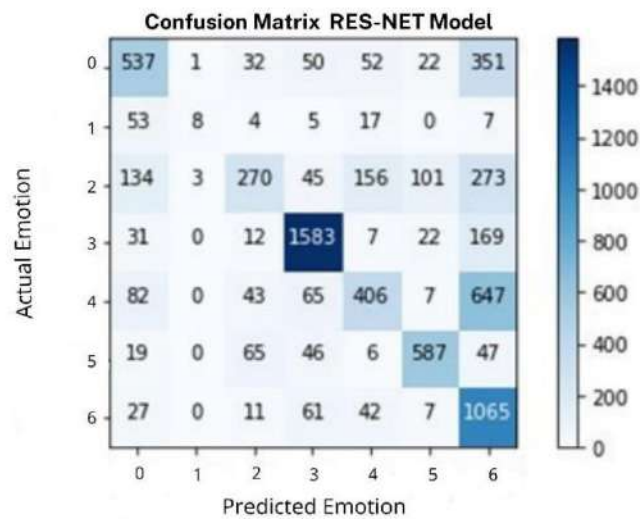


Figure 37 : Matrice de confusion pour le modèle RES-NET.

- **VGG (Visual Geometry Group)**

Le modèle VGG désigne une architecture de réseau de neurones convolutifs (CNN) qui a été développée par le groupe de Visual Geometry à l'Université d'Oxford.

Ses couches de convolutions sont extrêmement profondes, et sa structure se distingue par sa simplicité et son uniformité.

La conception de VGG se concentre principalement sur l'utilisation de couches de convolutions dotées de filtres de petite taille (généralement 3x3) ainsi que sur l'incorporation de couches de MaxPooling pour réduire les dimensions spatiales des données en entrée.

Le modèle VGG est célèbre pour sa structure à profondeur variable, qui se décline en différentes variantes : VGG16 possède 16 couches (composées de 13 couches de convolution et 3 couches entièrement connectées), tandis que VGG19 en compte 19 (avec 16 couches de convolution et 3 couches entièrement connectées).

Le modèle VGG est réputé pour son potentiel d'apprentissage élevé, mais il présente également un inconvénient en termes de coût computationnel et de mémoire, attribuable à sa profondeur importante.

### ❖ Architecture VGG

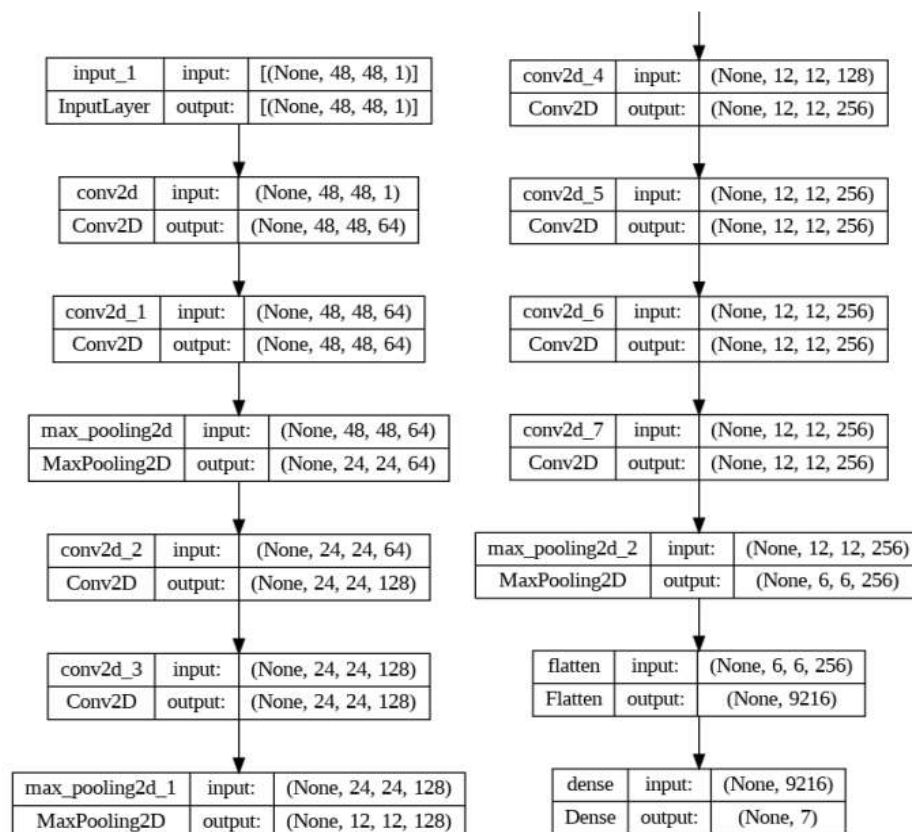


Figure 38 : Les composants de L'architecture VGG Modèle.

### ❖ Résultats expérimentaux et analyse de performance

Total params	Trainable params	Non-Trainable params
2,388,935	2,388,935	0

Tableau 11 : Taille des paramètres du modèle VGG.

Epoch	Loss	Accuracy	Val_loss	Val_accuracy
210/210	0.8292	0.6632	1.0698	0.5961

Tableau 12 : Résultats de Loss, Accuracy, Val\_Loss et Val\_Accuracy pour le modèle VGG.

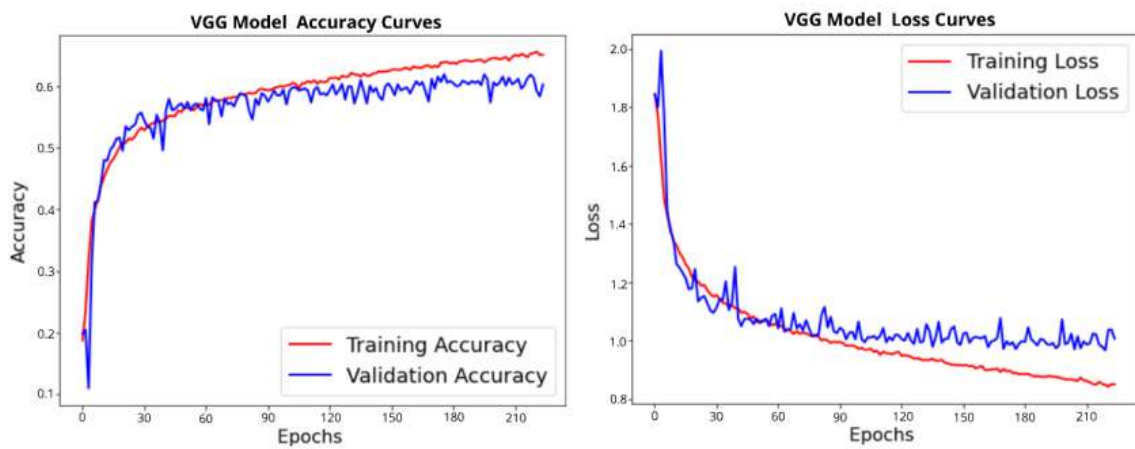


Figure 39 : Les courbes de précision et de perte pour le modèle VGG.

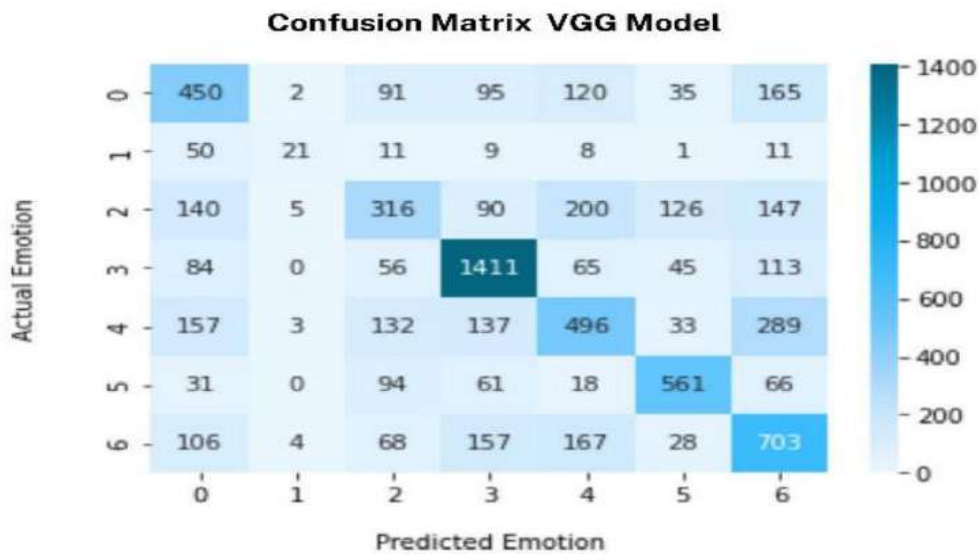


Figure 40 : Matrice de confusion pour le modèle VGG.

- **CNN-SVM**

Dans ce modèle, combiné CNN avec SVM la même architecture du modèle FER 2013

- ❖ **Résultats expérimentaux et analyse de performance**

Dans la dernière couche dense de sortie, un régularisateur L2 de 0,01 a été ajouté, ainsi qu'une fonction de perte de charnière au carré.

Epoch	Loss	Accuracy	Val_loss	Val_accuracy
<b>140/140</b>	<b>1.0453</b>	<b>0,6421</b>	<b>1.0582</b>	<b>0.6244</b>

Tableau 13 : Résultats de Loss, Accuracy, Val\_Loss et Val\_Accuracy pour le modèle CNNSVM.

Avec SVM, le résultat obtenu était de **1,0582** de perte et de **62,44 %** de précision

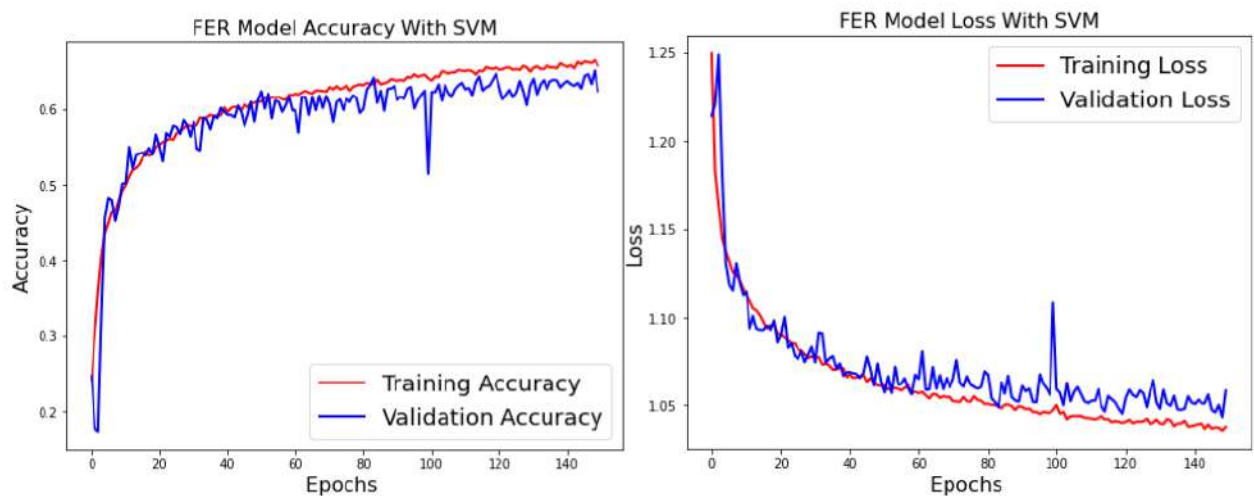


Figure 41 : The accuracy and loss curves for the FER model with SVM.

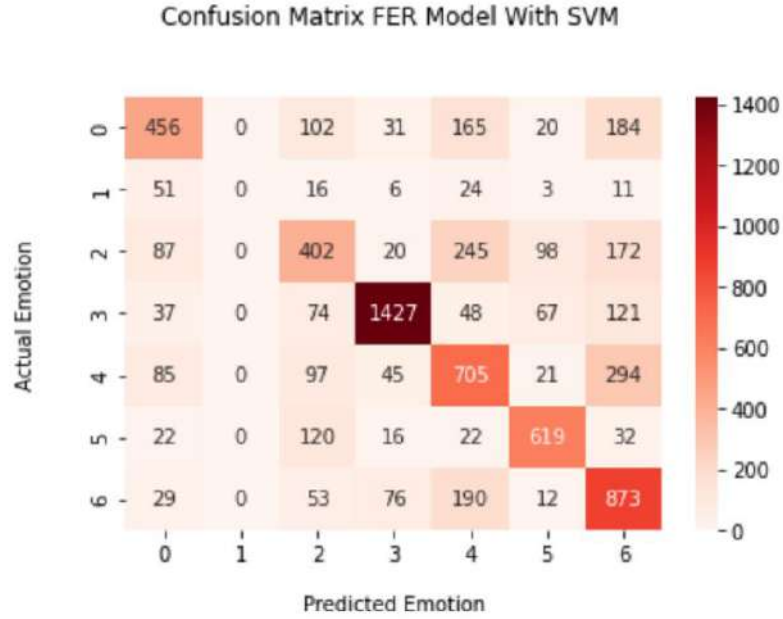


Figure 42: Matrice de confusion pour le modèle FER CNN avec SVM.

ID	Face Emotion	Precision	Recall	F1 Score
0	Angry	0.59452412	0.47599165	0.52869565
1	Disgust	0	0	0
2	Fear	0.46527778	0.39257812	0.42584746
3	Happy	0.88032079	0.80439684	0.84064801
4	Sad	0.50393138	0.56535686	0.53287982
5	Surprise	0.73690476	0.74488568	0.74087373
6	Neutral	0.51748666	0.7080292	0.59794521

Tableau 14 : Résultats de précision, de rappel et de score F1 pour le modèle FER avec SVM.

### 3 La comparaison

Dans cette section spécifique, nous allons comparer nos divers modèles FER dans le [Tableau 15].

<b>Model</b>	<b>Accuracy %</b>	<b>Loss</b>
<b>VGG</b>	<b>59.61%</b>	<b>1.0698</b>
<b>CNN-SVM</b>	<b>62.44%</b>	<b>1.0582</b>
<b>RES-Net</b>	<b>63.11%</b>	<b>0.9265</b>
<b>Séquentiel</b>	<b>66.01%</b>	<b>0.8967</b>
<b>Fer-modèle</b>	<b>69.80%</b>	<b>0.8450</b>

Tableau 15 : La différence de précision et les résultats de perte des modèles proposés.

En comparant ces modèles, voici quelques observations notables :

- Le Fer-modèle se distingue par sa précision supérieure parmi tous les modèles, atteignant 69,80 %. Cela suggère que ce modèle excelle dans la classification de la tâche spécifique...
- Le modèle Séquentiel se distingue également par sa performance élevée, affichant une précision de 66,01 %, ce qui le place comme le deuxième meilleur modèle de la liste.
- RES-Net suit de près avec une précision de 63,11 %. Bien qu'il soit légèrement moins performant que le modèle Séquentiel, il reste solide dans ses résultats.
- Les modèles CNN-SVM et VGG présentent des précisions légèrement inférieures, avec respectivement 62,44 % et 59,61 %. Cela les positionne en bas de la liste en termes de précision.

En ce qui concerne la perte, le modèle Fer-modèle affiche également la perte la plus faible, ce qui indique une meilleure capacité à minimiser les erreurs de classification.

## 4 Discussion

En résumé, le modèle Fer-modèle se démarque en tant que leader en termes de précision et de perte, suivi du modèle séquentiel et de RES-Net, tandis que les modèles CNN-SVM et VGG affichent des performances légèrement inférieures. Le choix du modèle à utiliser dépendra de l'importance que vous accordez à la précision et à la gestion de la perte dans votre application spécifique.

## 5 Le prototype

Description du programme :

L'interface graphique (GUI) de l'application de détection d'émotions en direct est conçue pour offrir une interface conviviale permettant de détecter en temps réel les émotions à travers un flux vidéo de la caméra. Voici comment elle fonctionne lorsqu'un utilisateur interagit avec :



Figure 43 : Page d'accueil.

L'interface de l'application de détection d'émotions en direct comprend deux boutons principaux :

**Mode Offline** : Détection des émotions à partir de vidéos enregistrées sur l'ordinateur, idéal pour l'analyse rétrospective.

**Mode Online** : Détection des émotions faciales en temps réel, utile pour la surveillance de sécurité et l'amélioration du service client.

## Mode Offline :

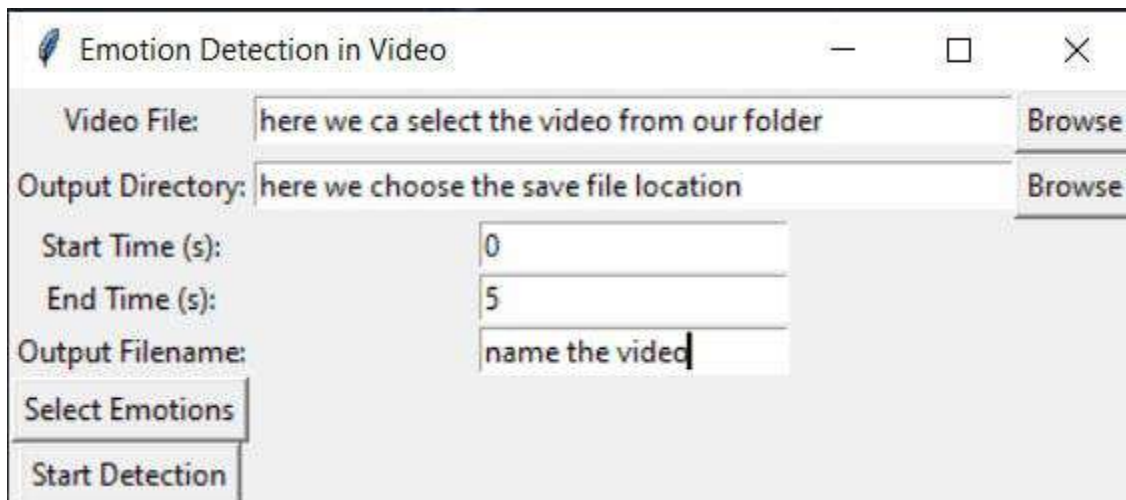


Figure 44 : interface offline.

### Explication de la partie offline :

1. **Video file** : Label "video File : " avec zone d'entrée et bouton "browse".

Sélectionne un fichier vidéo (ex : .mp4, .avi, .mov).

2. **Output Directory** : Label "output directory : " avec zone d'entrée et bouton "Browse" pour choisir le dossier ou faire la sauvegarde de la vidéo traitée.
3. **Start Time / End Time** : Choisissez le moment précis de début et de fin de la détection dans la vidéo à traiter.
4. **Output Filename** : Nomme le fichier vidéo de sortie.
5. **Sélection des Émotions** : Bouton "Sélectionner les Émotions" pour choisir les émotions à mettre en évidence.

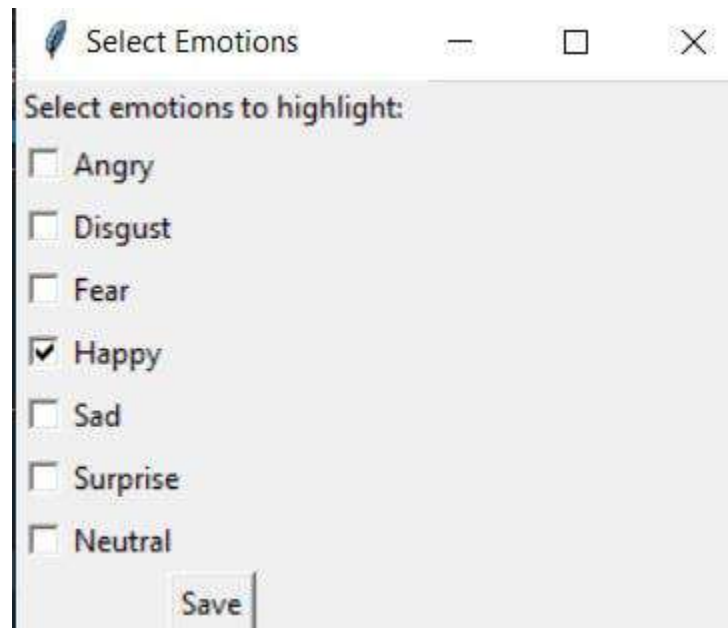


Figure 45 : Sélectionner l'émotion.

**Bouton Start detection :**

Bouton "Start detection" pour lancer le traitement de la vidéo.

Affiche une barre de progression et "Traitement..." pendant le processus.

**Barre de Progression :**

Montre l'avancement du traitement vidéo en temps réel.

**Terminé :**

Ferme automatiquement la fenêtre de progression après le traitement.

La vidéo de sortie, avec émotions mises en évidence, est enregistrée dans le répertoire choisi.

Cette interface simplifie la détection d'émotions dans les vidéos, permettant de choisir le fichier, configurer les paramètres et suivre facilement le processus de détection.

### Mode Online (live) :

Détection en temps réel des émotions voici l'interface avec les explications des boutons

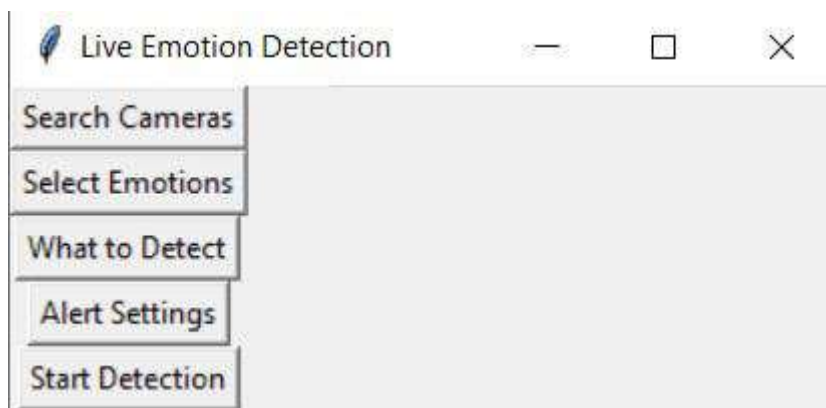


Figure 46 : interface online.

### Bouton Search Cameras:

Bouton "Search Cameras" : ouvre une fenêtre pour sélectionner les caméras disponibles avec cases à cocher, bouton de rafraîchissement et de sauvegarde.

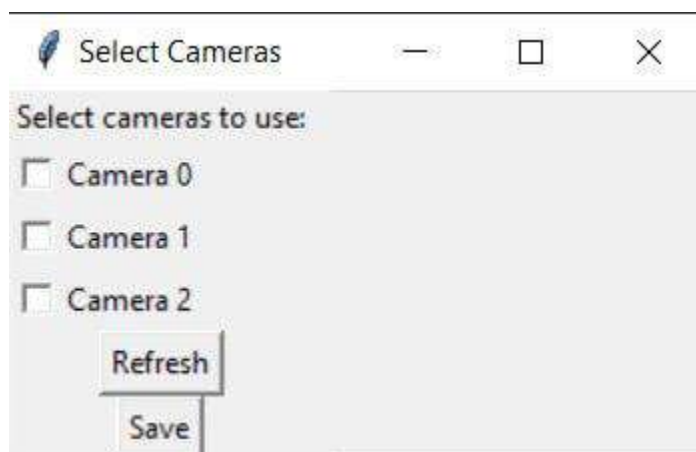


Figure 47 : interface sélection cameras.

### Bouton Select Emotions :

Bouton "Select Emotions" : ouvre une fenêtre pour choisir les émotions à détecter avec cases à cocher et bouton de sauvegarde.

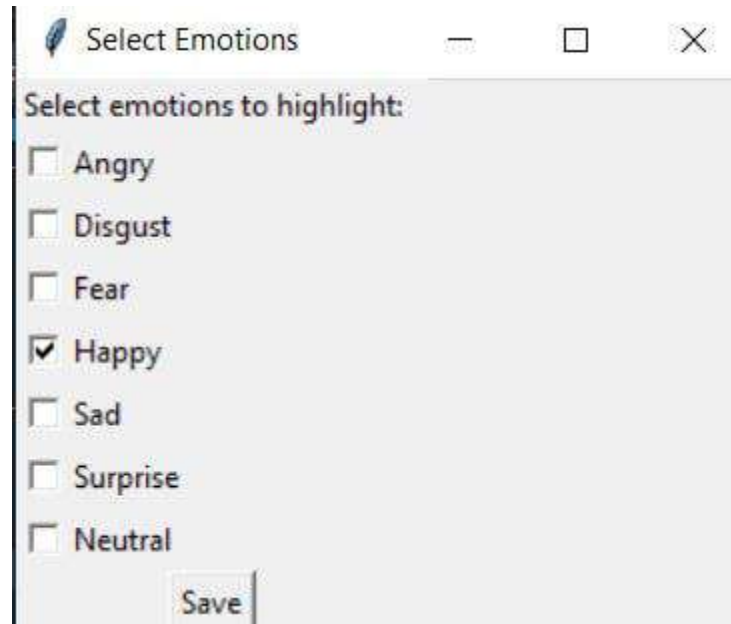


Figure 48 : interface sélectionner l'émotion.

### Bouton What to detect:

Le Bouton "What to detect " : ouvre une fenêtre pour configurer les paramètres de détection, afficher un cadre autour des visages détectés.

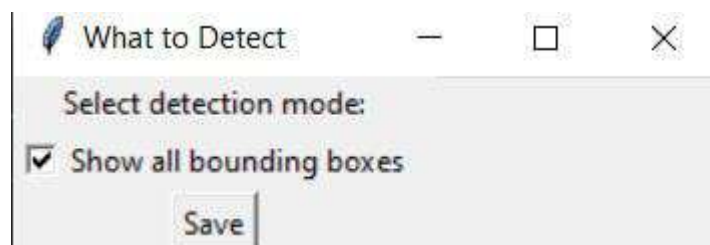


Figure 49 : What to detect.

### **Bouton Alert Settings:**

Bouton "Alert Settings" : configure les alertes, la durée d'affichage et le délai de récupération.

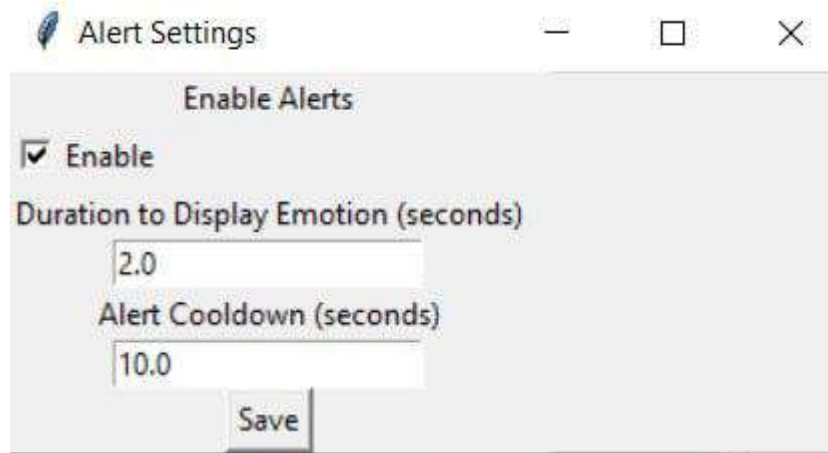


Figure 50 : paramètres des alertes.

### **Bouton Start detection :**

Bouton "Start detection" : lance la détection avec les caméras et paramètres choisis.

Affiche un message d'erreur si aucune caméra n'est sélectionnée.

### **Affichage du Flux Vidéo :**

Cadre pour afficher le flux vidéo en direct.

### **Start detection :**

Lance la détection et affiche le flux vidéo en direct avec les émotions détectées.

Voici un exemple de la détection, ici on a sélectionné l'émotion « Happy » et on a lancer la détection et voilà le résultat :

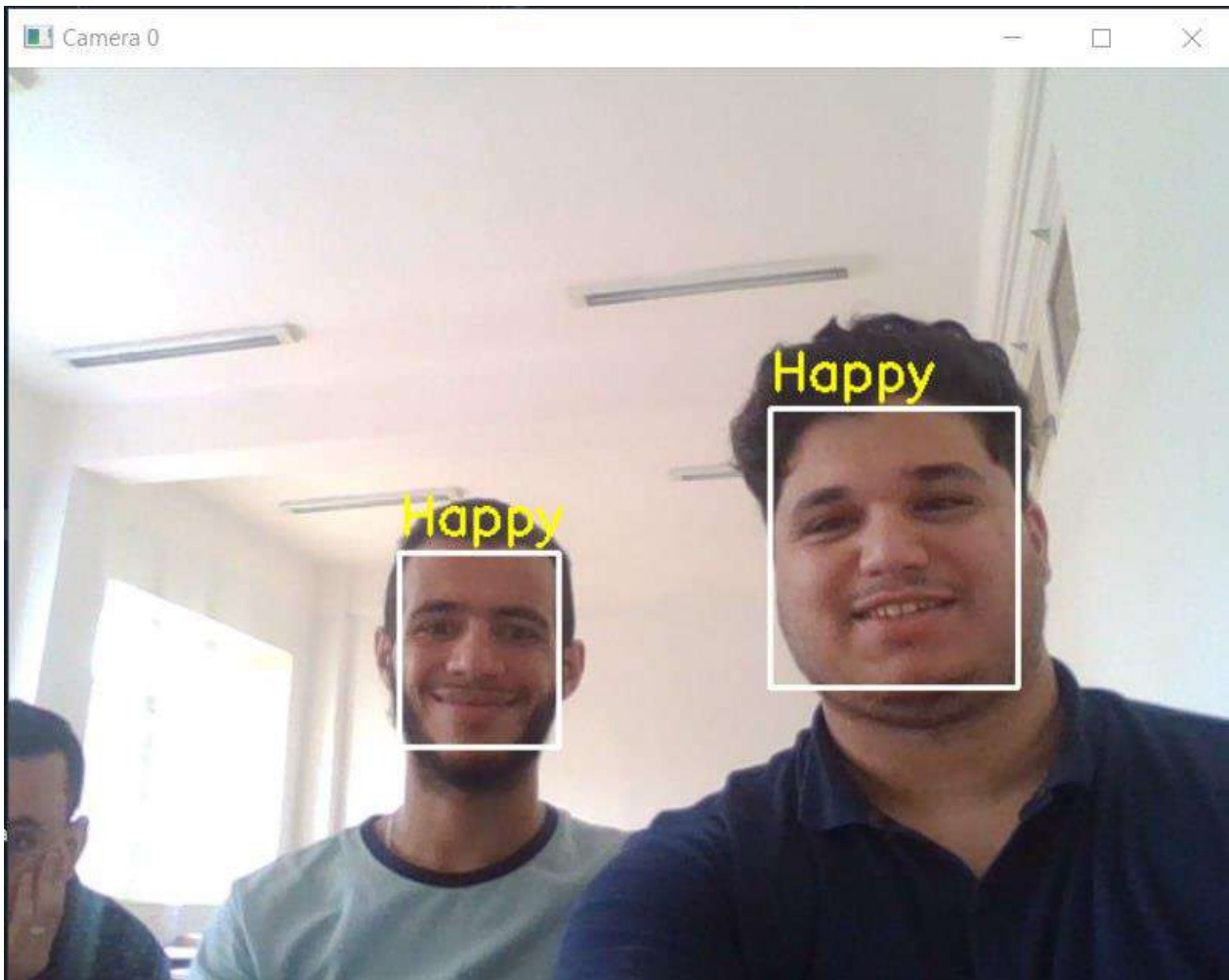


Figure 51: Résultat de la détection.

## **6 Conclusion**

Nous avons présenté dans ce chapitre le développement et les tests du système de détection d'émotion utilisant la base de données FER 2013 de Kaggle. Nous avons également présenté l'environnement de développement avec Python Google Colab et Spyder, ainsi que les résultats des tests suivis d'une discussion sur les performances du modèle. Puis on a présenté l'interface de notre programme avec explication des boutons et leur fonctionnement. Ces analyses fournissent un aperçu clair des capacités actuelles du système et des avenues potentielles pour son amélioration future.

# **Conclusion Générale**

## **Conclusion générale**

Dans ce mémoire de Master, nous nous sommes concentrés sur la reconnaissance faciales des émotions en utilisant des approches d'apprentissage profond. L'objectif principal était de développer une application pratique pour identifier les émotions d'une personne à partir des visages.

Nous avons exploré plusieurs algorithmes d'apprentissage profond, y compris les réseaux de neurones convolutifs (CNN), les réseaux de neurones artificiels, ainsi que des méthodes d'apprentissage par transfert avec différentes architectures comme RES-NET, VGG, et le modèle FER proposé. Les expériences ont été menées en ajustant minutieusement les hyperparamètres tels que la taille du lot, le nombre d'époques, le nombre de couches, et le type de couches (convolutives, d'abandon, BatchNormalization, MaxPooling) pour optimiser la précision et éviter le surajustement ou le sous-ajustement.

L'utilisation de l'ensemble de données FER2013, contenant sept émotions de base, a permis de construire et de tester notre modèle proposé. Ce modèle, un réseau de neurones convolutif profond, a démontré sa capacité à apprendre automatiquement des caractéristiques à partir des images d'origine, surpassant les autres approches telles que RES-NET, VGG, et SVM en termes de performance en apprentissage profond. Nous avons atteint une précision de 69.80% lors des tests, ce qui représente une amélioration significative par rapport aux architectures précédemment utilisées.

En conclusion, ce mémoire a ouvert de nouvelles perspectives dans le domaine de la reconnaissance des émotions à partir d'images faciales en utilisant des approches d'apprentissage profond. Les résultats obtenus montrent le potentiel de ces techniques pour des applications réelles telles que la détection d'émotions dans des environnements contrôlés et non contrôlés, offrant ainsi des opportunités pour de futures recherches et développements dans ce domaine passionnant et en constante évolution.

# **Bibliographie**

- [1] Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2), 124-129.
- [2] Ekman, P. (1992). Are there basic emotions? *Psychological Review*, 99(3), 550-553.
- [3] Keltner, D., & Haidt, J. (1999). Social Functions of Emotions at Four Levels of Analysis. *Cognition and Emotion*, 13(5), 505-521.
- [4] Fischer, A. H., & Manstead, A. S. R. (2008). Social Functions of Emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (pp. 456-468). The Guilford Press.
- [5] Turner, J. H., & Stets, J. E. (2005). *The Sociology of Emotions*. Cambridge University Press.
- [6] Saarni, C. (1999). *The Development of Emotional Competence*. The Guilford Press.
- [7] (Ekman, P., 1993, "Facial Expression and Emotion", *American Psychologist*, 48(4), 384-392).
- [8] Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2), 124-129.
- [9] R.Collobert et J.Weston : 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). 2008, July.
- [10] Ekman, P. (1993). Facial Expression and Emotion. *American Psychologist*, 48(4), 384-392.
- [11] Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44(4), 695-729.
- [12] LeDoux, J. E. (1995). Emotion: Clues from the brain. *Annual Review of Psychology*, 46(1), 209-235
- [13] Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto: Consulting Psychologists Press.
- [14] Ekman, P., & Rosenberg, E. L. (Eds.). (2005). *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. Oxford University Press.
- [15] Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review*, 97(3), 377-395.

- [16] Foued, N. (2019). Reconnaissance d'expression faciale à partir d'un visage réel. Algérie/Guelma.
- [17] Ambler, S. W. (2002). The object primer: Agile model-driven development with UML 2.0. Cambridge University Press.
- [18] Fowler, M., & Scott, K. (2000). UML distilled: Applying the standard object modeling language. Addison-Wesley Professional.
- [19] Larman, C. (2004). UML et les design patterns : Cours et exercices corrigés. Pearson Education.
- [20] A. Fathallah, L. Abdi et A. Douik : Facial expression recognition via deep learning. In 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA) (pp. 745-750), 2017.
- [21] Kh.Lekdioui : Reconnaissance d'états émotionnels par analyse visuelle du visage et apprentissage.
- [22] C.Qi et M.Li et Q.Wang, H.Zhang et J.Xing et Z.Gao et H.Zhang : Facial expressions recognition based on cognition and mapped binary patterns, 2018
- [23] A. Fathallah, L. Abdi et A. Douik : Facial expression recognition via deep learning. In 2017 IEEE/ACS.
- [24] C.Padgett et G.Cottrell : Representing face images for emotion classification. In Advances in neural information processing systems, pages 894–900, 1997.
- [25] Goodfellow, I., Bengio, Y., & Courville, A. (2016). L'apprentissage en profondeur. Presse du MIT.
- [26] Khadija Lekdioui. Reconnaissance d'états émotionnels par analyse visuelle du visage et apprentissage machine. Synthèse d'image et réalité virtuelle [cs.GR]. Université Bourgogne Franche-Comté ; Université Ibn Tofail. Faculté des sciences de Kenitra, 2018..
- [27] BELHADJ Mahdi ,Etude et simulation d'un syst`eme de reconnaissance des expressions faciale.universit´e de Biskra 2019.
- [28] Mane1, M. S. (May-June 2017). Intelligent Facial Emotion Recognition using modified-PSO. Kolhapur, India.
- [29] A. Fathallah, L. Abdi et A. Douik : Facial expression recognition via deep learning. In

2017 IEEE/ACS14th International Conference on Computer Systems and Applications (AICCSA) (pp. 745-750), 2017.

- [30] V.Perlibakas : Face recognition using principal component analysis and loggabor filters. arXiv preprint cs/0605025 evolution for feature selection in facial expression recognition systems. Expert Systems with Applications, 89, 129-137. 2006.
- [31] Dung Nguyen M.Sc., B.Sc : Multimodal Emotion Recognition Using Deep Learning Techniques,School of Electrical Engineering and Computer Science Science and Engineering Faculty Queensland University of Technology 2020.
- [w1] [https://courses.lumenlearning.com/waymaker-psychology/chapter/introduction-to-emotion\]](https://courses.lumenlearning.com/waymaker-psychology/chapter/introduction-to-emotion/)
- [w2] <https://amadium.com/les-emotions-le-coaching-et-le-marketing/>
- [w3] <https://www.cairn.info/30-grandes-notions-de-la-psychologie--9782100763474-page-31.htm>
- [w4] <https://www.pinterest.com/pin/212161832425357040/>
- [w5] <https://www.eiagroup.fr/domaines-expertise/expressions-faciales-et-micro-expressions/>
- [w6] <https://www.bernard-guevorts.com/article/les-emotions-en-detail-la-joie/>
- [w7] <https://www.cairn.info/revue-l-annee-psychologique1-2014-3-page-469.htm>
- [w8] <https://www.afis.org/La-complexe-interpretation-des-expressions-faciales>
- [w9] <https://dataanalyticspost.com/la-reconnaissance-des-expressions-faciales/>
- [w10] <https://www.vecteurdecroissance.com/blog/guide-micro-expressions>
- [w11] <https://www.cairn.info/revue-l-annee-psychologique1-2014-3-page-501.htm>
- [w12] <https://intelligence-artificielle.com/detection-des-emotions-ia/>
- [w13] [https://www.techno-science.net/glossaire-definition/Apprentissage-automatique.html#google\\_vignette](https://www.techno-science.net/glossaire-definition/Apprentissage-automatique.html#google_vignette)
- [w14] [https://fr.wikipedia.org/wiki/Apprentissage\\_non\\_supervis%C3%A9](https://fr.wikipedia.org/wiki/Apprentissage_non_supervis%C3%A9)
- [w15] <https://datascientest.com/apprentissage-non-supervise>
- [w16] <https://www.journaldunet.fr/intelligence-artificielle/guide-de-l-intelligence-artificielle/1501309-apprentissage-non-supervise/>
- [w17] <https://www.journaldunet.fr/intelligence-artificielle/guide-de-l-intelligence-artificielle/1501309-apprentissage-non-supervise/>
- [w18] [https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/26543781/apprentissage-supervise?utm\\_campaign=Redirection%20des%20anciens%20outils&utm\\_content=id\\_fiche](https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/26543781/apprentissage-supervise?utm_campaign=Redirection%20des%20anciens%20outils&utm_content=id_fiche)

%3D26543781&utm\_source=GDT

- [w19] <https://blog.octo.com/les-methodes-ensemblistes-pour-algorithmes-de-machine-learning>
- [w20] <https://www.axiocode.com/methode-agile-vs-classique-quelle-methode-utiliser/>
- [w21] <https://aws.amazon.com/fr/what-is/deep-learning/>
- [w22] <https://france.devoteam.com/paroles-dexperts/learning-by-doing-apprentissage-par-renforcement/>
- [w23] [https://fr.wikipedia.org/wiki/Apprentissage\\_automatique](https://fr.wikipedia.org/wiki/Apprentissage_automatique)
- [w24] <https://www.ibm.com/fr-fr/topics/machine-learning>
- [w25] [https://en.wikipedia.org/wiki/Machine\\_learning](https://en.wikipedia.org/wiki/Machine_learning)
- [w26] [https://en.wikipedia.org/wiki/Deep\\_learning](https://en.wikipedia.org/wiki/Deep_learning)
- [w27] <https://culturesciencesphysique.ens-lyon.fr/ressource/IA-apprentissage-Rousseau.xml>
- [w28] <https://theses.hal.science/tel-03783228/file/these.pdf>
- [w29] <https://www.geeksforgeeks.org/what-is-reinforcement-learning/>
- [w30] <https://towardsdatascience.com/reinforcement-learning-101-e24b50e1d292>
- [w31] <https://aws.amazon.com/fr/what-is/reinforcement-learning/>
- [w32] <https://ideascale.com/fr/blogues/uml-diagram-definition/>
- [w33] <https://www.datasciencetoday.net/index.php/fr/component/content/article?id=173:les-reseaux-de-neurones-convolutifs&Itemid=667>
- [w34] <https://www.python.org/>
- [w35] <https://research.google.com/colaboratory/faq.html>
- [w36] <https://docs.spyder-ide.org/current/index.html>
- [w37] <https://opencv.org/about/>
- [w38] <https://fr.wikipedia.org/wiki/Scikit-learn>
- [w39] [https://www.tutorialspoint.com/numpy/numpy\\_introduction.htm](https://www.tutorialspoint.com/numpy/numpy_introduction.htm)
- [w40] <http://www.keras.com>
- [w41] <https://pandas.pydata.org/>
- [w42] <http://www.tensorflow.com>
- [w43] <https://www.kaggle.com/datasets/msmbare/fer2013>
- [w44] <https://jason-adam.github.io/resnet50/>



**République algérienne démocratique et populaire**  
**Ministère de l'enseignement supérieur et de la recherche**  
**scientifique**



Université 20 août 1955- Skikda

Faculté de science

Département Informatique

**Titre du projet :**

**EmoSense**

**Un projet d'obtention d'un Master + Certificat d'une institution  
émergente (Startup) dans le cadre de l'arrêté ministériel 1275**



**EmoSense**

اموسانس

**Année universitaire :**

**2023\_2024**

1- Équipe de surveillance :

<b>Équipe de supervision</b>	
<b>Spécialisation :</b>  <b>Informatique</b>	<b>Superviseur principal (01) :</b>  <b>Mazouzi Smaine</b>
<b>Spécialisation :</b>  <b>Informatique</b>	<b>Superviseur principal (02) :</b>  <b>LalaBouali Yassmine</b>

2- Groupe de travail :

<b>Collège</b>	<b>Spécialisation</b>	<b>L'équipe de projet</b>
Université 20 août 1955- Skikda	Informatique master 2 IA	<b>Étudiant : Bouacha Rami Ahmed Mokhtar</b>
Université 20 août 1955- Skikda	Informatique master 2 IA	<b>Étudiant : Ben Haddad Aghiles</b>
Université 20 août 1955- Skikda	Informatique master 2 IA	<b>Étudiant : Kraim Raji</b>

# Remerciement

Premièrement, nous remercions Dieu tout puissant pour tous les bienfaits qu'Il nous a accordés et pour nous avoir donné la force et la volonté nécessaires pour accomplir ce travail.

Nous exprimons notre profonde gratitude au Dr. Mazouzi Smaine et à Madame Lala Bouali Yasmine, nos superviseurs, pour leur accompagnement tout au long de ce projet. Leurs informations précieuses et leurs conseils nous ont été d'une aide inestimable. Nous implorons Allah de guider leurs pas, de réaliser leurs aspirations et de les récompenser de tout bien.

Nous remercions également chaleureusement tous les responsables et travailleurs de l'incubateur et de l'entrepreneuriat de l'Université 20 Août 1955 à Skikda pour leur assistance et les facilités qu'ils nous ont accordées. Qu'Allah les récompense également de tout bien.

Nous exprimons notre respect aux membres du jury pour avoir accepté de participer et de juger notre travail. Nous remercions également tous les enseignants qui ont contribué à notre instruction durant nos années d'étude, ainsi que tous ceux qui nous ont aidés de près ou de loin à l'élaboration de ce modeste travail.

## Dédicace

Avec tout l'amour que nous portons dans nos cœurs, nous dédions ce mémoire de fin d'études universitaires à nos deux êtres respectifs qui nous sont les plus chers dans ce monde : Nos parents respectifs. Nous dédions ce modeste travail aussi à nos familles respectives, nos amis ainsi qu'à tous nos enseignants.

## Index des matières :

Liste des figures.....	6
Liste des tables.....	6
Introduction .....	7
Le premier axe : Présentation du projet (EmoSense) .....	9
1. L'idée de projet (la solution proposée) :.....	9
2. Les valeurs suggérées :.....	10
3. L'équipe de travail :.....	11
4. Les objectifs du projet : .....	12
5. Résumé des Objectifs.....	13
6. Calendrier de réalisation du projet EmoSense : .....	14
Deuxième axe : Aspects inovateurs .....	8
1. La nature des innovations.....	8
2. Les domaines d'innovation :.....	8
Troisième axe : Analyse stratégique du marché .....	11
1. Le segment du marché :.....	11
2. La mesure de l'intensité de la concurrence : .....	12
3. La stratégie marketing : .....	13
Le quatrième axe : Plan de production et organisation.....	17
1. Le processus de production :.....	17
2. L'approvisionnement .....	19
3. La main d'œuvre .....	19
4. Les principaux partenaires : .....	21
Le cinquième axe : Le plan financier.....	24
1. Les coûts et les charges.....	24
2. Le chiffre d'affaires .....	25
3. Les comptes de résultats escomptés .....	25
4. Plan trésorerie .....	26
Sixième axe : Prototype expérimental .....	26
1. Description du programme :.....	26
2. Explication de l'interface du programme .....	27
Septième axe : Business model canvas (BMC).....	35

## Liste des figures

Figure 1 : Page d'accueil. ....	27
Figure 2 : Page d'information.....	28
Figure 3 : Menu de la détection offline. ....	29
Figure 4 : Choisir les émotions.....	30
Figure 5 : Menu de la détection en temps réel. ....	31
Figure 6 : Paramètres d'alerte. ....	32
Figure 7 : Paramètres d'alerte. ....	33
Figure 8 : What to detect. ....	35
Figure 9 : Résultat de la détection. ....	36

## Liste des tables

Tableau 1 : Tableau chiffres d'affaires .....	25
Tableau 2 : résultats escomptés.....	25
Tableau 3 : BRF .....	25
Tableau 4 : Scenario optimiste. ....	26
Tableau 5 : Scénario pessimiste.....	27
Tableau 6 : BMC.....	35

# Introduction

EmoSense est une plate-forme innovante spécialisée dans la détection des émotions faciales à travers des technologies avancées d'intelligence artificielle. Notre programme permet non seulement d'analyser les émotions en temps réel mais aussi d'interpréter les réactions émotionnelles à partir de vidéos enregistrées. Conçue pour répondre aux besoins variés des entreprises, EmoSense offre des fonctionnalités robustes allant de la surveillance de sécurité à l'amélioration de l'expérience client. Grâce à notre engagement envers l'excellence technologique et à notre capacité d'innovation continue, nous transformons la manière dont les organisations comprennent et réagissent aux émotions humaines, ouvrant ainsi de nouvelles perspectives dans divers secteurs.

# **Le premier axe :**

# **Présentation du**

# **projet**

## Le premier axe : Présentation du projet (EmoSense)

### 1. L'idée de projet (la solution proposée)

Notre startup, EmoSense, est née de l'enthousiasme et de la passion d'un groupe d'étudiants en intelligence artificielle. Initialement, nous nous sommes concentrés sur la détection faciale, un domaine fascinant et pleine évolution. En approfondissant nos études, nous avons eu l'idée d'intégrer la reconnaissance des émotions dans nos systèmes de détection. L'idée a germé lorsque nous avons réalisé à quel point une telle technologie pourrait être utile au quotidien, notamment dans des domaines comme la sécurité.

Imaginez un système de surveillance qui ne se contente pas de détecter les visages, mais qui peut également comprendre les émotions des individus. Par exemple, dans un lieu public, un système capable de détecter des comportements suspects ou des états émotionnels de stress pourrait alerter les autorités de manière proactive, améliorant ainsi la sécurité globale.

Une autre source d'inspiration pour EmoSense est venue de notre expérience dans un restaurant. Nous avons observé une situation courante : un client attendait sa commande depuis trop longtemps et commençait à s'énerver. Nous avons alors imaginé un système qui pourrait détecter cette colère et alerter immédiatement le gérant ou le responsable des caméras de surveillance. Grâce à un simple bip, le gérant pourrait intervenir rapidement, envoyer un serveur pour résoudre le problème, et ainsi améliorer la satisfaction du client.

En réfléchissant à d'autres applications, nous avons vu le potentiel de notre technologie dans divers secteurs : du commerce de détail, où elle pourrait aider à comprendre les réactions des clients face à des produits ou des services, jusqu'à la santé mentale, en fournissant des outils pour suivre l'humeur des patients et offrir des interventions plus personnalisées.

Ainsi, EmoSense est né de la volonté de transformer la technologie de détection faciale en un outil puissant pour comprendre et réagir aux émotions humaines, rendant notre quotidien plus sûr et plus agréable.

## 2. Les valeurs suggérées

- **Sécurité renforcée** : Pour les entreprises utilisant des caméras de sécurité, EmoSense peut aider à détecter des comportements suspects ou des anomalies émotionnelles, contribuant ainsi à une meilleure prévention des incidents de sécurité.
- **Amélioration de l'expérience client** : En permettant aux restaurants et aux magasins de mieux comprendre les émotions des clients, ils peuvent adapter leurs services et offres pour améliorer la satisfaction et l'engagement des clients.
- **Optimisation des ressources humaines** : Les entreprises peuvent utiliser les données émotionnelles pour former et coacher leurs employés, en leur fournissant des feedbacks basés sur les interactions avec les clients, ce qui peut améliorer les performances du personnel.
- **Réactivité en temps réel** : La capacité à détecter et analyser les émotions en temps réel permet aux entreprises de réagir rapidement aux besoins et aux problèmes des clients, améliorant ainsi l'efficacité opérationnelle.
- **Avantage concurrentiel** : En adoptant une technologie innovante comme EmoSense, les entreprises peuvent se différencier de leurs concurrents en offrant un niveau de service personnalisé et attentif qui répond mieux aux attentes des clients.
- **Réduction des coûts** : En identifiant rapidement les problèmes et en améliorant la satisfaction client, EmoSense peut aider à réduire les coûts liés aux plaintes, aux retours et à la fidélisation des clients.
- **Responsabilité sociale** : En adoptant des pratiques durables et éthiques, nous soutenons les initiatives communautaires et environnementales, aidant à créer un impact positif sur la société dans son ensemble.
- **Amélioration de la qualité de vie** : En rendant les environnements plus sûrs et les interactions plus humaines, nous contribuons à une meilleure qualité de vie pour les individus au quotidien

### **3. L'équipe de travail**

L'équipe du projet est composée des étudiants Ben Haddad Aghiles, Bouacha Rami Ahmed Mokhtar, et Kraim Radji. L'équipe est responsable du développement des algorithmes d'intelligence artificielle et de leur optimisation. En outre, les tâches secondaires ont été réparties entre eux comme suit :

#### **Les tâches secondaires de l'étudiant Ben Haddad Aghiles : Conception et Développement Initial**

##### **1. Conception du Produit**

- Développe et affine les spécifications du produit en collaboration avec les membres de l'équipe.
- Crée des wireframes et des maquettes pour visualiser l'interface utilisateur.

##### **2. Développement Front-end**

- Responsable de la conception et du développement de l'interface utilisateur (UI).
- Intègre les maquettes graphiques et assure une expérience utilisateur fluide.

##### **3. Participation au Marketing**

- Collabore avec l'équipe marketing pour intégrer des éléments visuels et interactifs dans les campagnes.
- Contribue à la création de contenu marketing pour le site web et les réseaux sociaux.

#### **Les tâches secondaires de l'étudiant Bouacha Rami Ahmed Mokhtar : Infrastructure et Intégration**

##### **1. Développement Back-end**

- Gère la logique serveur et la base de données pour assurer la sécurité et la performance du produit.
- Implémente des API et assure l'intégration avec d'autres systèmes si nécessaire.

##### **2. Gestion des Données**

- Responsable du stockage et de la gestion des données utilisateur de manière sécurisée et conforme à la réglementation.
- Optimise les requêtes et la gestion des données pour améliorer les performances globales du système.

##### **3. Support Client**

- Assure la prise en charge initiale des requêtes et des problèmes des clients.
- Coordonne avec l'équipe principale de support client pour résoudre les problèmes complexes et assurer la satisfaction des utilisateurs.

### Les tâches secondaires de l'étudiant Kraim Radji : Tests et Qualité

#### 1. Tests et Assurance Qualité

- Conçoit et met en œuvre des plans de test complets pour garantir la qualité du produit.
- Effectue des tests fonctionnels, de performance et de sécurité réguliers.

#### 2. Déploiement et Maintenance

- Gère le déploiement du produit sur les environnements de test et de production.
- Assure la maintenance continue du logiciel, y compris les mises à jour et les correctifs nécessaires.

#### 3. Gestion Administrative et Financière

- Aide à la gestion des finances en suivant les dépenses liées au développement et à l'infrastructure.
- Participe à la gestion des opérations administratives telles que la facturation et les relations avec les fournisseurs.

### 4. Les objectifs du projet

#### Objectifs Stratégiques à Court Terme (6 à 12 mois)

✓ **Lancement du projet dans la wilaya de Skikda** : Nous visons à introduire notre programme sur le marché en commençant par notre wilaya, Skikda. Notre objectif est de faire connaître EmoSense et de démontrer son efficacité dans cette région.

✓ **Couvrir 20% des clients potentiels dans la région de Skikda** : Nous ciblons les secteurs tels que les restaurants, les commerces de détail, et autres établissements où notre service de détection des émotions peut apporter une réelle valeur ajoutée. Notre objectif est de couvrir 20% de ces clients potentiels dans la première année.

#### Objectifs Stratégiques à Moyen Terme (1 à 3 ans)

✓ **Couvrir 90% des clients potentiels de la région de Skikda** : Nous visons à étendre notre couverture à 90% des clients potentiels dans la wilaya de Skikda, en renforçant notre présence et en assurant une adoption plus large de notre système.

## Premier axe : présentation du projet

✓ **Expansion vers d'autres wilayas** : Nous prévoyons d'étendre notre service au-delà de Skikda, en entrant dans différentes wilayas du pays. Cette expansion nous permettra de toucher un marché plus large et de renforcer notre position en Algérie.

✓ **Vente en Ligne** : Commencer à vendre le programme Emosense en ligne pour atteindre les clients situés dans des zones éloignées ou difficilement accessibles physiquement.

### Objectifs Stratégiques à Long Terme (3 à 5 ans et au-delà)

✓ **Couvrir l'Algérie entière** : Notre objectif à long terme est de déployer EmoSense à travers toute l'Algérie. Nous ambitionnons de devenir le leader national en matière de détection des émotions, offrant nos services dans toutes les wilayas et couvrant ainsi l'ensemble du territoire algérien.

✓ **diversification des services** : Nous prévoyons d'intégrer des services différents comme l'installation des caméras de surveillance.

## 5. Résumé des Objectifs

### • Court Terme :

- Lancer le projet dans la wilaya de Skikda.
- Couvrir 20% des clients potentiels dans Skikda (restaurants, commerces de détail, etc.).

### • Moyen Terme :

- Couvrir 90% des clients potentiels à Skikda.
- S'étendre à d'autres wilayas pour toucher un marché plus large.
- Vend en ligne

### • Long Terme :

- Couvrir l'ensemble du territoire algérien et devenir le leader national en matière de détection des émotions.
- Diversification des services comme installation des caméras de surveillance.

## 6. Calendrier de réalisation du projet EmoSense

Tâche	Durée	Résultats Clés
Étude de marché et analyse des besoins locaux	1 mois	Rapport d'étude de marché, liste des besoins
Développement et adaptation du produit	2 mois	Version finale du produit, rapport de tests
Création de la présence en ligne et des supports marketing	1 mois	Site web opérationnel, supports marketing
Offrir le programme gratuitement à 20 clients pendant 1 mois et demi	1,5 mois	Feedback des 20 clients, rapport d'analyse
Formation de l'équipe de support client	1 mois	Équipe de support opérationnelle, documentation des processus
Création du Site Web	2 mois	un site pour vendre et avoir un guide d'installation et une section de feedback
Collecte et analyse du feedback annuel	1 mois en fin d'année	Rapport de feedback annuel, plan d'amélioration
Acquisition des premiers clients à Skikda	3 mois	Premiers clients acquis, feedback initial
Amélioration continue et personnalisation du produit	6 mois	Version améliorée du produit, satisfaction client
Expansion vers d'autres wilayas	12 mois	Présence dans plusieurs wilayas, augmentation de la part de marché
Concentration sur le marketing et la promotion	En continu sur 24 mois	Augmentation de la visibilité, notoriété accrue
Organisation et participation à des événements	En continu sur 24 mois	Augmentation de la visibilité, notoriété accrue
Partenariats avec des entreprises de caméras de surveillance	En continu sur 24 mois	Plus de publicité, collaborations renforcées

## Premier axe : présentation du projet

Collecte et analyse du feedback annuel	1 mois en fin d'année	Rapport de feedback annuel, plan d'amélioration
Couvrir 90% des clients potentiels à Skikda	6 mois	Couverture de 90%, forte présence locale
Expansion nationale et couverture complète de l'Algérie	24 mois	Couverture nationale, position de leader
Diversification des activités de la startup (ex: installation de caméras de surveillance)	12 mois	Nouvelles lignes de revenus, expansion des services

# Deuxième axe :

# Aspects

# innovateurs

## Deuxième axe : Aspects innovateurs

### 1. La nature des innovations

Technologique : Utilisation avancée de l'intelligence artificielle et de l'apprentissage automatique pour la détection d'émotions en temps réel à partir de flux vidéo de caméras de surveillance. Cela inclut le développement d'algorithmes de reconnaissance faciale et émotionnelle robustes pour identifier et interpréter les expressions faciales avec précision.

Produit : Création d'une solution intégrée comprenant à la fois du matériel (caméras de haute résolution adaptées à la reconnaissance faciale) et du logiciel (plateforme analytique et d'alerte en temps réel). Cette innovation vise à offrir une interface conviviale pour les utilisateurs finaux tout en garantissant des performances de détection fiables.

Processus : Adoption d'un processus de développement agile pour permettre des ajustements rapides et des améliorations continues basées sur les retours d'expérience des utilisateurs. Cela inclut des cycles itératifs de test, de validation et de déploiement pour optimiser l'efficacité opérationnelle et la satisfaction client.

Organisationnelle : Création d'une culture d'entreprise axée sur l'innovation et la collaboration, favorisant la recherche constante de nouvelles applications et améliorations technologiques. Cela implique également la formation continue du personnel pour maintenir une expertise de pointe dans le domaine de la reconnaissance d'émotions.

Modèle d'affaires : Introduction d'un modèle économique flexible basé sur l'abonnement ou la licence, permettant aux clients de choisir des solutions adaptées à leurs besoins spécifiques tout en assurant des revenus récurrents pour Emosense. Cela peut également inclure des services complémentaires tels que la maintenance, le support client avancé et la personnalisation de solutions.

### 2. Les domaines d'innovation

#### 1. Nouveau processus

Amélioration de l'efficacité opérationnelle à travers l'intégration de la technologie Emosense dans les systèmes de surveillance existants, réduisant ainsi le temps de réponse aux incidents émotionnels et améliorant la sécurité globale des environnements surveillés.

### 2. Nouvelles fonctionnalités :

- Développement de fonctionnalités avancées telles que la reconnaissance émotionnelle fine-grain, permettant une détection plus précise et une réponse adaptée aux nuances des émotions détectées, améliorant ainsi l'expérience utilisateur et la satisfaction client.

### 3. Nouveaux clients :

- Expansion de la clientèle pour inclure de nouveaux segments tels que les grands centres commerciaux, les complexes résidentiels et les institutions publiques, où la surveillance et la gestion des émotions peuvent jouer un rôle crucial dans la sécurité et la gestion des crises.

### 4. Nouvelles offres :

- Introduction de nouvelles offres de services complémentaires, comme des analyses avancées des données d'émotions pour optimiser les opérations de gestion de la clientèle et améliorer la prise de décision stratégique des entreprises clientes.

### 5. Nouveaux modèles :

- Adoption d'un modèle d'affaires innovant basé sur l'abonnement mensuel ou annuel pour l'utilisation de la technologie Emosense, offrant une flexibilité financière aux clients tout en générant des revenus récurrents pour Emosense. Cette approche peut inclure des packages personnalisés en fonction des besoins spécifiques de chaque client.

**Troisième axe :**

**Analyse**

**stratégique du**

**marché**

## Troisième axe : Analyse stratégique du marché

Pour la startup EmoSense, voici une analyse des segments de marché, en mettant l'accent sur les lieux publics à fort trafic où nos produits pourraient être déployés :

### 1. Le segment du marché

#### Marché Potentiel :

- **Qui achète nos produits ?** Principalement, les entreprises et les institutions situées dans des lieux publics à fort trafic tels que les centres commerciaux, les gares, les aéroports, les hôtels, les restaurants, et les institutions financières.
- **Qu'est-ce qui les motive à le faire ?** Améliorer la sécurité en détectant les émotions dans des environnements où il y a beaucoup de gens en mouvement. Optimiser le service client en comprenant mieux les réactions émotionnelles des visiteurs et clients.
- **Où sont-ils situés ?** Nous ciblons initialement Skikda pour le lancement, avec des plans d'expansion dans d'autres wilayas d'Algérie à moyen terme.
- **Combien sont-ils ?** Nous visons à couvrir environ 20% des entreprises situées dans des endroits publics à Skikda qui peuvent bénéficier de notre service, avec un objectif à long terme de couverture nationale.

#### Marché Cible (Segment) :

- Notre marché cible inclut les entreprises et institutions situées dans des endroits publics à fort trafic, où la détection d'émotions peut améliorer la sécurité et le service client.
- Les secteurs cibles comprennent les centres commerciaux, les gares, les aéroports, les hôtels, les restaurants, et les institutions financières.

#### Pourquoi avons-nous choisi ce marché cible ?

- Ces lieux publics à fort trafic offrent une grande opportunité pour la technologie de détection d'émotions, en raison de la diversité des interactions émotionnelles qui s'y produisent quotidiennement.
- La demande pour des solutions de sécurité et de service client améliorées dans ces environnements est élevée et croissante.

---

## Troisième axe : Analyse stratégique du marché

### Possibilité de Conclure des Contrats d'Achat avec Certains Clients Importants :

- Nous visons à conclure des contrats avec des clients importants comme des chaînes de centres commerciaux, des grandes gares et aéroports, ainsi que des hôtels et restaurants renommés.
- L'accès à ces clients dépendra de notre capacité à démontrer la valeur ajoutée de notre produit à travers des études de cas, des essais pilotes réussis, et des démonstrations efficaces.
- L'établissement de partenariats stratégiques avec des acteurs clés dans les secteurs de la sécurité et de la technologie est essentiel pour renforcer notre crédibilité et notre accès au marché.

En concentrant nos efforts sur ces segments de marché spécifiques, nous maximisons notre potentiel de croissance et d'impact en introduisant efficacement la technologie de détection d'émotions dans des lieux publics à forte fréquentation en Algérie.

## 2. La mesure de l'intensité de la concurrence :

### Absence de Concurrents Directs :

- Il n'y a pas de concurrents directs en Algérie proposant des systèmes similaires de détection d'émotions intégrées à des solutions de sécurité et de service client.
- Cela place Emosense dans une position unique pour capturer le marché émergent de la détection d'émotions, sans rival direct immédiat.

### Concurrents Indirects Potentiels :

- Bien qu'il n'y ait pas de concurrents directs, il pourrait y avoir des alternatives indirectes comme des solutions de sécurité traditionnelles sans analyse d'émotions ou des systèmes de service client sans fonctionnalités avancées d'interaction émotionnelle.
- Ces alternatives pourraient être considérées comme des concurrents indirects, bien que leur capacité à répondre aux besoins spécifiques d'Emosense soit limitée.

### Avantages de l'Absence de Concurrence :

- **Opportunité de Leader** : Emosense a l'opportunité de devenir le premier et le leader sur le marché algérien pour les solutions de détection d'émotions, établissant une forte présence et une base clientèle fidèle.
- **Développement de la Demande** : L'absence de concurrents directs signifie que la demande pour de telles technologies est sous-exploitée, offrant à Emosense une occasion unique de stimuler et de capturer cette demande naissante.

---

## Troisième axe : Analyse stratégique du marché

### Défis Potentiels :

- **Éducation du Marché** : Emosense devra investir dans l'éducation du marché sur les avantages et les applications de la détection d'émotions, puisqu'il s'agit d'une technologie nouvelle et potentiellement inconnue pour de nombreux acteurs du marché algérien.
- **Validation de la Technologie** : Prouver l'efficacité et les avantages de notre technologie à des clients potentiels peut être un défi initial, nécessitant des démonstrations et des études de cas convaincantes.

Cette approche permet à Emosense de se positionner stratégiquement sur un marché sans concurrent direct, tout en relevant les défis associés à l'éducation du marché et à la validation de la technologie.

## 3. La stratégie marketing

La stratégie marketing d'Emosense vise à sensibiliser et convaincre les clients potentiels d'adopter nos produits de détection d'émotions, en tenant compte de nos ressources financières et en équilibrant le mix marketing pour assurer le succès de nos initiatives.

### 1. Objectifs de la Stratégie Marketing :

- **Sensibilisation et Éducation** : Informer le marché algérien sur les avantages de la détection d'émotions dans les environnements publics à fort trafic.
- **Génération de la Demande** : Stimuler l'intérêt et la demande pour nos solutions auprès des entreprises cibles.
- **Positionnement** : Établir Emosense comme le leader de confiance en matière de technologie de détection d'émotions en Algérie.

### 2. Stratégies Marketing Clés

#### A. Segmentation et Ciblage

- Identifier et cibler les entreprises situées dans des lieux publics à fort trafic tels que les centres commerciaux, les gares, les aéroports, les hôtels, et les restaurants.
- Adapter notre message marketing pour répondre aux besoins spécifiques de chaque segment de clientèle.

---

## Troisième axe : Analyse stratégique du marché

### B. Positionnement

- Mettre en avant les avantages compétitifs d'Emosense, notamment sa précision dans la détection des émotions, son intégration facile avec les systèmes de sécurité existants, et son impact sur l'amélioration de la sécurité et du service client.
- Communiquer clairement notre proposition de valeur unique par rapport aux alternatives disponibles.

### C. Communication et Promotion :

- Utiliser une combinaison de canaux de communication adaptés au marché algérien, incluant les médias sociaux, le marketing digital, les relations publiques et les événements sectoriels.
- Organiser des démonstrations de produits, des webinaires et des sessions de formation pour éduquer les prospects et démontrer l'efficacité de notre technologie.

### D. Prix :

- Adopter une stratégie de prix compétitive qui reflète la valeur perçue de nos produits tout en étant accessible aux entreprises locales.
- Proposer des modèles de tarification flexibles pour répondre aux différents budgets et exigences des clients.

### E. Distribution :

- Adopter une stratégie de prix compétitive qui reflète la valeur perçue de nos produits tout en étant accessible aux entreprises locales.
- Proposer des modèles de tarification flexibles pour répondre aux différents budgets et exigences des clients.

## 3. Budget Marketing:

- Allouer judicieusement les ressources financières disponibles pour maximiser la portée et l'impact de nos campagnes marketing.
- Mesurer régulièrement le retour sur investissement (ROI) pour ajuster et optimiser nos initiatives marketing en fonction des résultats obtenus.

En mettant en œuvre cette stratégie marketing bien planifiée et équilibrée, Emosense est positionné pour attirer efficacement les clients potentiels, stimuler la demande pour ses produits innovants, et établir une présence solide sur le marché émergent de la détection d'émotions en Algérie.

# **Le quatrième axe : Plan de production et organisation**

## Le quatrième axe : Plan de production et organisation

### 1. Le processus de production

Pour EmoSense, le processus de production du programme de détection des émotions passe par plusieurs étapes essentielles pour garantir un produit de haute qualité. Voici une description détaillée de ce processus :

#### Étape 1 : Développement Initial

1. Front-end Développement
  - Conception et développement de l'interface utilisateur.
  - Mise en place de la structure de navigation et des écrans principaux.
2. Back-end Développement
  - Création des serveurs et bases de données.
  - Développement des API pour la communication entre le front-end et le back-end.
3. Algorithmes d'Intelligence Artificielle
  - Développement des algorithmes de détection d'émotions.
  - Intégration de modèles d'apprentissage automatique pour la reconnaissance des expressions faciales et des émotions.

#### Étape 2 : Intégration et Tests

1. Intégration des composants
  - Assemblage des éléments front-end, back-end et IA.
  - Assurer la compatibilité et la communication entre les différentes parties du programme.
2. Tests Unitaires et Fonctionnels
  - Exécuter des tests unitaires pour chaque composant du programme.
  - Réaliser des tests fonctionnels pour s'assurer que toutes les fonctionnalités répondent aux spécifications.
3. Tests d'Utilisabilité

- Faire tester le programme par des utilisateurs pour recueillir des retours sur l'expérience utilisateur.
- Ajuster l'interface et les fonctionnalités en fonction des retours.

### Étape 3 : Finalisation et Lancement

1. Améliorations et Corrections
  - Apporter les améliorations nécessaires et corriger les bugs identifiés lors des tests.
2. Préparation du Lancement
  - Préparer les supports marketing et la documentation utilisateur.
  - Mettre en place les infrastructures nécessaires pour le déploiement.
3. Lancement du Programme
  - Déployer le programme sur les serveurs et le rendre accessible aux clients.
  - Assurer le support technique pour les premiers utilisateurs.

### Étape 4 : Suivi et Amélioration Continue

1. Collecte de Feedback
  - Recueillir les retours des utilisateurs sur les performances et les fonctionnalités du programme.
2. Support et Maintenance
  - Offrir un support technique continu pour résoudre les problèmes et assister les utilisateurs.
  - Effectuer la maintenance régulière du programme pour assurer son bon fonctionnement et sa sécurité.
3. Améliorations et Mises à Jour
  - Intégrer les suggestions des utilisateurs et les innovations technologiques pour améliorer le programme.
  - Déployer des mises à jour régulières pour ajouter de nouvelles fonctionnalités et améliorer l'expérience utilisateur.

### 2. L'approvisionnement

#### Politique d'Achat

Étant donné que notre programme est déjà développé, notre politique d'approvisionnement se concentre sur les éléments nécessaires pour la maintenance, l'amélioration continue et le support.

#### 1. Matériels et Fournitures

- **Équipements Informatiques** : Ordinateurs pour les développeurs et le support technique.
- **Logiciels** : Licences de développement, outils de gestion de projets, outils de sécurité.

#### 2. Équipements Informatiques et Logiciels

- **Pc performant** : Pour les ordinateurs et serveurs.

### 3. La main d'œuvre

- **Nombre de Postes Créés par le Projet**

Le projet "EmoSense" est prévu de créer plusieurs postes pour assurer son bon fonctionnement et son expansion. Les postes potentiels incluent :

- **Développeurs**

- Nombre de postes : 3

- Description : Responsables du développement, de la maintenance et de l'amélioration continue du programme de détection des émotions.

- Localisation : Skikda.

- **Support Technique et Service Client**

- Nombre de postes : 1

- Description : Installation, configuration et support technique du programme chez les clients, gestion des relations clients, collecte de feedback, assistance à la clientèle.

- Localisation : Skikda.

- **Équipe Marketing et Communication**

---

## Quatrième axe : Plan de production et organisation

- Nombre de postes : 2

- Description : Développement et mise en œuvre de stratégies marketing pour promouvoir le programme, gestion des médias sociaux, création de contenu promotionnel.

- Localisation : Skikda.

- **Nature et Type de Main-d'œuvre**

La nature et le type de main-d'œuvre dont nous avons besoin pour "EmoSense" sont les suivants :

### **Développeurs**

- Compétences requises : Maîtrise avancée de Python, expérience en intelligence artificielle et en traitement d'images, capacité à travailler en équipe.

- Niveau d'expérience : Minimum 2 ans dans le développement logiciel.

### **Support Technique et Service Client**

- Compétences requises : Compétences techniques en configuration de logiciels, excellentes compétences en communication, expérience en service client.

- Niveau d'expérience : Minimum 1 an d'expérience.

### **Équipe Marketing et Communication**

- Compétences requises : Compétences en marketing digital, connaissances approfondies des médias sociaux, capacité à créer du contenu promotionnel efficace.

- Niveau d'expérience : Minimum 1 an d'expérience.

### **Possibilité de Recourir à la Manutention**

Pour le projet "EmoSense", la nécessité de recourir à la manutention physique est limitée en raison de la nature principalement logicielle du produit. Cependant, nous pourrions avoir besoin de services de manutention dans les cas suivants :

- Installation de Matériel : Transport et installation de matériel informatique pour les clients, si nécessaire.

- Événements et Présentations : Transport et mise en place de matériel pour les démonstrations et les présentations lors des événements marketing.

En structurant notre main-d'œuvre autour de ces postes clés, nous nous assurons que "EmoSense" est bien équipé pour répondre aux demandes des clients, assurer une croissance continue et maintenir un haut niveau de satisfaction client.

#### 4. Les principaux partenaires

Notre projet bénéficie du soutien essentiel de plusieurs parties prenantes interconnectées. Les incubateurs d'entreprises, y compris ceux des universités, nous offrent un cadre propice à l'innovation. Les banques, le Centre National du Registre du Commerce, la Direction Générale des Impôts et les investisseurs jouent des rôles clés en facilitant nos opérations et en soutenant notre croissance. Nos fournisseurs et importateurs garantissent un approvisionnement constant. Ensemble, ces partenaires forment un réseau robuste qui soutient notre développement sur le marché.

#### Répartition des Responsabilités

##### Ben Haddad Aghiles : Conception et Développement Initial

##### 1. Conception du Produit

- Développe et affine les spécifications du produit en collaboration avec l'équipe de direction.
- Créer des wireframes et des maquettes pour visualiser l'interface utilisateur.

##### 2. Développement Front-end

- Responsable de la conception et du développement de l'interface utilisateur (UI).
- Intègre les maquettes graphiques et assure une expérience utilisateur fluide.

##### 3. Participation au Marketing

- Collabore avec l'équipe marketing pour intégrer des éléments visuels et interactifs dans les campagnes.
- Contribue à la création de contenu marketing pour le site web et les réseaux sociaux.

### **Rami Ahmed Mokhtrar : Infrastructure et Intégration**

#### **1. Développement Back-end**

- Gère la logique serveur et la base de données pour assurer la sécurité et la performance du produit.
- Implémente des API et assure l'intégration avec d'autres systèmes si nécessaire.

#### **2. Gestion des Données**

- Responsable du stockage et de la gestion des données utilisateur de manière sécurisée et conforme à la réglementation.
- Optimise les requêtes et la gestion des données pour améliorer les performances globales du système.

#### **3. Support Client**

- Assure la prise en charge initiale des requêtes et des problèmes des clients.
- Coordonne avec l'équipe principale de support client pour résoudre les problèmes complexes et assurer la satisfaction des utilisateurs.

### **Kraim Radji : Tests et Qualité**

#### **1. Tests et Assurance Qualité**

- Conçoit et met en œuvre des plans de test complets pour garantir la qualité du produit.
- Effectue des tests fonctionnels, de performance et de sécurité réguliers.

#### **2. Déploiement et Maintenance**

- Gère le déploiement du produit sur les environnements de test et de production.
- Assure la maintenance continue du logiciel, y compris les mises à jour et les correctifs nécessaires.

#### **3. Gestion Administrative et Financière**

- Aide à la gestion des finances en suivant les dépenses liées au développement et à l'infrastructure.
- Participe à la gestion des opérations administratives telles que la facturation et les relations avec les fournisseurs.



**Le cinquième axe :**

**Le plan financier**

## Le cinquième axe : Le plan financier

### 1. Les coûts et les charges

Voici un résumé des principaux coûts associés à la réalisation du projet EmoSense :

- **Location** : 20 000 DA par mois
- **Assurances** : 50 000 DA
- **PC (3 unités)** : 200 000 DA chacun
- **Bureaux (3 unités)** : 80 000 DA chacun
- **Modem et réseau** : 80 000 DA + 20 000 DA d'abonnement mensuel
- **Imprimantes (2 unités)** : 50 000 DA chacune
- **Véhicule** : 3 000 000 DA
- **Marketing (publicité et événements)** : 500 000 DA
- **Salaires** :
  - 3 développeurs : 50 000 DA chacun par mois
  - 1 poste de support technique et service client : 30 000 DA par mois
  - 2 postes en marketing et communication : 30 000 DA chacun par mois

### Modes et Sources d'Obtention de Financement

- **Investissements personnels** : Apport initial des fondateurs.
- **Prêts bancaires** : Solliciter des prêts auprès des banques pour couvrir les coûts initiaux.
- **Incubateurs et subventions** : Tirer parti des programmes de subventions et de soutien des incubateurs, notamment ceux de l'université.
- **Investisseurs privés** : Chercher des investisseurs prêts à financer le projet en échange d'une participation au capital.

### Remboursement

Le plan de remboursement pour les prêts et autres sources de financement :

- **Échéancier des paiements** : Un calendrier détaillé indiquant les dates et montants des remboursements mensuels ou annuels.
- **Priorités de remboursement** : Définir l'ordre des priorités pour les différents types de financement (prêts bancaires, investisseurs privés, etc.).
- **Suivi financier** : Mettre en place un système de suivi pour s'assurer que tous les paiements sont effectués à temps et pour éviter les pénalités de retard.

## Sixième axe : Prototypage expérimental

### 2. Le chiffre d'affaires

Année	Nombre de ventes (Optimiste)	Chiffre d'affaires (Optimiste)(DA)	Nombre de ventes (Pessimiste)	Chiffre d'affaires (Pessimiste)(DA)
Année 1	1 000	10 000 000	0	0
Année 2	1500	15 000 000	0	0
Année 3	2000	20 000 000	100	1 000 000
<b>Total sur 3 ans</b>	<b>4500</b>	<b>45 000 000</b>	<b>100</b>	<b>1 000 000</b>

Tableau 1: Tableau chiffres d'affaires

On veut vendre notre produit à 10000 DA.

### 3. Les comptes de résultats escomptés

Année	Chiffre d'affaires (Optimiste) (DA)	Total des coûts (DA)	Bénéfice/Perte (Optimiste) (DA)	Chiffre d'affaires (Pessimiste) (DA)	Total des coûts (DA)	Bénéfice/Perte (Pessimiste) (DA)
Année 1	10 000 000	3 360 000	6 640 000	0	3 360 000	-3 360 000
Année 2	15 000 000	3 360 000	11 640 000	0	3 360 000	-3 360 000
Année 3	20 000 000	3 360 000	16 640 000	1 000 000	3 360 000	-2 360 000
<b>Total sur 3 ans</b>	<b>45 000 000</b>	<b>10 080 000</b>	<b>34 920 000</b>	<b>1 000 000</b>	<b>10 080 000</b>	<b>-9 080 000</b>

Tableau 2 : résultats escomptés

#### Besoin en Fonds de Roulement (BFR) :

Description	Montant (DA)
Stocks	0
Créance clients	0
Dettes fournisseurs	0
<b>BFR total</b>	<b>0</b>

Tableau 3 : BRF

## Sixième axe : Prototype expérimental

Le BFR est simplifié ici car il n'y a pas de stock, créances clients ou dettes fournisseurs significatives à ce stade.

### 4. Plan trésorerie

Scenario optimiste :

Mois	Revenus (DA)	Dépenses récurrentes (DA)	Dépenses uniques (DA)	Solde mensuel (DA)	Solde cumulé (DA)
Janvier	833 333	260 000	4570000	-3 996 667	-3 996 667
Février	833 333	260 000	0	573 333	-3 423 334
Mars	833 333	260 000	0	573 333	-2 850 001
Avril	833 333	260 000	0	573 333	-2 276 668
Mai	833 333	260 000	0	573 333	-1 703 335
Juin	833 333	260 000	0	573 333	-1 130 002
Juillet	833 333	260 000	0	573 333	-556 669
Août	833 333	260 000	0	573 333	16 664
Septembre	833 333	260 000	0	573 333	589 997
Octobre	833 333	260 000	0	573 333	1 163 330
Novembre	833 333	260 000	0	573 333	1 736 663
Décembre	833 333	260 000	0	573 333	2 309 996

Tableau 4 : Scenario optimiste.

## Sixième axe : Prototype expérimental

### Scénario pessimiste :

Mois	Revenus (DA)	Dépenses récurrentes (DA)	Dépenses uniques (DA)	Solde mensuel (DA)	Solde cumulé (DA)
Janvier	0	260 000	4570000	-3 830 000	-4 830 000
Février	0	260 000	0	-260 000	-5 090 000
Mars	0	260 000	0	-260 000	-5 350 000
Avril	0	260 000	0	-260 000	-5 610 000
Mai	0	260 000	0	-260 000	-5 870 000
Juin	0	260 000	0	-260 000	-6 130 000
Juillet	0	260 000	0	-260 000	-6 390 000
Août	0	260 000	0	-260 000	-6 650 000
Septembre	0	260 000	0	-260 000	-6 910 000
Octobre	0	260 000	0	-260 000	-7 170 000
Novembre	0	260 000	0	-260 000	-7 430 000
Décembre	0	260 000	0	-260 000	-7 690 000

Tableau 5 : Scénario pessimiste

**Le sixième axe :**

**Prototype**

**expérimental**

## Sixième axe : Prototype expérimental

### 1. Description du programme

EmoSense est un programme innovant développé pour la détection des émotions faciales, utilisant une approche avancée basée sur les réseaux de neurones convolutifs (CNN). Ce programme exploite la base de données FER 2013, composée d'images en niveaux de gris de visages centrés, afin d'analyser et classifier sept émotions principales : Colère, Dégoût, Peur, Joie, Tristesse, Surprise et Neutre.

Pour construire EmoSense, nous avons utilisé les environnements de développement Python, Collab et Spyder, ainsi que des bibliothèques de programmation telles que numpy, keras, tensorflow, pandas, scikit-learn et OpenCV.

Le choix d'un modèle CNN permet une analyse robuste et précise des expressions faciales en extrayant des caractéristiques pertinentes des images.

EmoSense offre deux modes principaux de détection d'émotions :

Détection en ligne (en direct) : Analyse des expressions faciales en temps réel, adaptée aux applications nécessitant une réponse rapide et en direct.

Détection hors ligne (vidéo enregistrée) : Analyse des émotions dans des vidéos pré-enregistrées, idéale pour l'analyse rétrospective et la recherche.

## 2. Explication de l'interface du programme

Bienvenue sur la page d'accueil principale de notre système de détection d'émotion EmoSense !

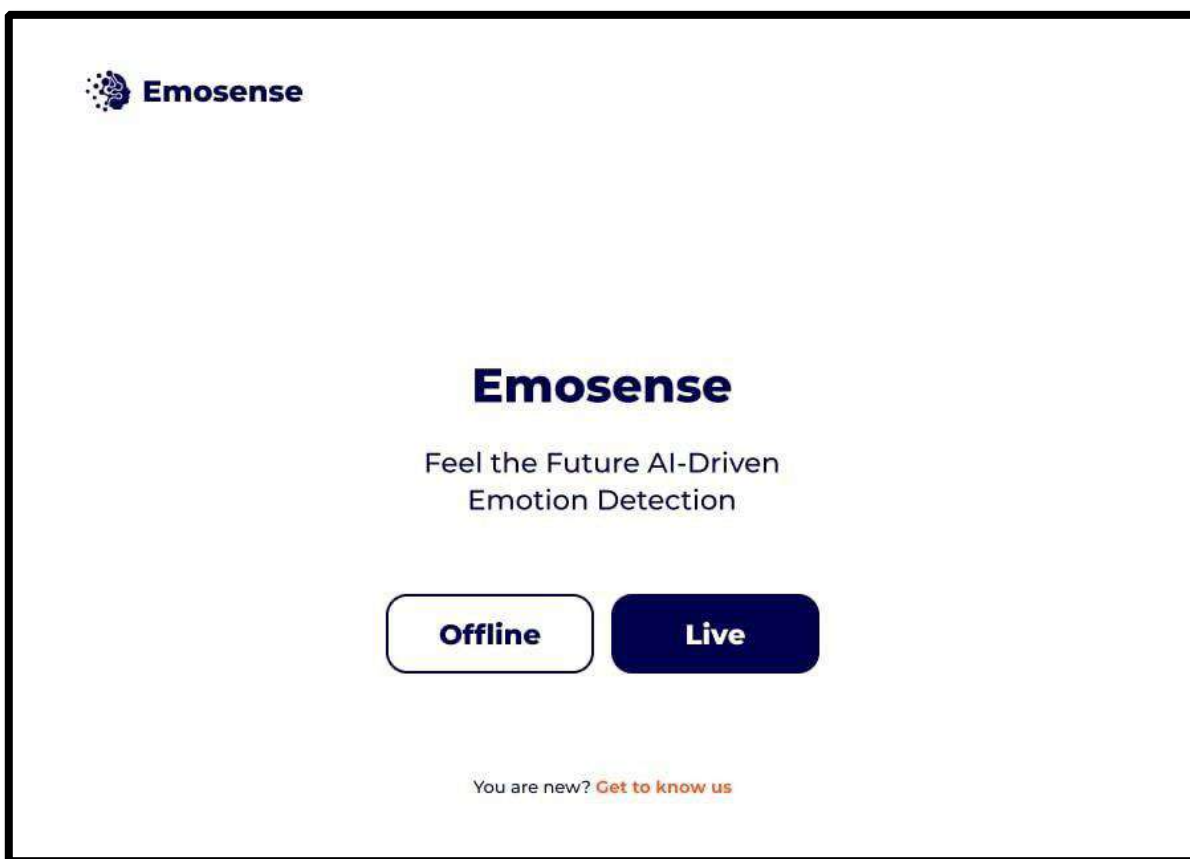


Figure 1 : Page d'accueil.

EmoSense offre deux modes principaux de détection des émotions faciales :

Elle contient deux boutons principaux :

### a. **Offline : Détection des Emotions a partir d'une vidéo enregistrées**

Le bouton Offline permet de détecter les émotions faciales à partir de vidéos déjà enregistrées ou stockées sur le PC. Ce mode est idéal pour analyser des enregistrements existants afin de mieux comprendre les réactions émotionnelles dans des contextes variés.

### b. **Live : Détection des Émotions en Temps Réel**

Le bouton Live est conçu pour la détection des émotions faciales en temps réel. En utilisant ce mode, vous pouvez capturer et analyser les émotions au fur et à mesure qu'elles se produisent, ce qui est particulièrement utile pour des applications comme la surveillance de sécurité, l'amélioration du service client, et bien plus encore.

Voici la page "Get to Know Us" de EmoSense !

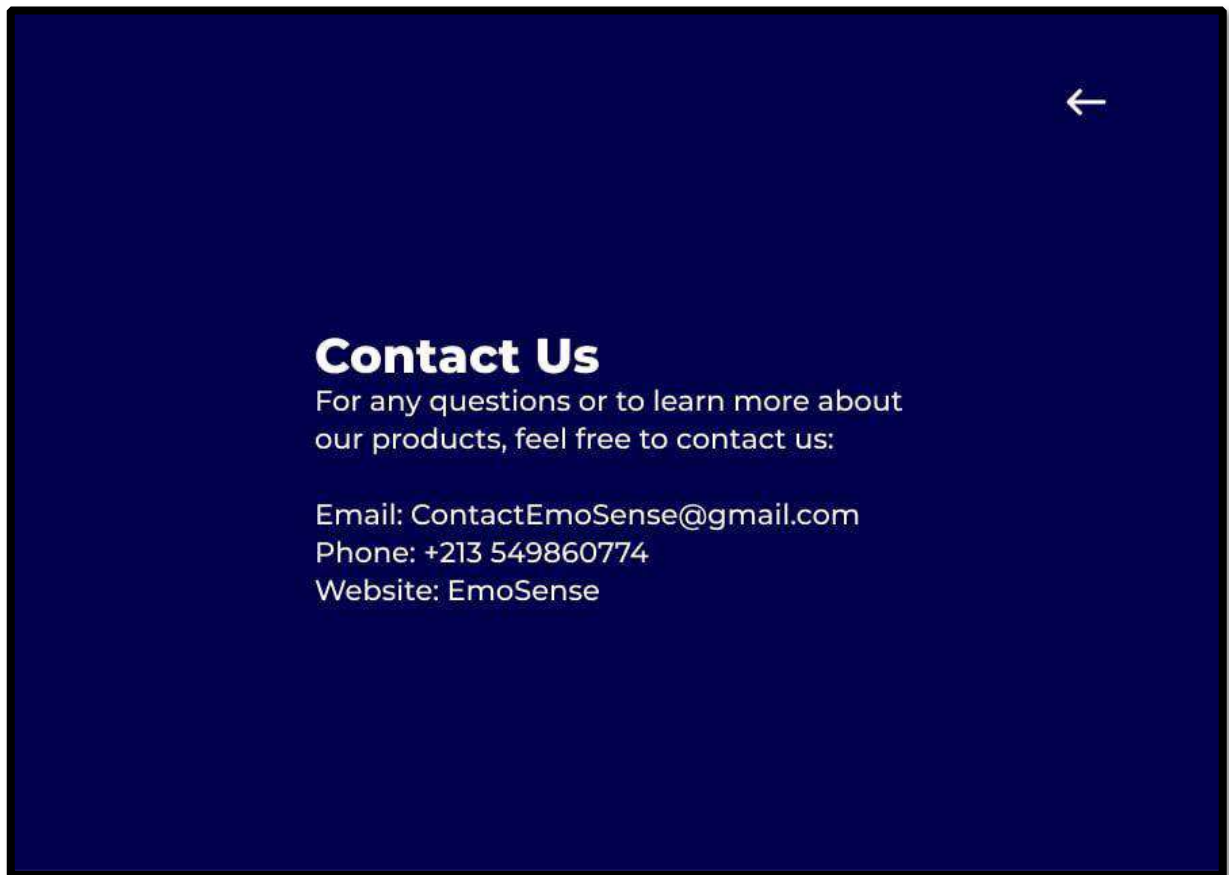


Figure 2 : Page d'information.

En cliquant sur le bouton "Get to Know Us" sur la page d'accueil, vous êtes dirigé vers cette page pour découvrir nos coordonnées et entrer en contact avec nous.

Explication des deux fonctionnalités principaux du programme

Offline (Détection à partir de Vidéos Enregistrées) :

Voici la page que vous aurez quand vous cliquerez sur offline !

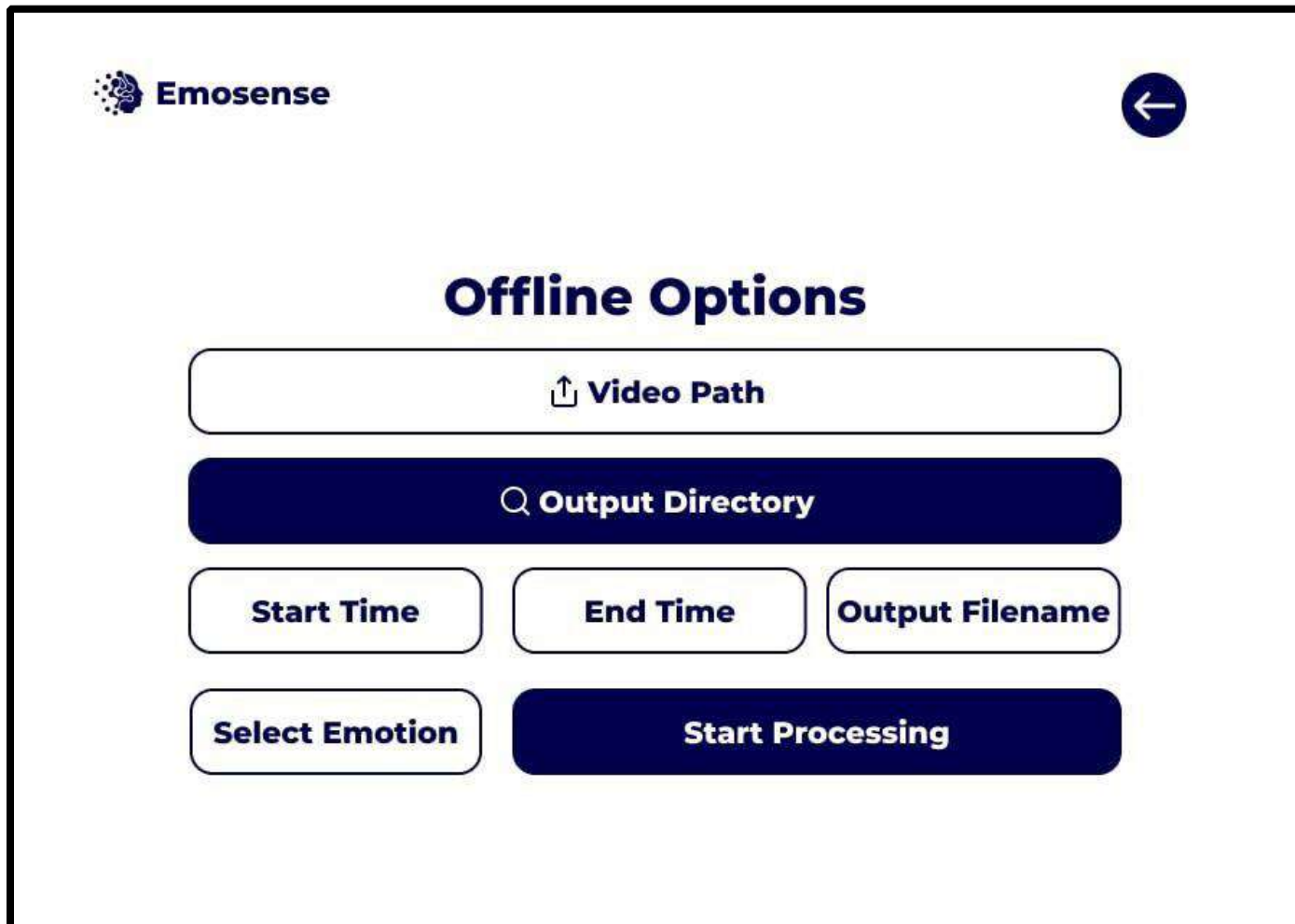


Figure 3 : Menu de la détection offline.

## Sixième axe : Prototype expérimental

Pour utiliser la fonctionnalité Offline de EmoSense, suivez ces étapes simples :

### 1. Importer la Vidéo :

- Cliquez sur le bouton "Vidéo Path" pour importer la vidéo souhaitée depuis votre PC.

### 2. Options de Configuration :

- **Start time/End time** : Choisissez le moment précis de début et de fin de la détection dans la vidéo. Par exemple, vous pouvez démarrer la détection à la 8ème minute et la terminer à la 10ème minute.
- **Output Filename** : Sélectionnez ou enregistrez la vidéo de sortie après le traitement.
- **Select emotion** : Sélectionnez parmi les 7 émotions disponibles celle que vous souhaitez détecter.

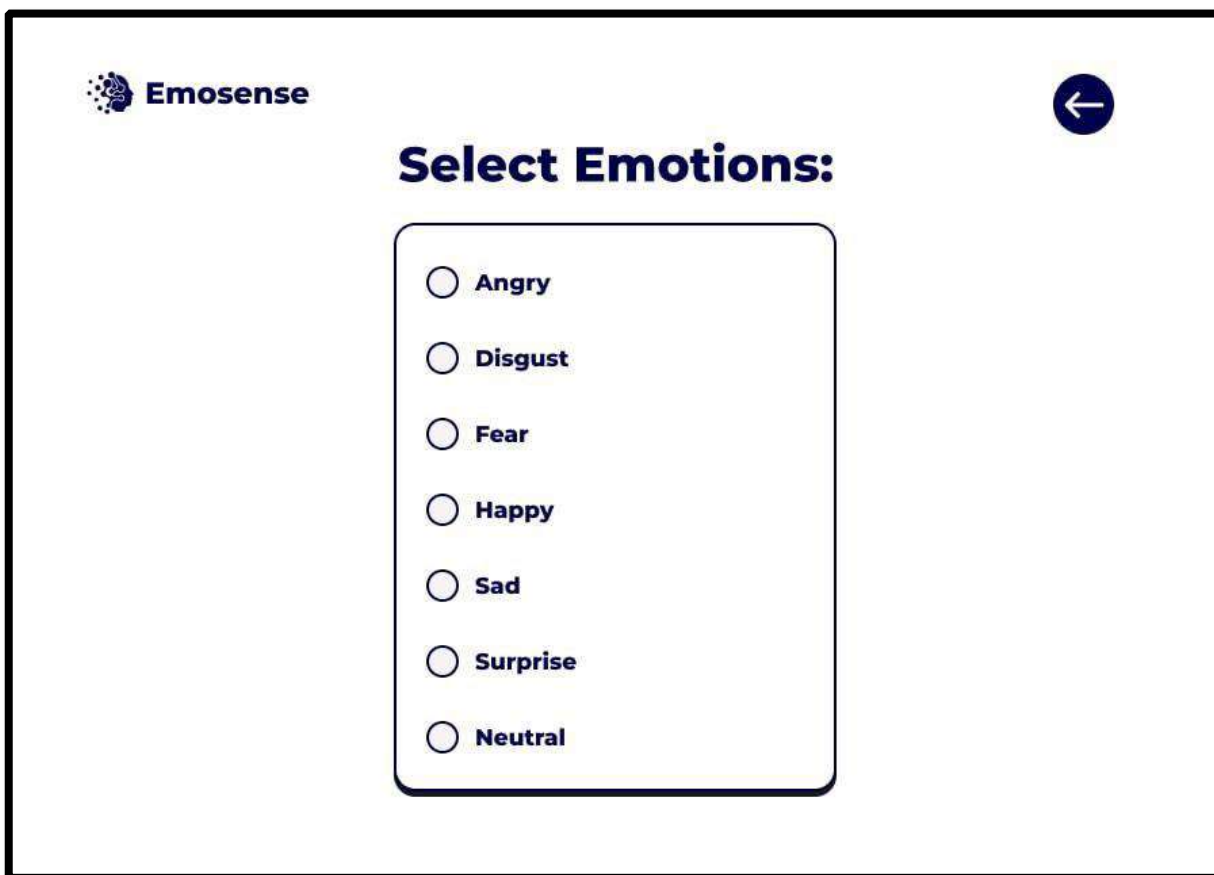


Figure 4 : Choisir les émotions.

### 3. Démarrer le Traitement :

- Une fois toutes les configurations faites, cliquez sur le bouton "Start processing" pour lancer le processus de détection d'émotions.

Live (Détection en temps réel) :

Voici l'interface pour la détection en direct

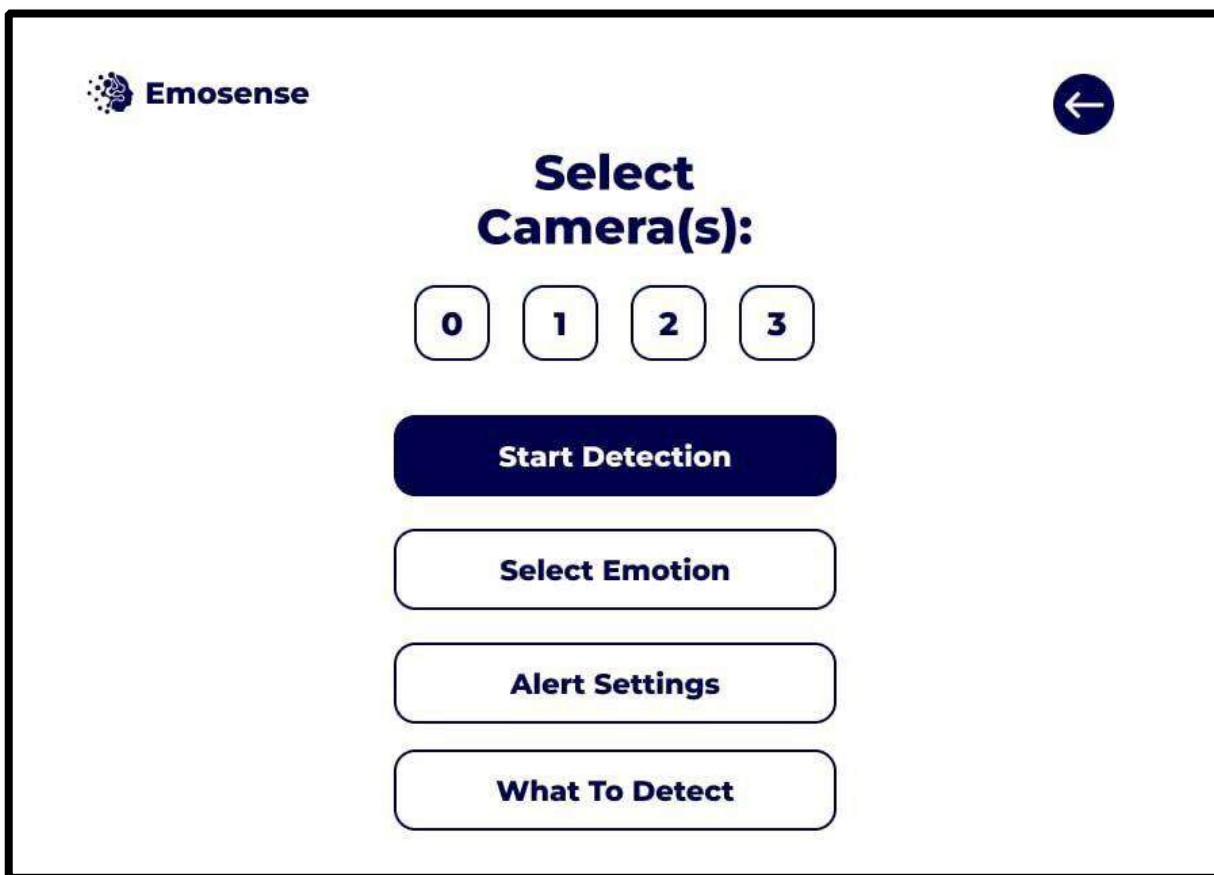


Figure 5 : Menu de la détection en temps réel.

Pour la détection en temps réel avec EmoSense, voici comment procéder :

1. **Interface de Détection en Direct :**
  - Sélectionnez les caméras à utiliser pour la détection en direct.
2. **Options Supplémentaires :**
  1. **Alerte setting :**

C'est pour activer une alerte (Notification) sur l'émotion voulue

Le bouton "Alerte Setting" (Paramètres d'Alerte) permet d'activer les alertes et configurer les paramètres des alertes qui seront générées par le système lorsque certaines émotions ou comportements sont détectés.

Deux options sont disponibles :

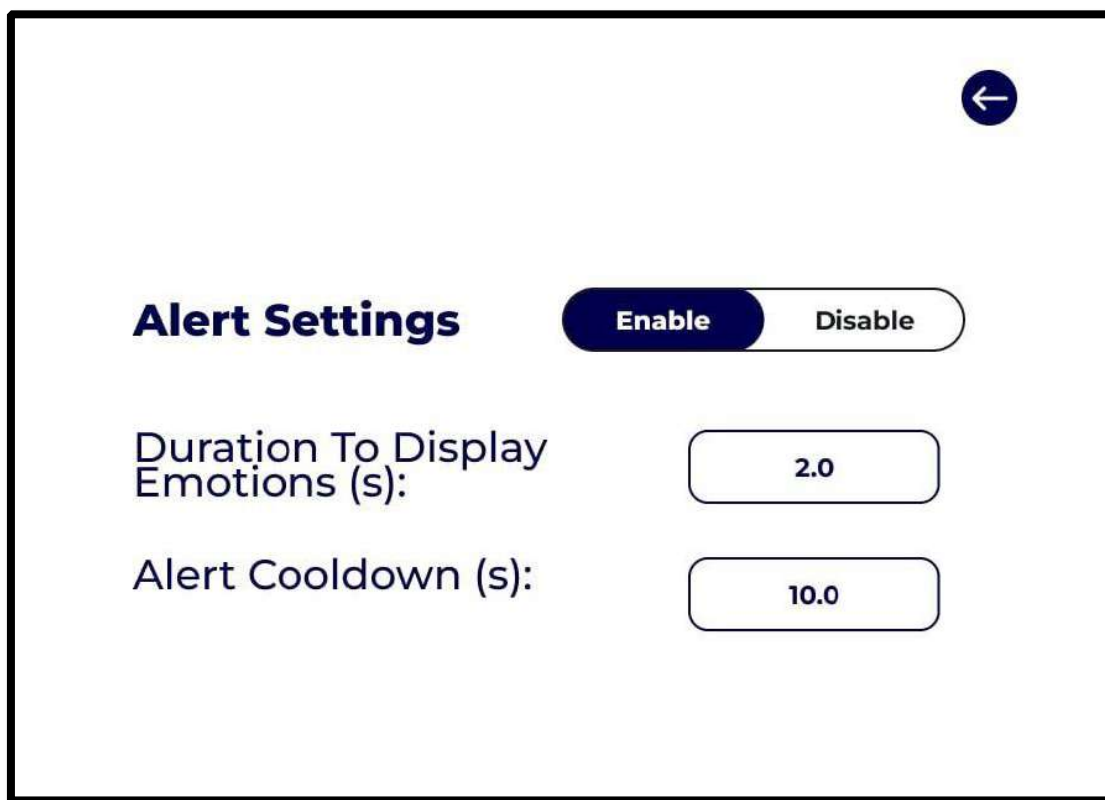


Figure 6 : Paramètres d'alerte.

**Le bouton "Duration to Display Emotions" (Durée d'affichage des émotions) :**

Permet de configurer la durée pendant laquelle les résultats des émotions détectées seront affichés à l'écran après leur détection. Voici une explication détaillée de son fonctionnement et de son utilité :

L'utilisateur peut sélectionner la durée souhaitée à partir d'un menu déroulant, de boutons d'options, ou en entrant une valeur manuellement. Par exemple, les options pourraient inclure 5 secondes, 10 secondes, 30 secondes, etc.

Le bouton peut également être un curseur que l'utilisateur peut glisser pour choisir une durée précise.

Une fois la durée sélectionnée, le programme affichera les résultats de la détection d'émotions à l'écran pour cette durée spécifique après chaque détection.

Si une nouvelle émotion est détectée avant l'expiration de la durée précédente, l'affichage peut être mis à jour immédiatement ou après la durée spécifiée, selon la configuration.

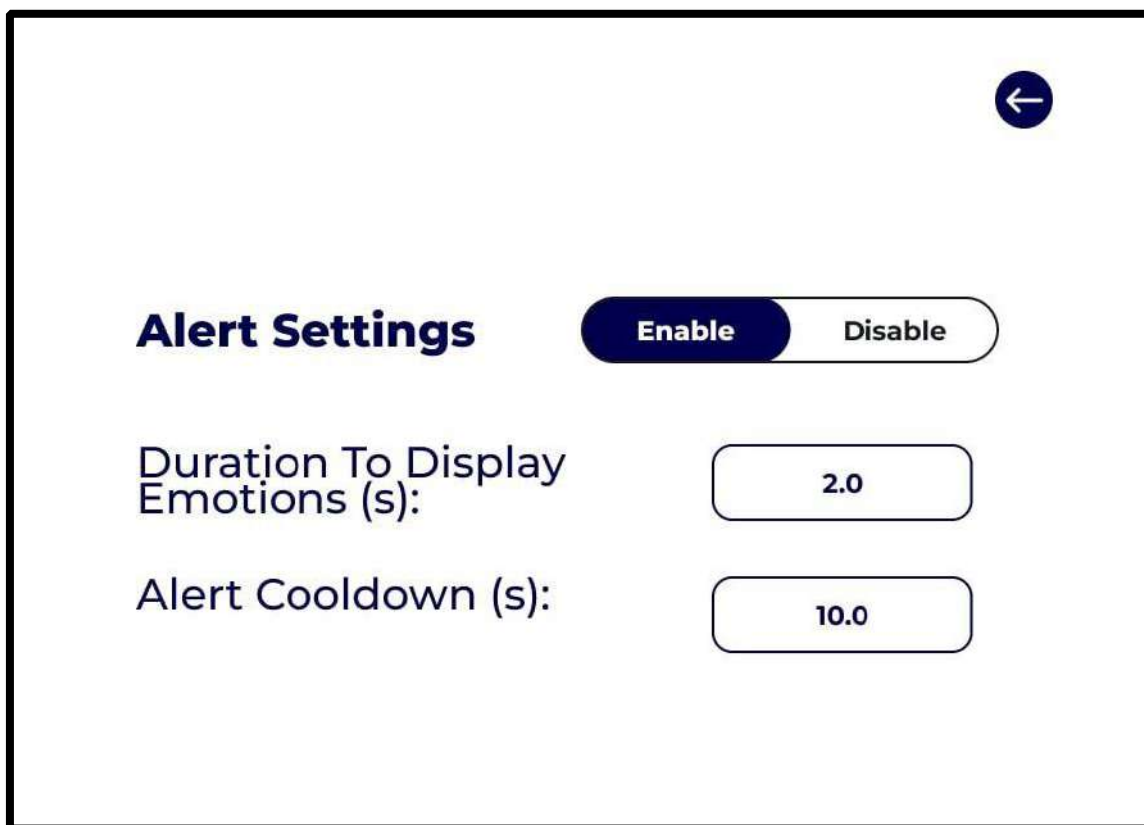


Figure 7 : Paramètres d'alerte.

**Le bouton "Alert Cooldown" (Temps de Réinitialisation des Alertes) :**

Permet de configurer la durée minimale entre deux alertes consécutives pour la même émotion. Cela aide à éviter les alertes excessives et à garantir que les notifications restent significatives et gérables. Voici une explication détaillée de son fonctionnement et de son utilité :

L'utilisateur peut sélectionner la durée de cooldown à partir d'un menu déroulant, de boutons d'options, ou en entrant une valeur manuellement. Par exemple, les options pourraient inclure 1 minute, 5 minutes, 10 minutes, etc.

Une fois la durée sélectionnée, le programme utilisera cette valeur pour limiter la fréquence des alertes pour une même émotion. Si une émotion est détectée et qu'une alerte est envoyée, le système n'enverra pas une autre alerte pour cette émotion tant que la durée de cooldown n'est pas écoulée.

Si une nouvelle émotion différente est détectée pendant le cooldown, le système peut générer une alerte distincte selon la configuration.



## 2. What to detect (Detection mode):

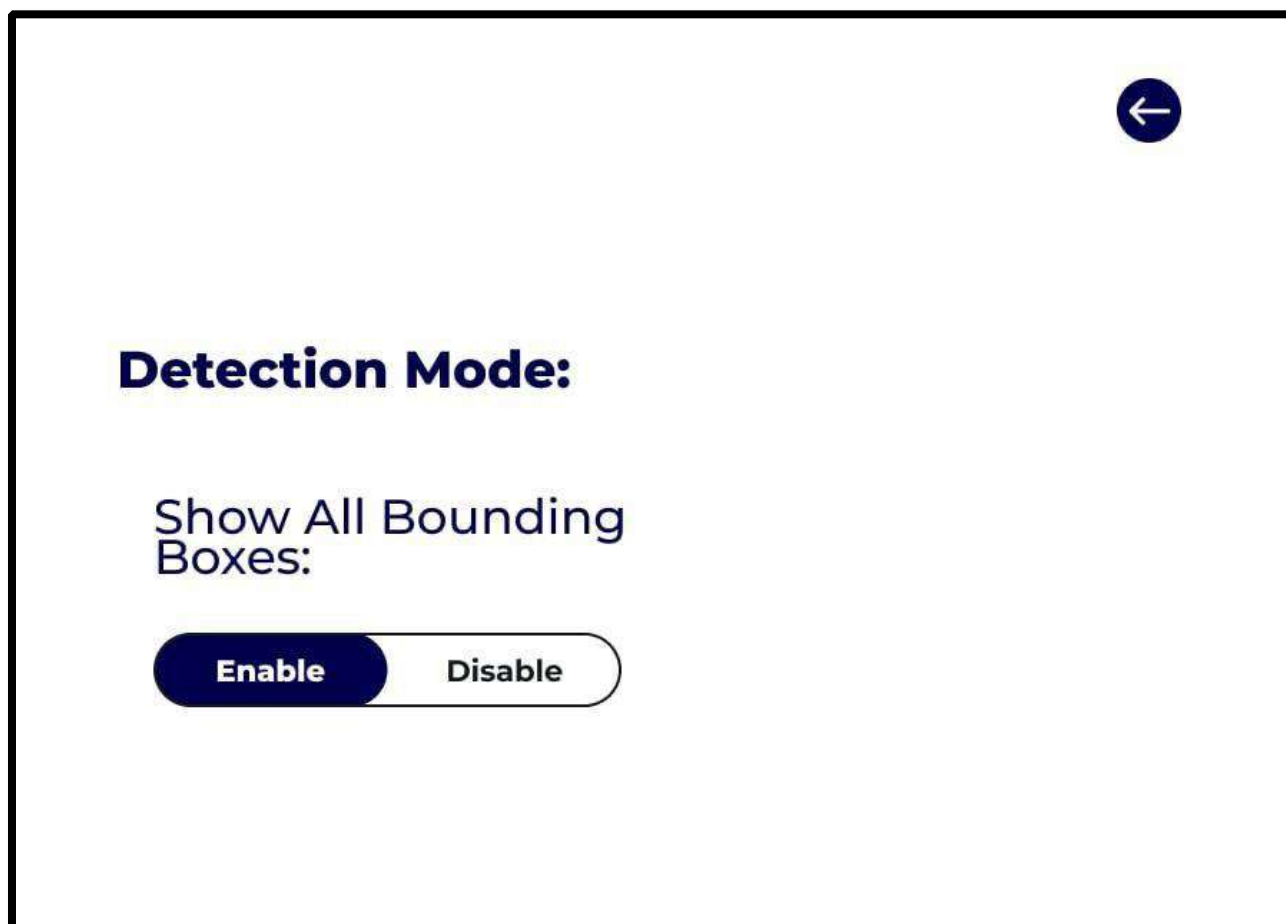


Figure 8 : What to detect.

Le bouton "Détection Mode" permet de configurer le mode de détection des émotions, en offrant la possibilité d'afficher ou de masquer les boîtes de délimitation (bounding boxes) autour des visages détectés. Voici une explication détaillée de son fonctionnement et de son utilité :

### Fonctionnement du Bouton "Detection Mode"

1. **Activation du Mode de Détection :**
  - Lorsque l'utilisateur clique sur le bouton "Detection Mode", une fenêtre ou un panneau de configuration s'ouvre, affichant l'option "Show all bounding boxes" (Afficher toutes les boîtes de délimitation) avec les choix "Enable" (Activer) ou "Disable" (Désactiver).
2. **Sélection de l'Option :**
  - L'utilisateur peut choisir d'activer ou de désactiver l'affichage des boîtes de délimitation en sélectionnant l'option appropriée.
  - En mode "Enable", le programme affichera des boîtes de délimitation autour des visages détectés à l'écran, facilitant ainsi la visualisation des zones où les émotions sont détectées.

## Sixième axe : Prototype expérimental

- En mode "Disable", les boîtes de délimitation ne seront pas affichées, ce qui peut rendre l'interface plus propre et moins encombrée.

En affichant les boîtes de délimitation, les utilisateurs peuvent voir exactement quels visages sont détectés et analysés, ce qui peut être utile pour vérifier l'exactitude de la détection.

La possibilité de désactiver ces boîtes permet de réduire les distractions visuelles et de rendre l'interface plus agréable à utiliser dans certaines situations.

### Démarrage de la détection

Après avoir fait tous les réglages nécessaires dans la fenêtre des paramètres, l'utilisateur peut appliquer ces paramètres et démarrer la détection en cliquant sur le bouton "Start Detection".

Une fois les paramètres appliqués, le programme commence à détecter les visages et les émotions en temps réel selon les configurations définies.

Voici la fenêtre après la détection :

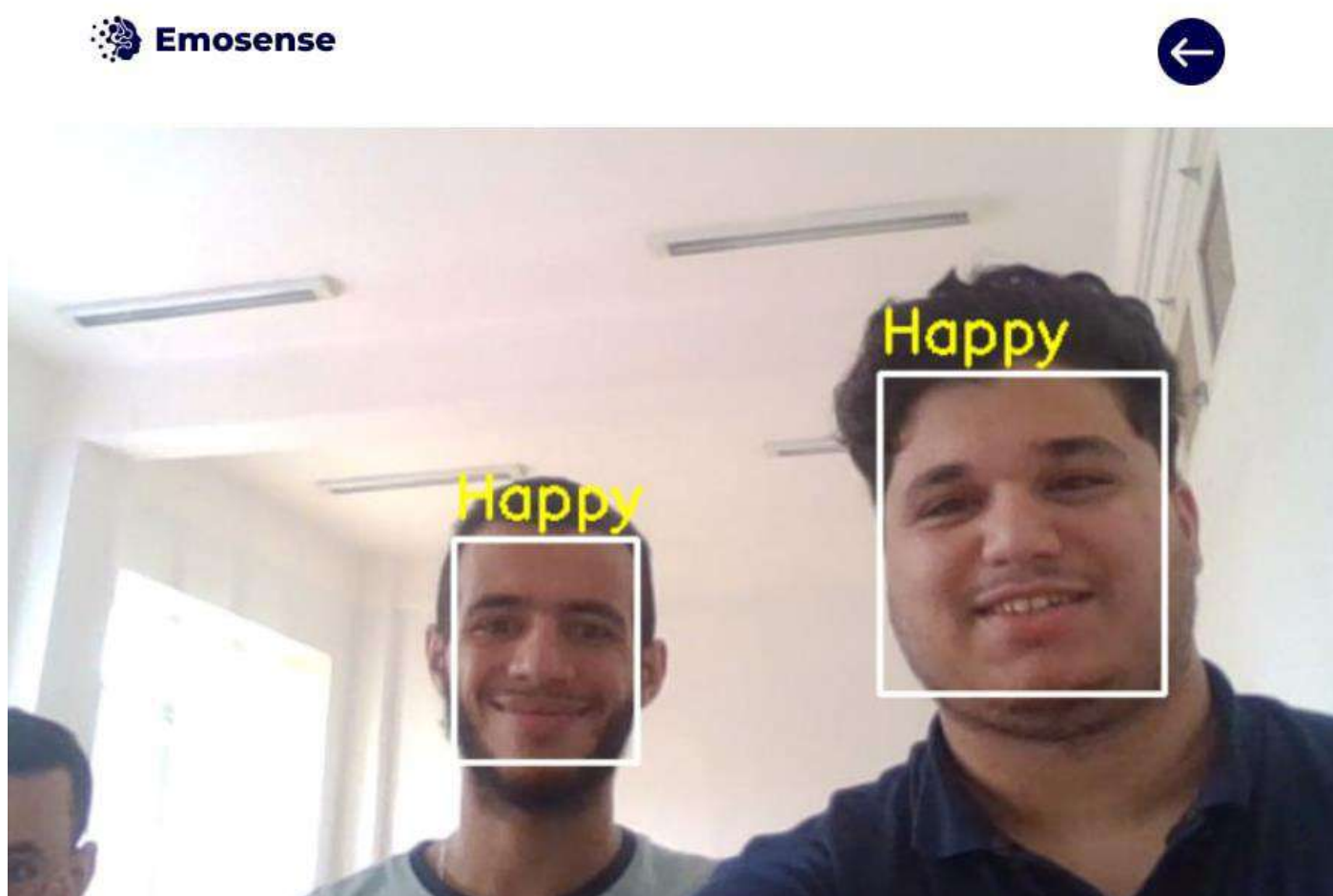


Figure 9 : Résultat de la détection.

**Le septième axe :**

**Business Model**

**Canvas (BMC)**

## Septième axe : Business model canvas (BMC)

- Modèle d'entreprise du projet EmoSense :

Partenaires	Principales activités	Valeur fournie	Relations avec les clients	Segments de clientèle
<ul style="list-style-type: none"> <li>-Incubateur de l'université</li> <li>-Agences de marketing</li> <li>-Entreprises de sécurité et de surveillance</li> <li>-Partenaires sectoriels</li> <li>-Universités et centres de recherche</li> </ul>	<ul style="list-style-type: none"> <li>Développement du logiciel de détection des émotions</li> <li>-Marketing et promotion</li> <li>-Recherche et développement</li> <li>-Vente du système de détection des émotions</li> </ul>	<ul style="list-style-type: none"> <li>-Technologie de détection d'émotions avancée</li> <li>-Amélioration du service client</li> <li>-Sécurité renforcée</li> <li>-Réactivité en temps réel</li> <li>-Avantage concurrence</li> </ul>	<ul style="list-style-type: none"> <li>-Réseaux sociaux</li> <li>-Support technique personnalisé</li> <li>-formations spécialisées</li> <li>démonstrations de produit</li> </ul>	<ul style="list-style-type: none"> <li>-Restaurants et établissements de restauration</li> <li>-Entreprises de sécurité et gestion de la sécurité</li> <li>-Agences gouvernementales</li> <li>-Centres de contact et services à la clientèle</li> <li>-Restaurants et établissements de restauration</li> <li>-Hôpitaux et établissements de santé</li> <li>-Transport et logistique</li> <li>-Hôtellerie</li> </ul>
	<b>Ressources primaires</b>		<b>Canaux</b>	
	<ul style="list-style-type: none"> <li>-Données de formation et d'entraînement</li> <li>-Équipe de développement et d'ingénierie</li> <li>-Équipe support et service client</li> <li>-Infrastructure informatique</li> </ul>		<ul style="list-style-type: none"> <li>-Site web , réseaux sociaux</li> <li>-plates-formes de vente en ligne</li> <li>-Vente direct (déplacement dans les locaux clients)</li> </ul>	
<b>Coûts</b>		<b>Sources de revenus</b>		
<ul style="list-style-type: none"> <li>- Développement et maintenance du logiciel</li> <li>-Support technique et service client</li> <li>-Marketing et promotion</li> <li>-salaires</li> <li>-frais location du bureaux et coûts services publics</li> <li>-Abonnement aux logiciels et outils nécessaires a l'exploitation de l'entreprise</li> <li>-frais des déplacements</li> </ul>		<ul style="list-style-type: none"> <li>-Ventes de licences de logiciel</li> <li>-Abonnement mensuels ou annuels</li> <li>-Service de personnalisation</li> <li>-Support et maintenance</li> </ul>		

Tableau 6 : BMC.