Clothing Texture Synthesis using CycleGAN

Miguel Ferrer Avila, Ji Hoon Andy Kim
{mferrera, jkim4223}

Abstract and Background

In recent years, we have seen an explosion of deep learning in diverse fields, driven by the rise in computing power, the increase in availability of data, and the scalability of large systems. Within the fashion industry, attribute and category matching as well as recommendations systems are extensively used to market clothing[2]. However, one of the more unique applications of deep learning is to perform unpaired image-to-image translation using CycleGAN to generate images from domain A to domain B[1]. In this project, we combine these ideas and perform image-to-image translation on articles of clothing, in our case, translation between denim and leather clothing, and hope to generate novel fashion designs.

Dataset

- 646 RGB images of leather clothing and 711 RGB images of denim clothing from both the DeepFashion and ImageNet datasets[3].
- Images are approximately 300 x 300
- 20% held out for validation

Figure 1: Sample Images from the dataset

Data Processing and Augmentation

- Downsampled images to 286 x 286
- Random cropping on images to 256 x 256
- Random horizontal flipping on images
- Rotation of images
- Center image and normalize the pixel values

Methodology and Metrics

- CycleGAN architecture as done by Zhu et al.[1]

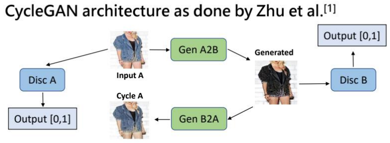

Figure 2: CycleGAN Model

- Adversarial Loss

$$\mathcal{L}_{GAN}(G, D_B, B, A) = \mathbb{E}[\log(D_B(b))] + \mathbb{E}[\log(1 - D_B(G(a)))] + \mathbb{E}[\log(D_B(G(a)))]$$

- Cyclic Loss

$$\mathcal{L}_{cyclic} = \mathbb{E}||x - F(G(x))||_1 + \mathbb{E}||y - G(F(y))||_1$$

- Objective

$$(G^*, F^*) = \arg\min_{G,F}\max_{D_A, D_B} \mathcal{L}_{GAN}(G, D_B, A, B) + \mathcal{L}_{GAN}(F, D_A, A, B) + \lambda\mathcal{L}_{cyclic}$$

Architecture

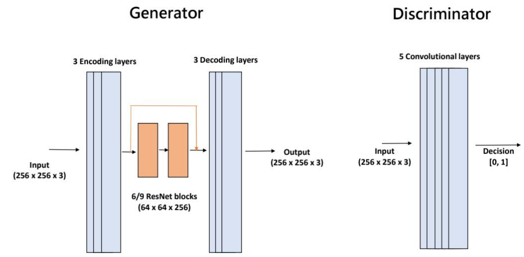

Figure 3: Left - Generator architecture. **Right** – Discriminator Architecture

Results

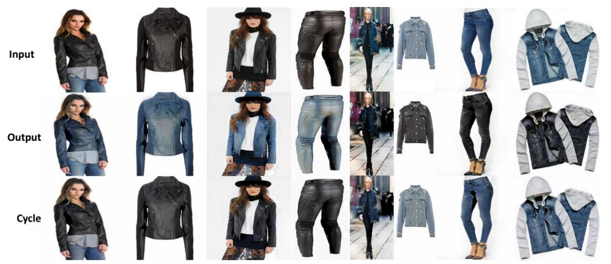

Figure 4: Image-to-Image Translation output on our validation set. Top row shows original image, second row show generated image and third rows shows the recocovered image. Model is trained for 200 epochs with Adam optimization using a batch size of 16.

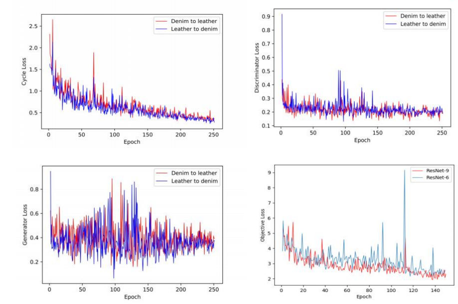

Figure 5: Top Left – Cyclic Loss, **Top Right**- Discriminator Loss. **Bottom Left –** Loss of a ResNet-9 generator. **Bottom Right –** Comparison in objective loss between Resnet-9 and ResNet-6 Architectures for the generator

Conclusion and Future Work

- CycleGAN works well in generating novel fashion designs
- ResNet-9 architecture works best and could use more training with more data
- Use inception scoring metric in addition to the current loss metric
- Experiment with DenseNet and UNet designs for generator
- Perform additional data curation (Upsample the number of images for a more balanced dataset, increase training size, etc...)

References

[1] J. Zhu, T. Park, P. Isola, A. Efros. 2017. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. *IEEE* (2017)

[2] A. Sun, Y. Tay, S. Zhang, L. Yao. 2018. Deep Learning based Recommender System: A Survey and New Perspectives. *Comput. Surveys 1, 1 (2018)*, 1-35.

[3] P. Luo, X. Wang, X. Tang, Z. Liu, S. Yan. 2016. DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations. *European Conference on Computer Vision (ECCV) (October 2016).*