

# Projet HAI927I

Rapport de la semaine 2 – Génération de visages avec émotions

Lucas Jalbaud et Bastian Langouet

Novembre 2025

## Introduction

Ce document présente les avancées réalisées lors de la deuxième semaine du projet d’Image, consacré à la génération de visages avec émotions. Nous y exposons également les objectifs fixés pour la semaine à venir.

## 1 Objectifs de la semaine

Nos objectifs pour cette seconde semaine étaient de nous familiariser avec l’encodeur proposé par DeepFace. Le but est de récupérer les parties de code pertinentes pour notre projet et de les adapter afin de mieux détecter les émotions, suite à des problèmes de détection de certaines émotions, comme la peur.

## 2 Reprise de l’encodeur

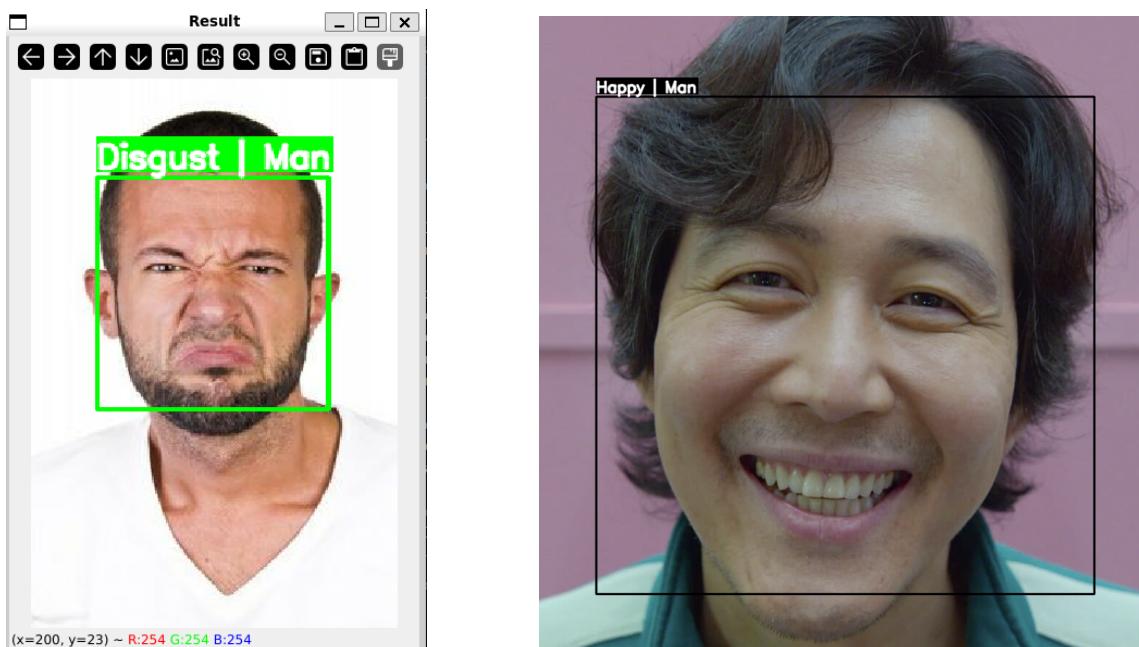
Nous avons, dans un premier temps, analysé le code de DeepFace afin de comprendre comment l’encodeur analyse les images et renvoie les émotions. Ceci nous a permis de programmer notre propre encodeur, qui contient plusieurs réseaux CNN :

- Reconnaissance de visage : cette première couche permet à notre encodeur de détecter les visages sur une image. Cette analyse se base sur le modèle OpenCV pour la détection de visages. On obtient alors une boîte englobante du visage à analyser, ce qui nous permet de rogner l’image et d’aligner les visages dans le cadre. Ainsi, nous envoyons une information pré-traitée aux couches suivantes.
- Reconnaissance de genre : cette seconde couche permet à l’encodeur d’estimer le genre de la personne sur l’image. Pour cela, nous utilisons la reconnaissance faciale basée sur le modèle VGGFace.
- Reconnaissance des émotions : pour finir, une fois l’image pré-traitée, nous passons à la classification des émotions. Elle s’effectue à l’aide de poids pré-calculés pendant la phase d’entraînement du modèle, avec trois couches de convolution.

### 3 Problème de l'encodeur

Avec cette implémentation, nous avons remarqué que le modèle pouvait aisément reconnaître les émotions telles que la joie ou le dégoût, mais avait beaucoup plus de mal à faire la différence entre la surprise et la peur. Cela provient du fait que la surprise et la peur ont de nombreuses caractéristiques communes (dans le dataset), et l'entraînement du modèle n'a sûrement pas été optimisé pour capter efficacement les différences. Nous avons donc ensuite consacré notre temps à reprogrammer l'entraînement du modèle d'émotion à l'aide de la base d'images FER-2013. Cependant, après l'entraînement, nous obtenons des résultats similaires. Après recherche, nous avons pu constater que la base d'images comprend de grandes disparités dans le nombre d'images par classe, ainsi qu'une très faible qualité d'image, entraînant un manque d'efficacité de l'apprentissage. Nous avons donc essayé de fine-tuner le modèle d'apprentissage : nous constituons un premier modèle classique, puis un second qui se concentre sur les classes sous représentées et qui se base sur le premier modèle pour accélérer l'apprentissage. Les résultats ne semblent pas convaincants. La base d'image semble être le principal élément problématique.

### 4 Résultats obtenus par notre propre encodeur



### 5 Objectifs de la semaine prochaine

Pour la semaine à venir, nous avons pour objectif de terminer l'entraînement de notre encodeur afin d'obtenir de meilleurs résultats. En parallèle, nous allons commencer le développement du décodeur pour la mise en place de la VQ-VAE.

## Références

- [1] Serengil, S. \*DeepFace : A Lightweight Face Recognition and Facial Attribute Analysis Framework for Python.\* Disponible sur : <https://github.com/serengil/deepface>
- [2] Pythonia Formation. \*Apprendre la computer vision : reconnaissance faciale en Python.\* Disponible sur : <https://www.pythoniaformation.com/blog/tutoriels-python-par-categories/apprendre-la-computer-vision/reconnaissance-faciale>
- [3] EWA Direct. \*Proceedings of ACE — Article 16681.\* Disponible sur : <https://www.ewadirect.com/proceedings/ace/article/view/16681>
- [4] FER-2013. \*Banque d'images pour l'entraînement.\* Disponible sur : <https://www.kaggle.com/datasets/msambare/fer2013?resource=download>