



Master 2 IMAGINE

**DIAB Ingo
LECOURT Florian**

**Compte Rendu 4 :
Édition du genre d'un portrait**

Année Universitaire 2023-2024

Table des matières

I	Partie sans Deep Learning	2
II	Partie avec Deep Learning	6
1	Observations	7
2	Problèmes avec CycleGan	7

I Partie sans Deep Learning

Concernant la partie sans Deep Learning, nous avons utilisé les handles fournis par dlib afin de mieux détourer les features que nous souhaitons extraire. Pour cela, nous utilisons une forme convexe englobant les handles de la feature voulue. Nous pouvons agrandir la forme convexe en "poussant" chaque sommet de l'enveloppe convexe depuis le centroïde de l'enveloppe.

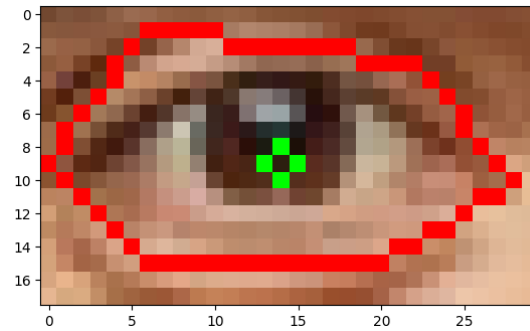


FIGURE 1 – Enveloppe convexe (rouge) et son centroïde (vert)

A partir de cette forme convexe, nous pouvons calculer un masque. Tous les pixels à l'intérieur de cette forme convexe sont en blanc sur le masque, les autres sont en noir.

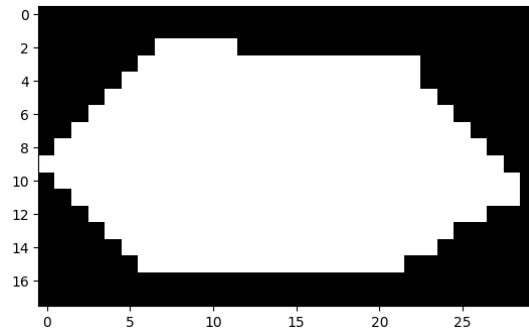


FIGURE 2 – Mask de l'oeil

En utilisant le masque ainsi que l'image extraite, nous pouvons effectuer un meilleur collage des features.



FIGURE 3 – Original

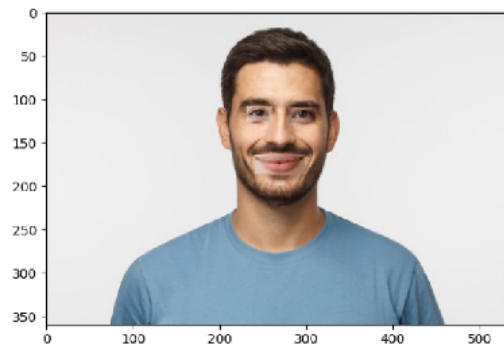


FIGURE 4 – Sans mask

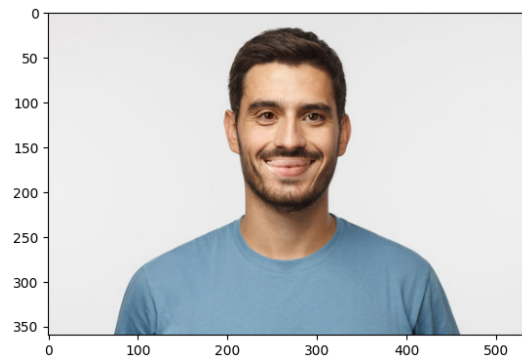


FIGURE 5 – Avec mask

Nous avons ensuite effectuer une triangulation de Delaunay, chaque tri-
angle des deux images sont mis en relation (ils relient les mêmes handles).



FIGURE 6 – Même triangles sur les deux visages

Cette triangulation nous permettrait "d'annuler" la rotation de l'individu sur le portrait mais nous permet aussi de cibler un "groupe de triangles" à partir duquel on pourrait extraire la barbe de l'individu.

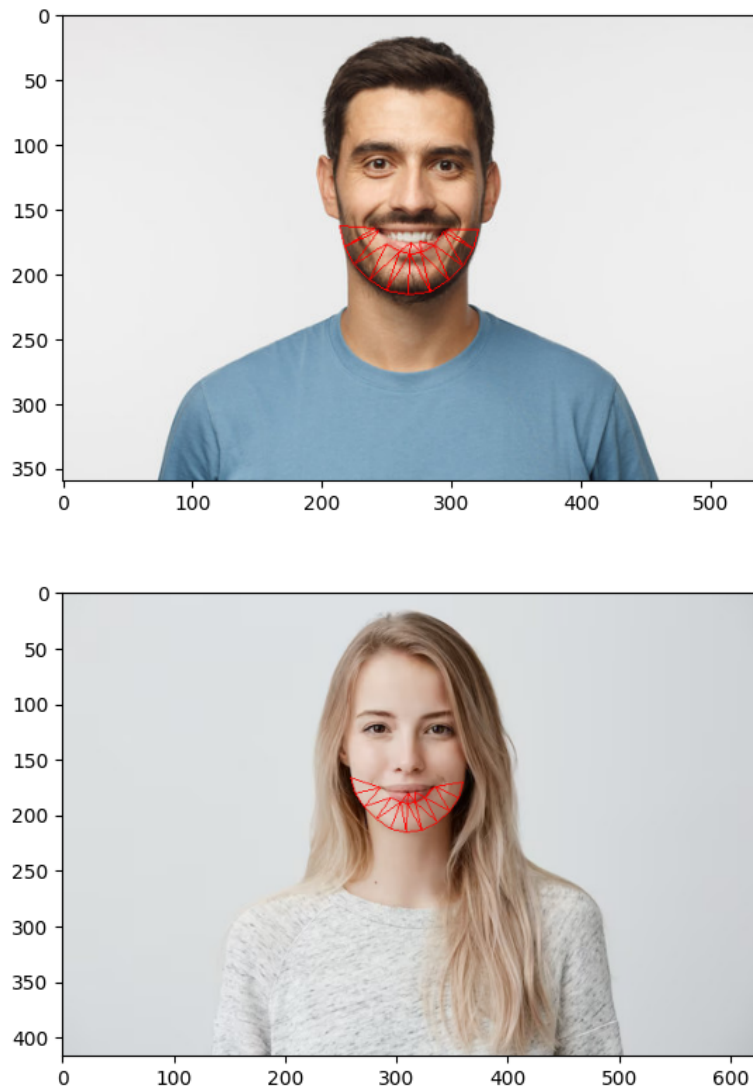


FIGURE 7 – Triangles représentant la barbe sur les deux visages

Le but maintenant serait de transformer ces triangles pour qu'ils aient la même dimension sur les deux images, les coller d'une image à l'autre et faire correspondre la couleur au reste de la peau.

II Partie avec Deep Learning

Comme évoqué lors du compte-rendu précédent, nous avons changé de générateur d'unet à resnet pour tenter d'obtenir de meilleurs résultats. Nous avons à nouveau entraîné notre réseau, avec les paramètres suivants :

- Taille de batch : 1000 (1000 photos de célébrités hommes, 1000 photos de célébrités femmes)
- Nombre d'epochs : 100
- Optimizer : Adam



FIGURE 8 – Résultats après 100 epochs

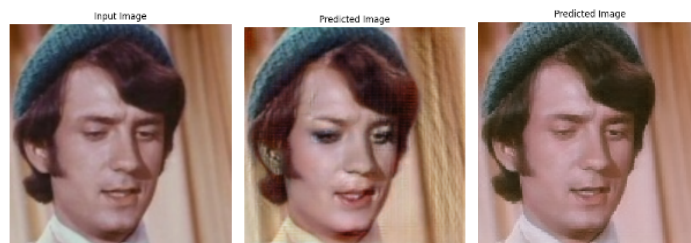


FIGURE 9 – Comparaison entre l'image originale, les résultats obtenues en utilisant le générateur Unet et utilisant le générateur ResNet



FIGURE 10 – Comparaison entre l'image originale, les résultats obtenues en utilisant le générateur Unet et utilisant le générateur ResNet

1 Observations

Contrairement aux images prédites avec l'utilisation du générateur Unet, les images prédites avec le générateur ResNet présentent moins de bruits dans le fond de l'image et des visages de qualité égale à l'image d'entrée. Cependant, même avec un nombre d'époques doublés (100 contre 50 epochs lors de l'utilisation du générateur Unet), les visages obtenus subissent peu de transformations comparés aux résultats obtenus précédemment.

2 Problèmes avec CycleGan

Après avoir échangés avec nos enseignants et avoir effectués de nouvelles recherches, les problèmes cités dans ce compte-rendu et dans le compte-rendu précédent proviennent de l'architecture CycleGan. CycleGAN est conçu pour la traduction d'images entre domaines sans paires d'entraînement. Il utilise deux générateurs et deux discriminateurs. Chaque générateur traduit une image d'un domaine à un autre (de femme à homme et vice-versa dans notre cas). Cependant, CycleGAN traite les images en se concentrant sur des parties spécifiques ou des caractéristiques locales, se concentrant sur la modification des caractéristiques de femmes pour les transformer en hommes, sans nécessairement comprendre ou modifier le contexte global de l'image.