

CR 1 Projet Face swap

CONDE SALAZAR Arthur, BONETTI Timothée

1 État de l'art

Lors de nos recherches nous avons identifié plusieurs phases d'évolution des techniques d'échanges de visage.

1.1 *Morphable 3D Faces et Poisson Blending*

Les premières techniques consistaient à modéliser les visages texturés en 3D à partir d'une ou plusieurs photos.

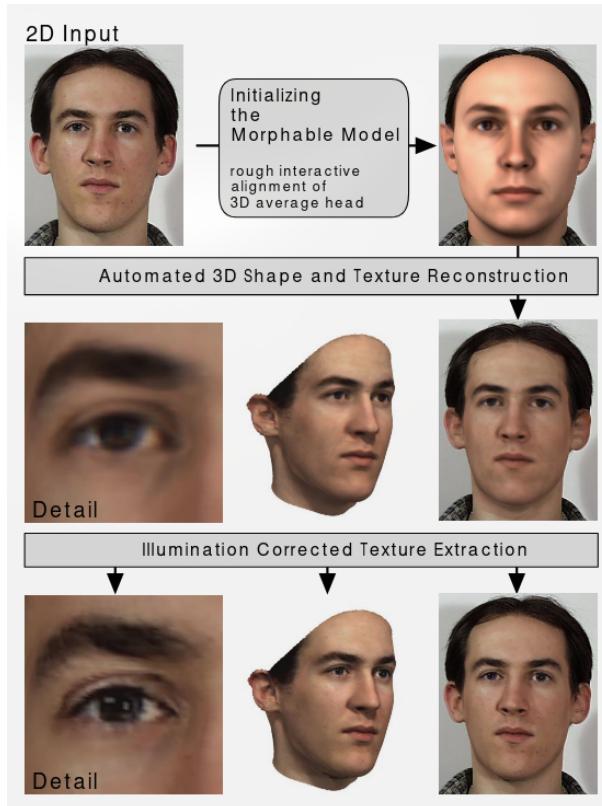


FIGURE 1 – Extraction de visage 3D texturé, extrait de *Blanz, V., & Vetter, T. (1999). A Morphable Model for the Synthesis of 3D Faces*

Une fois le visage extrait, on peut le plaquer sur une autre image comme démontré par les chercheurs *Blanz, V., Scherbaum, K., Vetter, T., & Seidel, H.-P. (2004). Exchanging Faces in Images.*

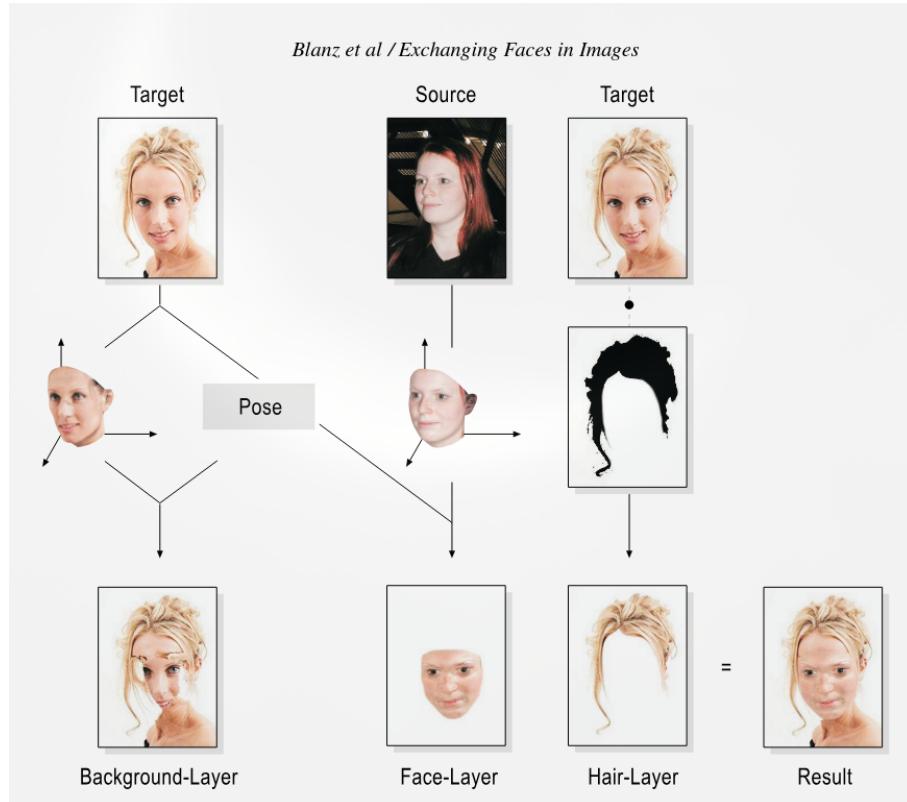


FIGURE 2 – Extrait de *Exchanging Faces in Images*.

Le mélange de Poisson peut aussi être utilisé, il permet de mélanger deux images en éliminant les bords et transitions brusques mais aussi de réajuster les couleurs. La technique a notamment été présentée dans le papier *Pérez, P., Gangnet, M., & Blake, A. (2003). Poisson Image Editing.*

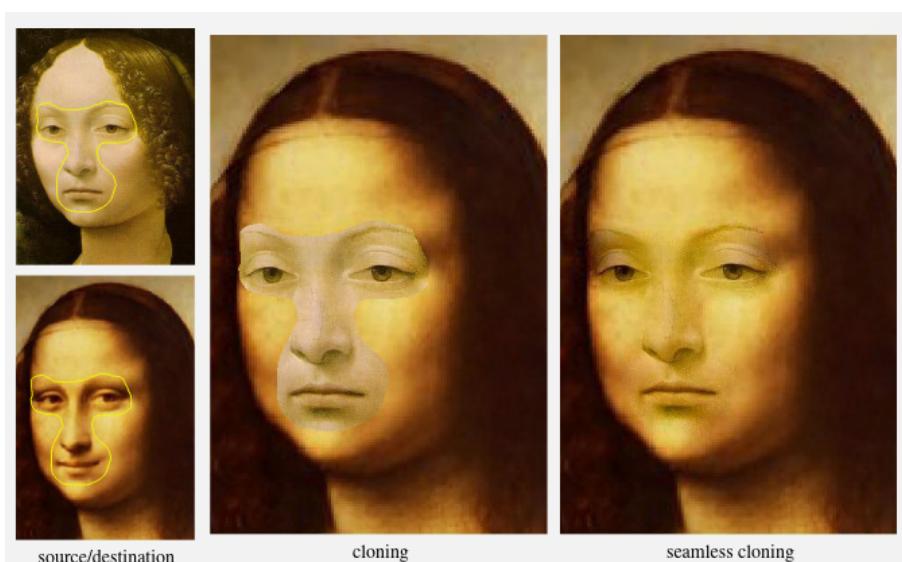


FIGURE 3 – Extrait de *Poisson Image Editing*.

1.2 CNN, Autoencodeur et GAN

Une des premières utilisations de CNN pour du Face Swap est faite par *Korshunova, I., Shi, W., Dambre, J., & Theis, L. (2017)* dans *Fast Face-Swap Using Convolutional Neural Networks*. Ils ont notamment entraîné un modèle nommé *CageNet* pour échanger le visage de célébrité (du dataset CelebA) avec celui de Nicolas Cage.



Figure 7: Left: original image, middle and right: CageNet trained on 256×256 images with style weights $\alpha = 80$ and $\alpha = 120$ respectively. Note how facial expression is altered in the latter case.

FIGURE 4 – Extrait de *Fast Face-Swap Using Convolutional Neural Networks*.

Les premiers DeepFakes réalisés à l'aide d'autoencodeurs sont réalisés par le projet communautaire Github [faceswap](#) en 2017 et décrits dans l'article [Understanding the Technology Behind DeepFakes](#).

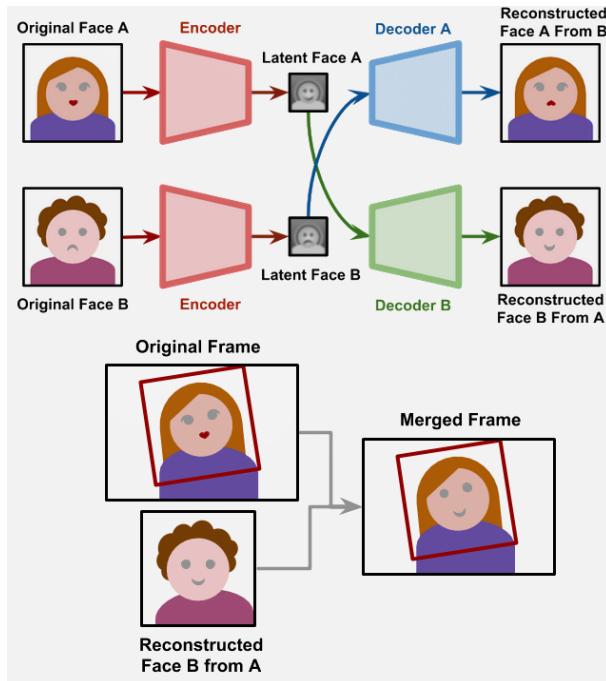


FIGURE 5 – Extrait de *Understanding the Technology Behind DeepFakes*



FIGURE 6 – FaceSwap avec le modèle phazeA avec une architecture d'autoencodeur

Une approche du face swap avec des GAN est d'abord décrite dans Nirkin, Y., Keller, Y., & Hassner, T. (2019). *FSGAN : Subject Agnostic Face Swapping and Reenactment*. Elle utilise un GAN pour séparer l'identité et l'expression, ce qui permet un échange de visage et une reproduction des expressions sur n'importe quel sujet.

La technique par GAN sera plus tard améliorée par Li, L., Bao, J., Yang, H., Chen, D., & Wen, F. (2019) dans *FaceShifter : Towards High Fidelity And Occlusion Aware Face Swapping*.



FIGURE 7 – Extrait de *FSGAN : Subject Agnostic Face Swapping and Reenactment*

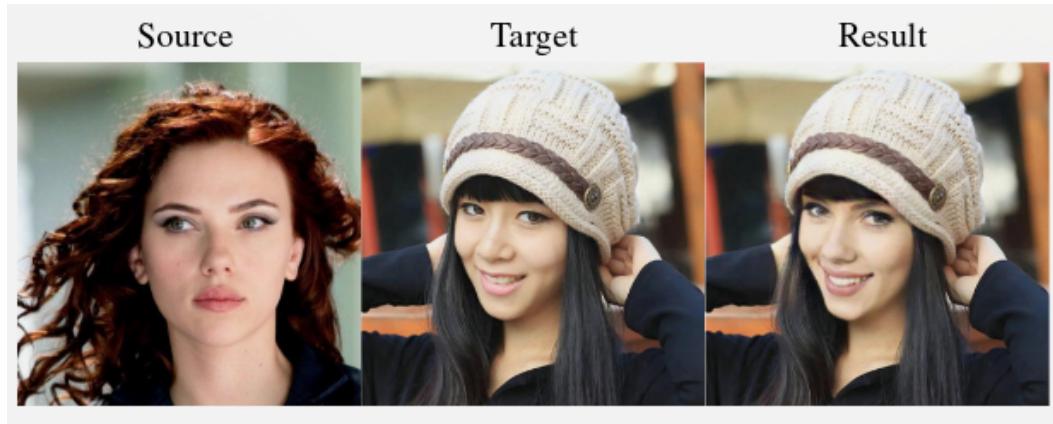


FIGURE 8 – Extrait de *FaceShifter : Towards High Fidelity And Occlusion Aware Face Swapping*

1.3 Modèle de diffusion et Transformers

Très récemment, de nouvelles techniques se développent, suivant le développement des nouvelles techniques d'IA générative.

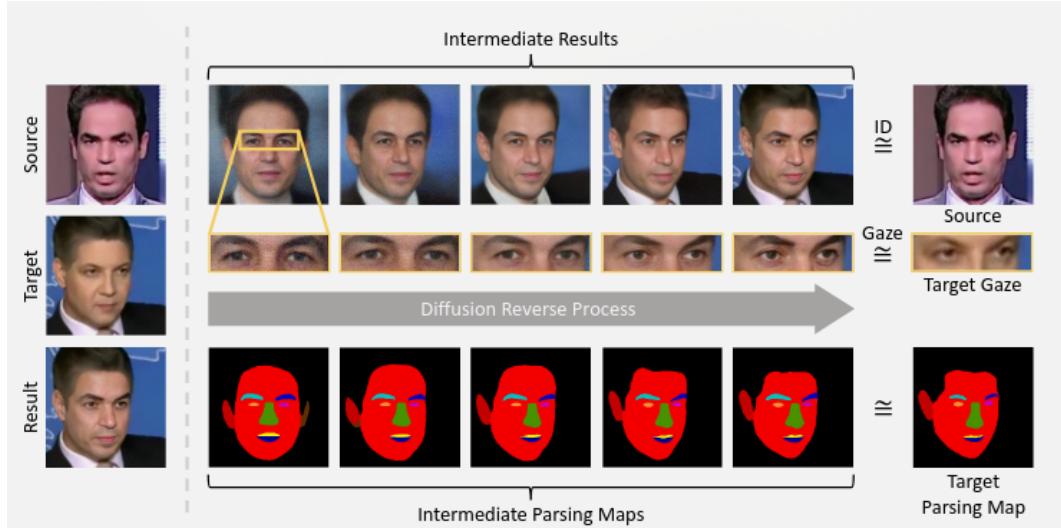


FIGURE 9 – Extrait de Kim, K., Kim, Y., Cho, S., Seo, J., Nam, J., Lee, K., Kim, S., & Lee, K.-H. (2022). *DiffFace : Diffusion-based Face Swapping with Facial Guidance*.

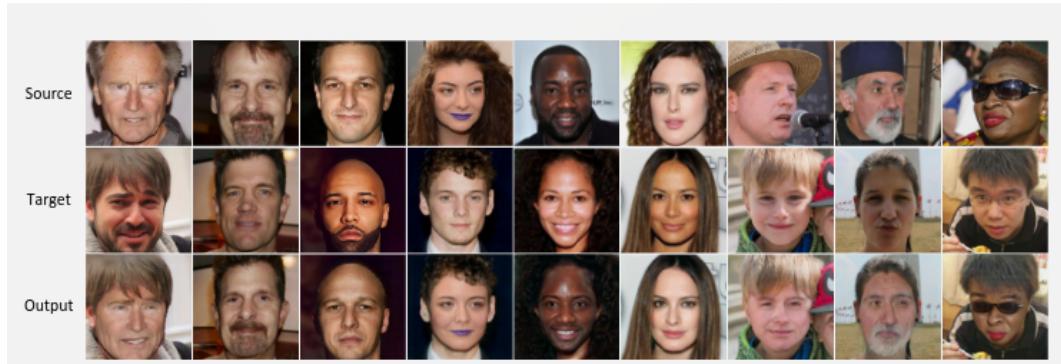


FIGURE 10 – Extrait de Cui, K., Wu, R., Zhan, F., & Lu, S. (2023). *Face Transformer : Towards High Fidelity and Accurate Face Swapping*.

2 Notre feuille de route

Comme demandé par M. Puech, nous allons commencé par développer un classifieur permettant de reconnaître nos camarades de classes (si ils acceptent d'être pris en photo).

Nous allons ensuite entraîné un décodeur qui nous permettra de reconstruire le mieux possible les images à partir de leur représentation dans l'espace latent.

Enfin pour faire du faceswap, nous pensons utiliser l'approche décrite dans l'article FSGAN, les chercheurs fournissent aussi un modèle sur [Github](#).