AgingGAN: Age Progression with CycleGAN

Jie Chen (jiechen8), Junwen Bu (junwenbu), and Yu Zhao (zhaoyu92)

CS230 Deep Learning, Stanford University

Introduction

Age progression, the process of aesthetically rendering a facial image with simulated effect of growing old, has attracted much attention from the Deep Learning and Computer Vision community. It is a challenging task because the patterns of aging that we want to capture could be easily affected by the various conditions of the input image. Further, the scarcity of paired data – two images of the same person taken at different time (20+ years apart) -- prevented existing solutions to achieve good performance.

In this project, we proposed a simple, yet intuitive deep learning model based on **CycleGAN** [1] that can generate predictive images of people's look after certain years based on their current images, without the need of paired dataset.

Dataset & Features

IMDB-WIKI [2]

- Group A (age 20~30)
  - 5,004 images (3,165 male, 1,839 female).
  - filtered with *face_score* > 3.
- Group B (age 50+)
  - 2,779 images (2,209 male, 570 female)
  - filtered with *face_score* > 1.

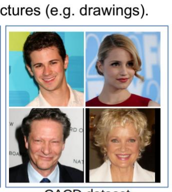
Cross-Age Celebrity (CACD) [3]

- Group A (age 20~30)
  - 2,200 images randomly taken from pool of 39,069.
- Group B (age 50+)
  - 2,200 images randomly taken from pool of 33,872.

Modifications

- Resized to 256 x 256.
- Removed Grayscale images
- Removed images that are not pictures (e.g. drawings).

IMDB-WIKI dataset



CACD dataset

Method and Model

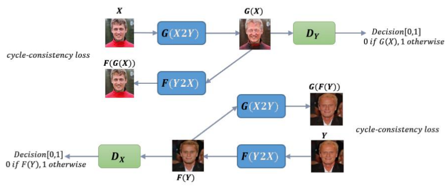


Figure 1 - Network Architecture: Two mapping functions G and F associated adversarial discriminators D_Y and D_X. Introduce *cycle-consistency* loss to regularize the model

$$L_{GAN}(G, D_Y, X, Y) = \mathbb{E}_{y\sim p_{data}(y)}[log D_Y(y)] + \mathbb{E}_{x\sim p_{data}(x)}[\log(1 - D_Y(G(x))]$$

$$L_{GAN}(F, D_X, Y, X) = \mathbb{E}_{x\sim p_{data}(x)}[log D_X(x)] + \mathbb{E}_{y\sim p_{data}(y)}[\log(1 - D_X(F(y))]$$

$$L_{cyc}(G, F) = \mathbb{E}_{x\sim p_{data}(x)}[\|F(G(x)) - x\|_1] + \mathbb{E}_{y\sim p_{data}(y)}[\|G(F(y)) - y\|_1]$$

Objective: $G^*, F^* = \arg \min_{G,F} \max_{D_X, D_Y} [L_{GAN}(G, D_Y, X, Y) + L_{GAN}(F, D_X, Y, X) + \lambda L_{cyc}(G, F)]$

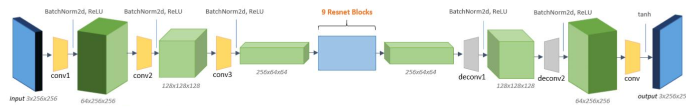


Figure 2 – Generator: Three components 1) Encoder 2) Transformer 3) Decoder

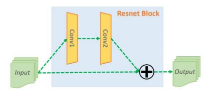


Figure 3 – Resnet Block

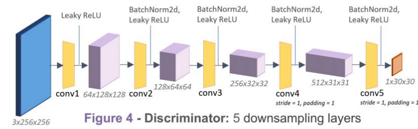


Figure 4 - Discriminator: 5 downsampling layers

In practice, using ***least-squares* loss**, e.g. $L_{GAN}(G, D, X, Y)$:

Train G to minimize $\mathbb{E}_{x\sim p_{data}(x)}[(D(G(x)) - 1)^2]$

Train D to minimize $\mathbb{E}_{y\sim p_{data}(y)}[(D(y) - 1)^2] + \mathbb{E}_{x\sim p_{data}(x)}[D(G(x))^2]$

Results

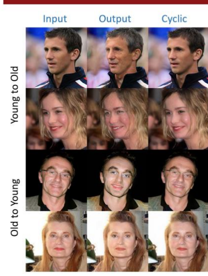


Figure 5 – CycleGAN In Action: Images are processed through "young to old to young" half or "old to young to old" half. We can see that the cyclic image is very similar to the original image.

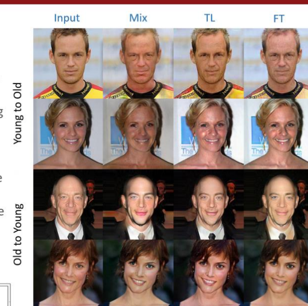


Figure 6 – Speed-up learning: "Mix" is the model trained with male and female images (as baseline); "FT" is fine tuning on top of *horse2zebra* model; "TL" is transfer learning with *horse2zebra* model.

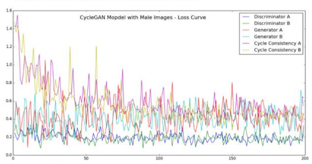


Figure 7 – Loss Curve: sample loss curve for training dataset of male images only

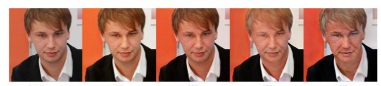


Figure 8 – Growing Old: cross epoch effects

Acknowledgements: CS230 teaching staff

Conclusions

1. CycleGAN can generate quality age progression images.
2. The aging effects will increase as # of epoch increases, but such effect become less and less apparent after 200 epochs.
3. Transfer learning and fine tuning using other trained model (*horse2zebra* model in our case) can be applied to accelerate training but will slightly compromise the quality of the output.
4. The choice of dataset can severely affect the performance of the model (CACD dataset has horrible results).

Future Work

- Investigate the correlation between the Cycle-Consistency cost and image quality.
- Increase training set size to 20~50K.
- Explore models support facial geometric changes.

References

1. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. in ICCV, 2017.
2. IMDB-WIKI – 500k+ face images with age and gender labels, https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/.
3. Cross-Age Reference Coding for Age-Invariant Face Recognition and Retrieval, http://bcsiriuschen.github.io/CARC/.