FacePaint: Localized Style Transfer of Facial Expressions

Sasha Harrison | Frits van Paasschen
aharris6@stanford.edu | fritsvp@stanford.edu

CS 229
Fall 2019

Motivation & Definition

In recent years, Generative Adversarial Networks (GANs) have become the state of the art in artificial image generation. We applied **Neural Style Transfer**, Image Segmentation, and GANs to a new application, altering facial expressions in photos. This generative task may have useful applications in special effects. or artistic tooling. As input our model takes an image of a person smiling, and as output produces an image with a neutral facial expression.

Problem Statement: Is it possible to use neural networks to change a person's facial expression?

Custom Dataset

- For SA-GAN, used only **JAFFE** dataset
- JAFFE consists of 213 images of 10 distinct Japanese women
- Each makes 6 different facial expressions (anger, disgust, fear, happiness, sadness,and surprise)
- For Cycle-GAN, simplified to **2-class problem**, padded dataset with faces from FEI Dataset
- Cycle-GAN training set had 446 images, each with 1 channel

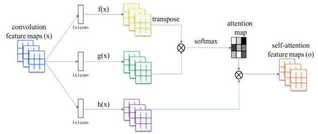
Happy

Sad

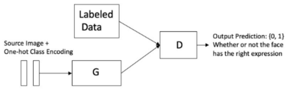
Neutral

Approaches

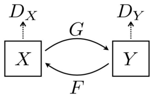
- Experimented with **various architectures** and models
- **Style Transfer + Segmentation**: transplant face + smooth
- **Self-Attention C-GAN**: CGAN + module to model long-range spatial feature dependencies
- **Cycle GAN:** GAN transforms images between two classes, enforcing cycle consistency



Self-Attention Module



SA-Gan Architecture



Cycle-GAN

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

$$\begin{aligned} \mathcal{L}(G, F, D_X, D_Y) = & \mathcal{L}_{\text{GAN}}(G, D_Y, X, Y) \\ & + \mathcal{L}_{\text{GAN}}(F, D_X, Y, X) \\ & + \lambda \mathcal{L}_{\text{cyc}}(G, F), \end{aligned}$$

Cycle Consistency Loss

Conclusions and Future Work

Both GANs were able to learn **non-trivial expression changes** on input images. Cycle-GAN cycle-consistent loss **greatly increased performance**, and produced more realistic results. This is likely because reducing the problem to two-class simplifies the prediction. For example, the model seemed to localize changes to the mouth area.

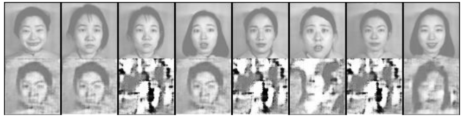
Future Work:

- Expand Cycle-Gan implementation to multiclass problem
- Experiment with **Style-GAN** architecture for better resolution

Experiments and Results

- **Self-Attention (SA-CGAN): Completely learned image transformation**, better quality than CGAN, but needs improvement
- **Mode collapse** frequent with SA-CGAN
- **Cycle-GAN:** Trained for 350 iterations, lr = 0.0005
- Learned high quality transformation between Happy and Neutral expressions.
- Micro and macro-level facial feature adjustments show clear model ability to interpret and manipulate expression

SA-CGAN Results



Cycle-GAN Results

Normalized Test Set L1 Distance
Cycle-GAN: 134.59

Orig. Img (H)

Fake Img (N)

Orig Img (N)

Fake Img (H)

References

- Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." Proceedings of the IEEE international conference on computer vision. 2017.
- Zhang, Han, et al. "Self-attention generative adversarial networks." arXiv preprint arXiv:1805.08318 (2018).
- Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.