

PROJET TRAITEMENT ET ANALYSE D'IMAGES

RECONNAISSANCE DE VISAGES



Date : 7 février 2024

Professeur responsable : Vozel Benoit

Equipe :

- Antoine MAURAIS
- Sylvain MESTRE

Résumé

Le projet consistait à développer un système de reconnaissance de visages en deux phases : la première explorant la reconnaissance par attributs via des méthodes comme les cascades de Haar et les Eigenfaces, et la seconde utilisant un réseau neuronal convolutionnel pour classer les images en hommes et femmes. Bien que des résultats prometteurs aient été obtenus, le modèle présentait des limitations, notamment en termes de généralisation sur la base de la longueur des cheveux. Des améliorations potentielles ont été identifiées, telles que l'affinement des données d'entraînement.

Table des Matières:

Liste des Figures et des Tableaux.....	3
I. Introduction.....	4
II. Phase 1 : Détection et Reconnaissance par Attributs.....	5
III. Phase 2 : Reconnaissance par Réseaux de Neurones.....	7
Choix de l'architecture du réseau neuronal et justification.....	7
Détails du processus d'apprentissage et choix des hyperparamètres.....	7
Résultats obtenus et analyse.....	7
Problème rencontrés.....	9
IV. Conclusion et Perspectives.....	10

Liste des Figures et des Tableaux

Figure 1 : Analyse de l'image lors de la phase 1.....	6
Figure 2 : Distances entre le nouveau vecteur et chaque vecteur moyen des classes.....	7
Figure 3 : Résultat obtenu lors de la phase 1.....	7
Figure 4 : Explication de nos résultats.....	9
Figure 5 : Résultats problématiques.....	10

I. Introduction

La reconnaissance de visages est une problématique fondamentale dans le domaine de la vision par ordinateur, consistant à identifier et vérifier les personnes présentes dans une image ou une vidéo à partir de leurs visages. Bien que cette tâche soit aisée pour l'œil humain, elle reste un défi majeur pour les systèmes informatiques, jusqu'à l'avènement récent de l'apprentissage profond.

Ce projet a pour objectif de mettre en place un système de reconnaissance de visage. Il se déroule en deux phases, chacune visant à obtenir des résultats exploitables de reconnaissance de visages, voire d'expressions. Dans la première phase, différentes solutions de reconnaissance d'attributs (forme, éléments clés) dans une image seront explorées. Ensuite dans la seconde phase nous examinerons les images à l'aide d'approches d'apprentissage profond.

Ce rapport est structuré en 4 parties majeures. La première partie présentera la phase 1 qui concerne la détection et la reconnaissance par attributs. Ensuite nous aborderons la deuxième phase avec la reconnaissance à l'aide de réseau de neurones. Puis nous passerons en revue les résultats obtenus afin de voir si ceux-ci correspondent à nos attentes. Enfin nous concluons ce rapport en rappelant les points majeurs.

II. Phase 1 : Détection et Reconnaissance par Attributs

La première phase du projet consistait à détecter et reconnaître les visages, et dire à quelle classe appartient l'image, en utilisant des attributs spécifiques avec une approche de classification supervisée.

Pour la détection des attributs, deux méthodes étaient proposées : les cascades de Haar et les Eigenfaces, toutes deux implémentées avec OpenCV, une bibliothèque de programmation open source spécialisée dans la vision par ordinateur. Les cascades de Haar utilisent des features de Haar pour détecter les objets, tandis que les Eigenfaces sont une approche basée sur l'analyse en composantes principales pour la reconnaissance faciale.

L'utilisation des API suivantes était recommandée pour mettre en œuvre ces méthodes :

- OpenCV Homepage : <https://opencv.org/>
- OpenCV GitHub Project : <https://github.com/opencv/opencv>
- matplotlib.patches.Rectangle et Circle pour la visualisation des détections
- Documentation OpenCV pour la détection de visages avec les cascades de Haar, l'entraînement des classificateurs en cascade, et l'implémentation des Eigenfaces.

Dans le cadre de notre projet, nous avons débuté par la détection d'un élément spécifique du visage, choisissant les yeux comme point de départ. Suite à l'identification des yeux, nous avons vectorisé cette zone de l'image, la convertissant en une série de valeurs numériques. Ce processus a été réalisé pour 30 catégories d'images, nous permettant ainsi de déterminer une moyenne des vecteurs pour chaque catégorie.

Nous avons ensuite utilisé ces informations pour évaluer la proximité entre le vecteur nouvellement calculé et les vecteurs moyens des catégories préétablies, facilitant l'identification de la catégorie la plus proche.

Il convient de mentionner que notre méthode pourrait être améliorée de plusieurs manières. Par exemple, intégrer d'autres parties du visage dans le processus de vectorisation pourrait améliorer la précision de nos résultats. Une stratégie de comparaison et d'évaluation des résultats pour différentes zones du visage permettrait une estimation plus exacte et fiable de la catégorie adéquate.

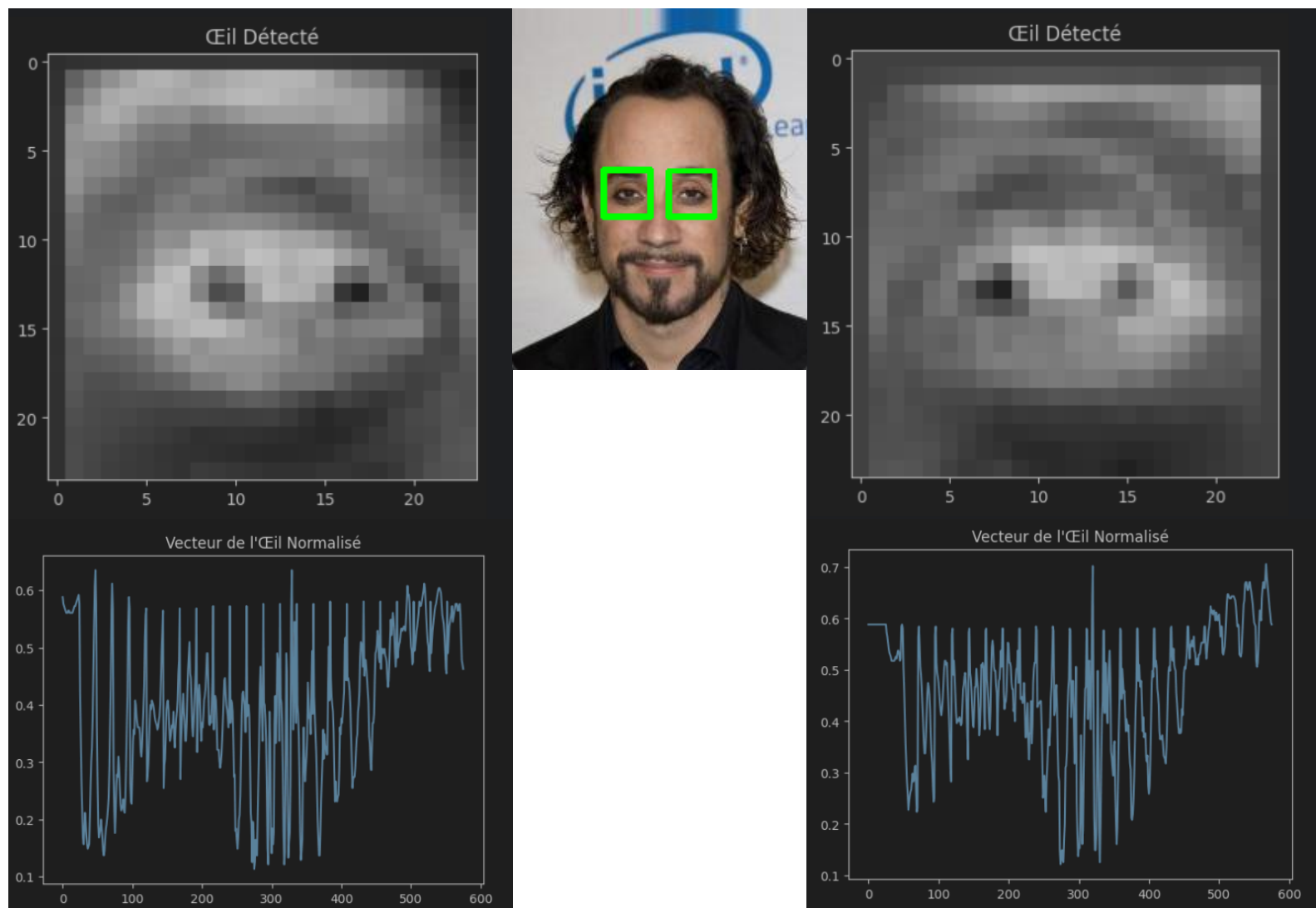


Figure 1 : Analyse de l'image lors de la phase 1

Sur les images ci-dessus, nous pouvons observer les différents yeux détectés et la mise en vecteur de chaque œil.

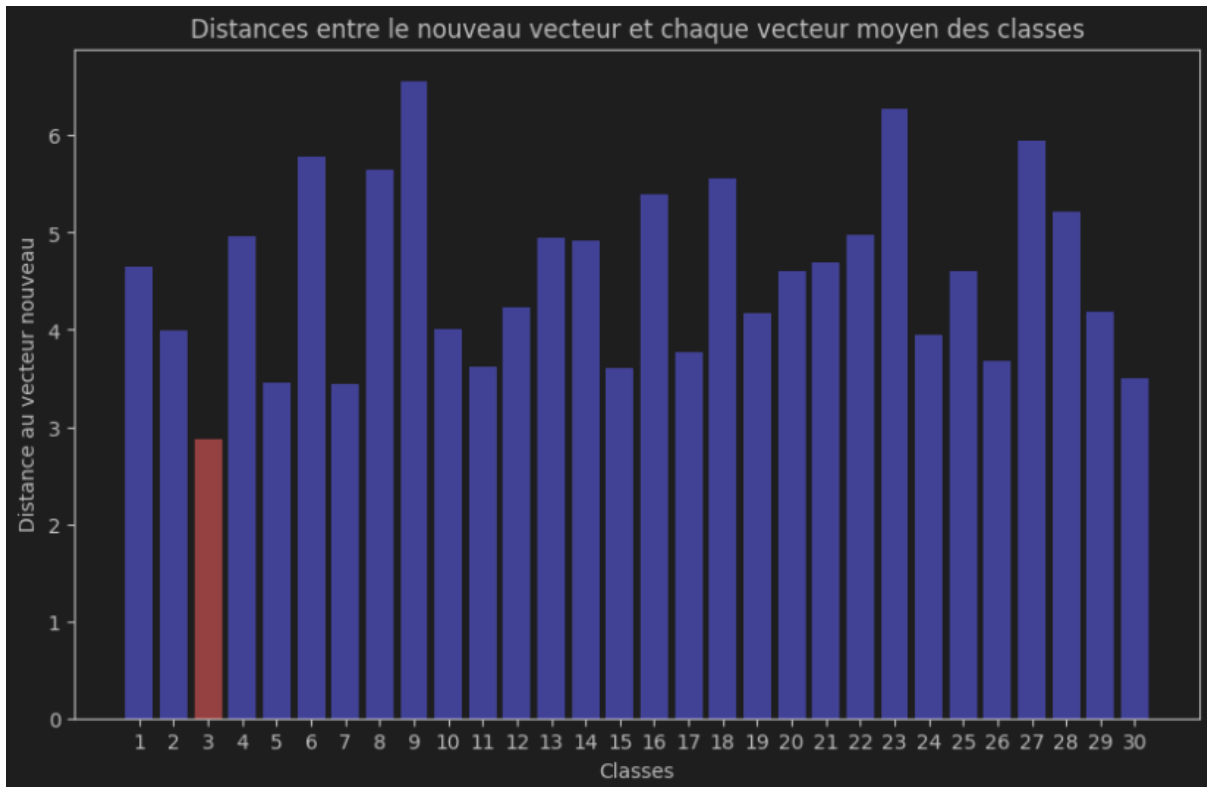


Figure 2 : Distances entre le nouveau vecteur et chaque vecteur moyen des classes

Le graphique illustre les distances calculées entre un nouveau vecteur caractéristique (de l'image d'entrée) et les vecteurs moyens de chaque classe, on retrouve les valeurs dans le fichier 'mean_vectors.json'. Les barres représentent la distance euclidienne entre le vecteur caractéristique de l'image soumise et chaque vecteur moyen, pour les 30 premières classes issues du fichier 'sorted_by_label.json'. Le vecteur moyen le plus similaires au vecteur de l'image analysée sont mis en évidence en rouge, signalant la classe la plus proche par rapport à l'image entrée.

Ensuite, une image aléatoire de la classe trouvée est affichée pour confirmer le résultat.

Avec l'image qu'on a entré au début, on peut trouver le résultat suivant :

La classe prédite pour la nouvelle image est : 3

Figure 3 : Résultat obtenu lors de la phase 1



III. Phase 2 : Reconnaissance par Réseaux de Neurones

Choix de l'architecture du réseau neuronal et justification

Nous avons adopté une architecture de réseau neuronal convolutionnel (CNN) pour classer les images en deux catégories : les hommes et les femmes. Les réseaux neuronaux convolutifs sont très efficaces pour les tâches de reconnaissance d'images en raison de leur capacité à capturer les hiérarchies spatiales dans les images. Le modèle a été construit à partir de zéro et se compose de couches convolutives avec activation ReLU, de couches de mise en commun maximale, d'une couche d'aplatissement et de couches entièrement connectées avec dropout pour la régularisation.

Détails du processus d'apprentissage et choix des hyperparamètres

Des techniques d'augmentation des données telles que la rotation, les décalages largeur/hauteur, le cisaillement, le zoom et le retournement horizontal ont été appliquées pour atténuer le surajustement et améliorer la généralisation du modèle sur des données limitées. Nous avons utilisé 'adam' comme optimiseur en raison de ses capacités d'apprentissage adaptatif, 'binary_crossentropy' comme fonction de perte pour la classification binaire, et 'accuracy' comme métrique d'évaluation.

Résultats obtenus et analyse

Le modèle a été entraîné pendant 15 époques avec un arrêt précoce et une réduction du taux d'apprentissage sur le plateau pour éviter le surajustement. Le processus d'apprentissage s'est appuyé sur des rappels pour sauvegarder le meilleur modèle sur la base de la précision de validation.

Avec notre méthode basée sur l'apprentissage profond, nous avons obtenu des résultats prometteurs lors de la phase de test. Il est important de comprendre l'interprétation de nos résultats : un score tendant vers 1 indique un haut niveau de confiance que l'individu est un homme. Inversement, un score proche de 0 traduit un haut niveau de confiance que l'individu est une femme, comme illustré dans l'image ci-dessous.

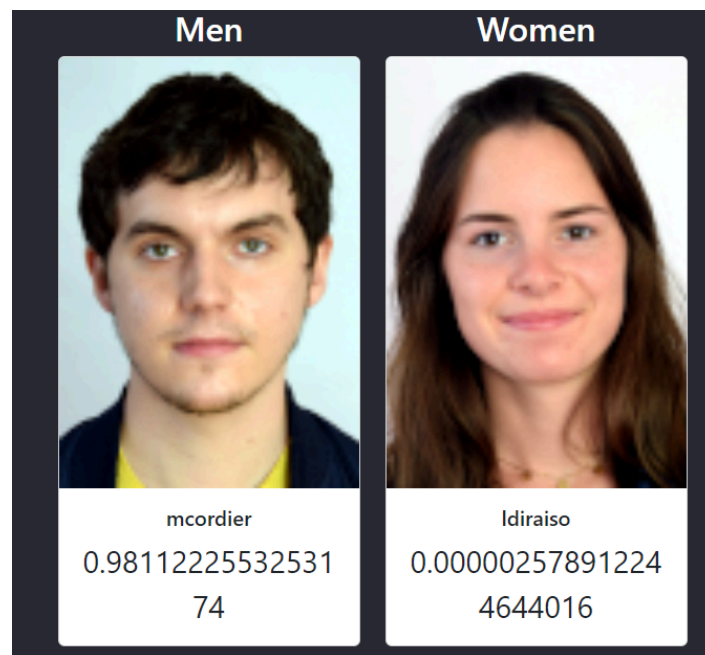


Figure 4 : Explication de nos résultats

Pour analyser nos résultats plus efficacement, nous avons créé une petite application web en React. Cette application permet de charger notre fichier de résultats et de les afficher directement. En développant davantage cette idée, il serait possible d'implémenter une fonctionnalité permettant d'envoyer des images à notre modèle via l'application web et d'en afficher les résultats immédiatement.

L'application web n'est actuellement pas accessible, car nous devons manuellement fournir à l'application les fichiers de résultats et les photos testées.

Problème rencontrés

En effet, nous n'avons pas eu l'opportunité de consacrer suffisamment de temps à l'analyse des données fournies à notre modèle. Par conséquent, nous remarquons que le modèle tend à faire des généralisations assez approximatives.

Nos observations révèlent que le modèle accorde une importance considérable à la longueur des cheveux.

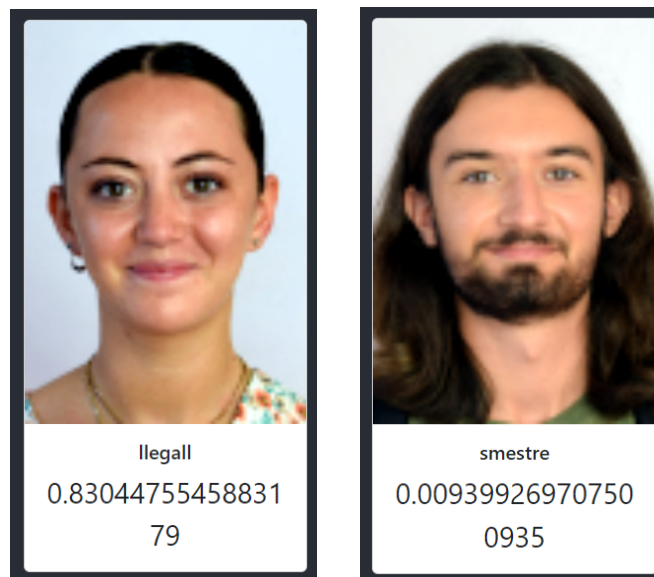


Figure 5 : Résultats problématiques

Comme il est visible sur les images mentionnées précédemment, une femme aux cheveux attachés est classifiée comme un homme, tandis qu'un homme aux cheveux longs est identifié comme une femme.

Cette problématique pourrait être aisément résolue en affinant les données fournies à notre modèle. Une solution serait d'inclure davantage de femmes aux cheveux attachés et d'hommes aux cheveux longs dans notre jeu de données.

IV. Conclusion et Perspectives

Ce projet de reconnaissance de visages a été une expérience enrichissante, offrant un aperçu précieux des défis et des possibilités dans le domaine de l'analyse d'images et de l'apprentissage automatique. Bien que nous ayons rencontré des difficultés, notamment en termes de temps disponible pour peaufiner notre modèle et de limitations dans l'analyse des données, les résultats obtenus ouvrent la voie à des améliorations significatives et à l'exploration de nouvelles méthodes.

Nos travaux ont mis en évidence l'importance d'une sélection et d'un traitement méticuleux des données pour améliorer la précision des modèles d'apprentissage profond. La question de la généralisation basée sur des attributs tels que la longueur des cheveux souligne le besoin de diversifier et d'équilibrer les jeux de données pour éviter les biais et augmenter la robustesse des systèmes de reconnaissance faciale.

En termes de perspectives, plusieurs pistes d'amélioration se dessinent :

1. **Enrichissement des données** : Incorporer un ensemble de données plus varié et équilibré, comprenant une gamme plus large de caractéristiques physiques et de contextes, pour renforcer la capacité du modèle à généraliser à partir d'exemples moins représentés.
2. **Expérimentation avec des architectures de modèles plus avancées** : Tester d'autres configurations de réseaux neuronaux convolutionnels ou explorer des modèles récents de l'état de l'art pour améliorer la précision et la fiabilité de la reconnaissance.
3. **Optimisation des hyperparamètres** : Approfondir le tuning des hyperparamètres du modèle pour trouver le meilleur équilibre entre capacité de généralisation et performance spécifique sur le jeu de données cible.
4. **Développement d'interfaces utilisateurs plus sophistiquées** : Améliorer l'application web pour permettre un traitement et une analyse des images en temps réel, offrant ainsi une plateforme plus interactive et conviviale pour les utilisateurs finaux.

En conclusion, ce projet nous a permis de comprendre l'importance de l'approche méthodique et de la réflexion critique dans le développement de solutions de reconnaissance faciale. Il ouvre également des pistes passionnantes pour des recherches futures, avec l'objectif d'atteindre des systèmes plus précis, équitables et accessibles.