



MACHINE LEARNING CHALLENGE

Rapport MLChallenge

Élèves :

Bastien TAROT

Raphael MONNIER

Tanguy POINGT

Allan PLANCHENAULT

Enseignant :

Thibault GEOFFROY

7 septembre 2024

Table des matières

1	Introduction	2
1.1	Objectifs du Projet	2
1.2	Etat de l'art	2
2	Analyse des Données	2
2.1	Description des Données	2
2.2	Exploration des Données	3
3	Prétraitement et Extraction des Features	4
3.1	Prétraitement des Données	4
3.2	Extraction des Features	5
3.3	Modèles et Méthodologies	5
4	Modèle(s) choisi(s)	5
4.1	Paramétrage et Entraînement	5
4.2	Expérimentations	6
4.3	Résultats	6
5	Évaluation des Modèles	6
5.1	Analyse des Résultats	6
5.2	Discussion et Limites	6
6	Discussion des Résultats	6
6.1	Limites	6
6.2	Conclusion et Perspectives	6
7	Bibliographie / Références	6

1 Introduction

La reconnaissance des émotions est un domaine en pleine expansion qui suscite l'intérêt de nombreux chercheurs en raison des défis complexes qu'il pose. Malgré les avancées réalisées, ce problème reste encore partiellement résolu. Ce champ d'étude est vaste, mais pour ce projet, nous nous concentrerons sur les émotions dites de base selon le modèle d'Ekman : la joie, la colère, le dégoût, la tristesse, la peur et la surprise, en ajoutant également l'absence d'émotion, c'est-à-dire l'état « neutre ».

1.1 Objectifs du Projet

Pour développer un système de reconnaissance des émotions en utilisant une approche de machine learning, il est généralement recommandé de suivre un pipeline structuré.

Dans le cadre de ce projet, après la mise en place des différents sujets d'introduction nous aborderons les grands thèmes suivants :

1.Prétraitement et Extraction des Features : Les étapes nécessaires pour préparer les données et extraire les caractéristiques pertinentes pour la reconnaissance des émotions.

2.Modèles Choisis : Les différents modèles de machine learning sélectionnés pour cette tâche et les raisons de leur choix.

3.Évaluation des Modèles : Les méthodes et les métriques utilisées pour évaluer la performance des modèles de reconnaissance d'émotions.

4.Discussion des Résultats : Une analyse critique des résultats obtenus, mettant en évidence les points forts et les limitations des approches adoptées.

Cette structure permettra de couvrir de manière exhaustive les aspects clés du développement d'un système de reconnaissance des émotions basé sur l'apprentissage automatique.

1.2 Etat de l'art

2 Analyse des Données

2.1 Description des Données

Le visage manifeste près de 2/3 des émotions chez un humain. [1] Les zones les plus démonstratives sont principalement situées au niveau des lèvres, des sourcils, des yeux mais également des yeux et du nez (malgré qu'ils sont moins significatifs sur leurs représentations). [1] Les données sont représentées dans un CSV. Il y a au total 978 observations, toutes sous le même format. Chacune d'entre elles est représentée sur une même ligne et 138 colonnes. La première colonne est réservée pour l'ID de l'image. La deuxième colonne représente simplement le label de l'observation. Pour les 136 colonnes restantes, elles sont

séparées en deux sous-parties. En effet, les 64 premières valeurs et les 64 dernières valeurs sont respectivement les coordonnées en X et en Y de chaque points encadrant le visage et plus précisément les zones mentionnées plus tôt.

2.2 Exploration des Données

Nous allons dans un premier temps nous familiariser avec les données afin de comprendre leur structure mais également leur distribution. Nous remarquons que notre jeu de données est de manière générale bien réparti comme présenté sur la figure 1. Nous notons tout de même une quantité moins importante de données pour les labels 'happy' et 'neutral'.

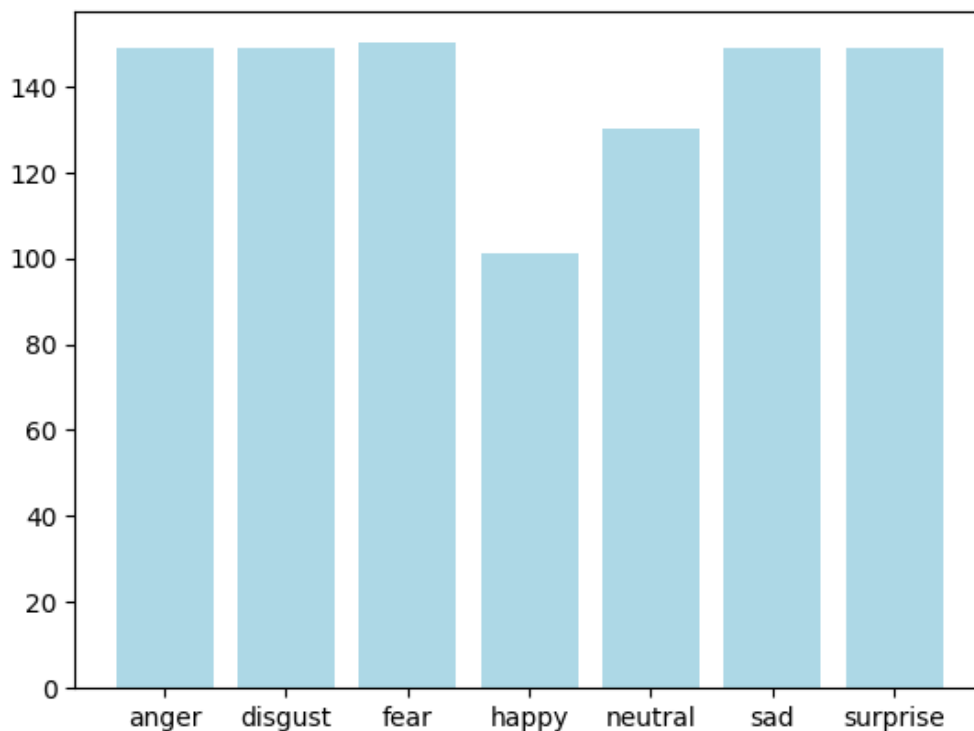


FIGURE 1 – Distribution des émotions dans le jeu de données

Comme précisé dans la partie précédente, nous disposons d'un fichier CSV contenant les coordonnées des points clés du visage. Pour nous familiariser avec les données, nous allons visualiser quelques visages marqués de ces points clés comme présenté sur les figures 2 et 3 :

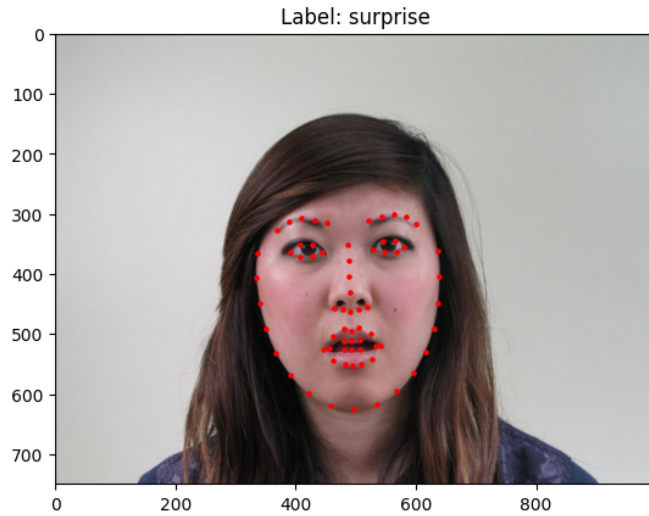


FIGURE 2 – Visage surpris avec ses points clés

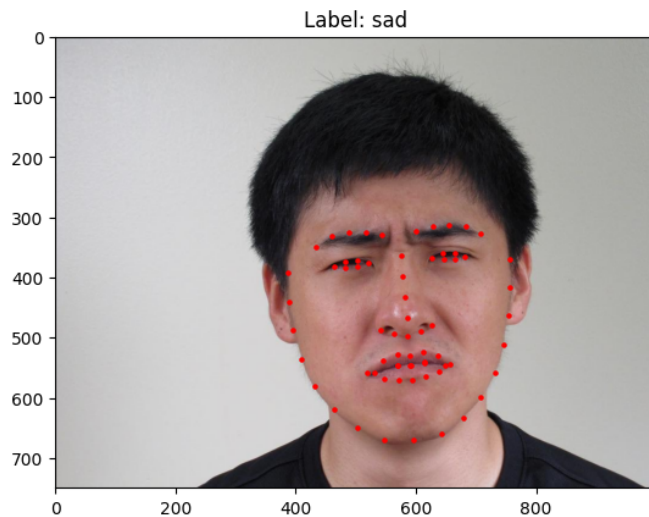


FIGURE 3 – Visage triste avec ses points clés

3 Prétraitement et Extraction des Features

3.1 Prétraitement des Données

Comme spécifié auparavant, nous disposons donc des 138 caractéristiques représentant les coordonnées des points repères du visage. Avant de pouvoir construire notre modèle, il est essentiel de normaliser ces données. Nous les normalisons donc selon la manière suivante [2] :

$$X = \frac{(X - X_{mean})}{X_{std}} \quad (1)$$

En parallèle, nous devons également appliquer un traitement sur les labels des don-

nées. En effet ceux-ci correspondent à une variable qualitative parmi les 7 émotions. Nous devons donc l'encoder en entier entre 0 et 6, chaque entier correspondant à une émotion en particulier. Une fois ces traitements appliqués, nous pouvons commencer à réfléchir au modèle à utiliser.

Nous avons également tenter de travailler avec les images. En effet, les repères de visage présentent tout de même plusieurs inconvénients et une partie de l'information est perdue. Les images quant à elle permettent d'avoir les textures, les rides du visage et autres informations susceptibles d'aider dans la reconnaissance d'émotions. Néanmoins, de nouveau défis surviennent lorsque nous utilisons les images, comme la luminosité ou encore l'orientation du visage. Nous supposons également que la construction du modèle ainsi que sa complexité seront décuplées.

3.2 Extraction des Features

3.3 Modèles et Méthodologies

4 Modèle(s) choisi(s)

4.1 Paramétrage et Entraînement

L'utilisation d'un modèle Support Vector Machine (SVM) n'est pas nouvelle est a déjà prouvé ses performances dans la reconnaissance d'émotions [2]. Nous avons donc décidé d'utiliser ce modèle couplé à une GridSearch afin de trouver les hyperparamètres optimaux pour le modèle.

GridSearch est une méthode utilisée afin de trouver les hyperparamètres optimaux pour un modèle. Elle prend en paramètre une grille d'hyperparamètres et va tester ensuite chaque combinaison pour faire ressortir la meilleure d'entre elle. Pour vérifier la performance des paramètres, GridSearch utilise la méthode "k-fold cross-validation" qui consiste à diviser les données d'entraînement en plusieurs sous-ensemble. Le modèle est entraîné ensuite sur l'un de ces sous-ensemble et est évalué sur le reste des données. Ce processus est répété k fois et permet d'éviter le surapprentissage du modèle. La meilleure combinaison d'hyperparamètres est ensuite utilisée pour entraîner le modèle sur l'ensemble des données d'entraînement. Il est important de noter que même si cette méthode permet de trouver les hyperparamètres optimaux, elle reste tout de même très gourmande en temps et en ressources.

Nous avons donc décidé d'utiliser cette méthode afin d'obtenir les meilleurs hyperparamètres pour le modèle SVM.

4.2 Expérimentations

4.3 Résultats

5 Évaluation des Modèles

5.1 Analyse des Résultats

5.2 Discussion et Limites

6 Discussion des Résultats

6.1 Limites

6.2 Conclusion et Perspectives

7 Bibliographie / Références

Références

- [1] B. Ko, “A Brief Review of Facial Emotion Recognition Based on Visual Information,” vol. 18, no. 2, p. 401.
- [2] L. Kalapala, H. Yadav, H. Kharwar, and S. Susan, *Facial Expression Recognition from 3D Facial Landmarks Reconstructed from Images*.