Deep Face Swap with GAN

Chi Wang, Jinil Jang
{chiwang, jinil}@stanford.edu

Introduction

Human image synthetic technology plays an important role in movie editing. In nowadays, AI based approach, such as “DeepFake”, is getting attention, but it has constrained with factors such as light condition or skin tones. We overcome the limitation with GAN.

Dataset

- Collected 500 celebrity – “George Clooney” – images from Google
- 200 our own images (200/Chi or 200/Jinil) for model training
- Randomly split 500 celebrity’s pictures to 400 / 50 / 50 as train, validation and test sets.
- In pre-train, detected faces are extracted from images, warped, and scaled to 64 x 64 x 3.

Methods

- MSE loss for the Discriminator
- SSIM loss for the Autoencoder

\[
SSIM(x, y) = \frac{(2μ_xμ_y + c_1)(2c_2σ_{xy} + c_2)}{(μ_x^2 + μ_y^2 + c_1)(σ_x^2 + σ_y^2 + c_2)}
\]

- ADAM with
  - learning rate = 5e^{-3}, β_1 = 0.5, β_2 = 0.999

Models

- Autoencoder
- GAN Architecture

Approach

Training

Performed on a GPU-enabled EC2 instance per each algorithm (Baseline, More Intermediate layers, and GAN) in parallel over 72 hours.

Results

- Training Loss
- Quality Comparison
- Discussion

Discussion

- To get reasonable output, it’s required at least 500 dataset, and over 50 hour training process on GPU-enabled machine. Long training process made us hard to tune hyperparameters.
- If there is multiple faces in an image, all faces were swapped with source face. Face recognition was needed to swap only the target face.
- After 40K epoch, learning was slow and gave blur output. Early stopping and applying GAN on output result can reduce train process.

Future Directions

- Bad output example:
  - Face edge is not smooth with bear.
  - Black/white image gets blur result.
- Future Work:
  - Data Augmentation: Generate Black/White images to improve model performance on dark environment.
  - GAN on the swapped image to make edge smooth.

References