Makeup Removal System with Deep Learning

Mingchen Li, Yiyang Li, Yifan He

{limc, yiyang7, heyifan}@stanford.edu

CS 230 Deep Learning, Stanford University

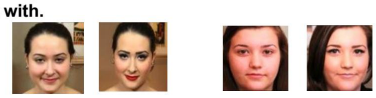
Objective

This project we are doing is a makeup removal deep learning system which can erase people's makeup. This novel system carries many practical applications, e.g. security applications on makeup face verification, social network apps for consumers since facial cosmetics can sometimes be deceivable.

Dataset

The dataset comes from the Internet and other papers. We sent emails to the authors of the papers to ask for it. We gathered five dataset (FAM, MIFS, MIW, VMU, YMU) with around 2000 images. Most people have one image without makeup and one with, while some people have one image without makeup and three images with (lips, eyes, full).

FAM and MIFS are come with plain photos that we can use directly. Other three datasets are boxed as Matlab matrices, so we have to decode them into photos to visualize the data. Some of the sample images are showing below, with the left one without makeup, the right one with.

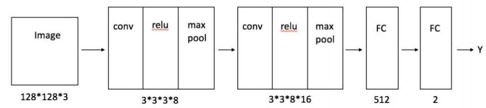


Preprocessing

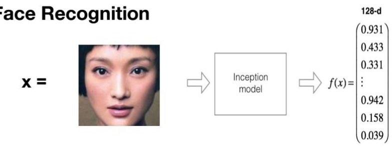
- Resize the images to 96*96*3 using Matlab
- Manually go through all the data (around 2000 images), and throw away the bad ones.
- Flip the image left to right as data augmentation

Models

1. CNN

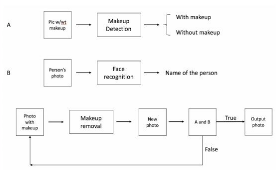


2. Face Recognition

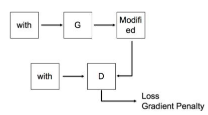


3. GAN

①



②



The first GAN showing here uses the normal GAN's loss function, which trains slowly and results are not satisfying. We adopted the WGAN-GP architecture, which is the second GAN on the left.

WGAN-GP is different with traditional GAN for it adds a gradient penalty on the discriminator, so that the training process is easier than that of the normal GAN.

We use a CNN plus U-NET in the generator to generate makeup-removed images, and a normal CNN in the discriminator.

Results

Training GAN is a very time consuming and somewhat frustrating job. We've tried lots of different architecture in the generator and discriminator, and still cannot get satisfying results. The image one the left showing blow is from a WGAN-GP trained on a pure noisy input and an image without makeup. The results are closed to the without make image, but still need refinements.

The image one the right showing below is from the same WGAN-GP, with the architecture of the generator been modified to take an with makeup image as an input and generate a without makeup image.

Future work

- Continue tuning the hyper parameters
- Try different architecture for G and D
- Refine the loss function. Ex: add feature encoding difference, face symmetry loss…

References

[1] Yi L; Lingxiao S; Xiang W; Ran H; Tieniu T. 2017. Anti-Makeup: Learning A Bi-Level Adversarial Network for Makeup-Invariant Face Verification. In *Computer Vision and Pattern Recognition.*

[2] Jun-Yan Z; Taesung P; Phillip I; Alexei E. 2018. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In *Computer Vision and Pattern Recognition.*

[3] Jun-Yan Z;, Richard Z; Deepak P; Trevor D; Alexei E; Oliver W; Eli S. 2017. Toward Multimodal Image-to-Image Translation. In *Computer Vision and Pattern Recognition.*

[4] Ishaan G; Faruk A; Martin A; Vincent D; Aaron C. 2017. Improved Training of Wasserstein GANs, In *Learning*.