Thomas FERMELI-FURIC EPITA ING3

Deep Learning
-
CycleGAN Project Report

Link to the GitHub Repository

https://github.com/thomasfermelifuric/cyclegan

Introduction and Architecture

The objective of this project is to utilize CycleGAN to translate facial features between two distinct age groups: young (15–40 years) and old (60–90 years). This application demonstrates the potential of CycleGAN for unpaired image-to-image translation tasks, such as aging and rejuvenation of facial features.

CycleGAN (Cycle-Consistent Generative Adversarial Network) was introduced by Zhu et al. in 2017 as a method for unpaired image-to-image translation. It has been applied to tasks such as style transfer, domain adaptation, and object transfiguration. CycleGAN was chosen as the architecture due to its ability to learn mappings between two unpaired domains X and Y by enforcing a cycle consistency loss. This loss ensures that a sample $x \in X$ translated to domain Y and then back to domain X will reconstruct the original x. The model comprises two generators ($G: X \rightarrow Y$ and $F: Y \rightarrow X$) and two discriminators (DX and DY), trained adversarially to enhance the quality of the generated images

Dataset Details

The dataset used is a filtered version of the UTKFace dataset, sourced from Kaggle (https://www.kaggle.com/datasets/jangedoo/utkface-new). The dataset contains facial images labeled by age, gender, and ethnicity.

To improve the learning process, the dataset was filtered to include only images of white individuals, aged 15–40 for the "young" category and 60–90 for the "old" category. This simplification focuses on reducing variability across ethnicity, enabling the model to concentrate on age-related transformations. The filtering is not motivated by bias but rather to enhance the model's focus on age-related transformations in facial features. Including a broader ethnic range could introduce additional complexities in learning, potentially affecting translation quality due to unpaired data.

Domain X ("Young Faces"): Contains filtered images of individuals aged 15–40.

Domain Y ("Old Faces"): Contains filtered images of individuals aged 60–90.

Since my domain Y contains only 920 elements, data augmentation techniques, such as horizontal flipping, could have been applied to increase variability and robustness during training. However, the goal here was not to push the performance of the model, but rather to get the idea that the translation of an image from domain X to Y is possible when trained on unpaired-image, even with a simple model architecture and a thin dataset. For this reason, we made the most simplified experiments we could.

Experiments and Results

The CycleGAN architecture was implemented using PyTorch. The generators employed ResNet blocks for efficient feature learning, while PatchGAN discriminators were used to classify real versus fake image patches.

The Adam optimizer was used with a learning rate of 1e-3. Training was conducted over 20 epochs, with a batch size of 1 due to the limited dataset size. The best model we got was saved after the 19th epoch. We observed that acceptable results were quickly obtained, followed by a complete degeneration of the model performance. After 20-30 epochs, the training loss started increasing again and the generated images lost all of their delicacy, with people looking like an eight-year-old kid’s Santa Claus drawing.

At the end, the generated images successfully displayed transformations, with the model capturing key aging features such as wrinkles, skin sagging, and changes in facial structure. However, challenges were noted in generating consistent results for certain edge cases (e.g., individuals with glasses or distinct hairstyles). Here is an example of a well-working image translation :

Image Originale

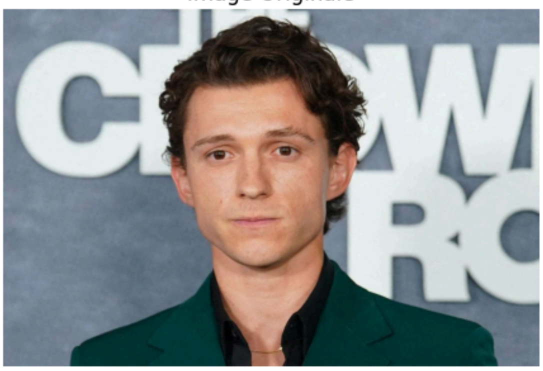

Image Générée

Another interesting point about CycleGAN is that we trained two generators. This way, we can of course also translate images from old to young faces using the same process. In this project, the old-to-young translation was not that good. As we can see, our generators still have trouble with hair colour and density, which makes the old-to-young translation a bit disappointing. Here is an example of a translation where the model still succeeded in smoothing the skin and darkening the hair :

Conclusion

This project successfully demonstrates the use of CycleGAN for unpaired image-to-image translation in facial aging and rejuvenation. By filtering the dataset and focusing on clear age group distinctions, the model achieved notable results while using a small model trained on an even smaller dataset.

This experiment highlights the potential of CycleGAN in applications requiring domain-specific transformations while emphasizing the importance of thoughtful dataset preparation to maximize model performance.