**Literature Review**

**- Abulaiti Akeli**

**Introduction**

This paper aims to review some of the scholar papers which are related to my project-to make a filter that can turn any real-life images into anime style images. Since my project is heavily based on image style transfer, I was looking for the papers which related to this topic. I searched with broader concepts and topics first and then slowly narrowed down and went deep for specific topics based on what I found and learned. The following review will also be conducted in this manner.

**First Paper**

The first paper I am reviewing is called “Advanced Deep Learning Techniques for Image Style Transfer: A Survey”. [1] This paper itself is a research and summarization of various image style transfer techniques. It starts with the style transfer techniques without using neural networks, which are a little bit old fashion and less advanced, but relatively easy to use and can come in handy under certain circumstances, like texture synthesis, image filtering, or image analogy. Next, it moves on to the methods with using the neural networks, which include slow transfer based on online image optimization and the fast transfer based on online model optimization, which can be further divided into PerStyle-Per-Model, MultipleStyle-Per-Model and Arbitrary-Style-Per-Model. Finally, this paper also covered the applications of these methods.

I think this paper is a nice starting point for my project since it is very general and informative. It doesn’t go in very deep for each method but it covers a lot of them with simple definition and application. It is very helpful to someone like me who just started the project and wonder where to start from. It gives me a rough idea on the concepts and how can I approach my project, what should I further looking into for next.

**Second Paper**

The second paper is called “Image Style Transfer Using Convolutional Neural Networks”. [2] Unlike the first paper, this focuses more narrowly and deeply into how to apply CNN to image style transfer. It introduces “A Neural Algorithm of Artistic Style”, a new algorithm to perform image style transfer. To achieve the objection, they started with normalizing the basic VGG networks, which they called deep image representations. Then they performed gradient descent to visualize the image information that is encoded at different layers of the hierarchy, which they called content representations. They used a feature space designed to capture texture information to obtain a representation of the style of an input image. Finally, they could perform the style transfer which synthesize a new image that simultaneously matches the content representation of the style representation. Their results shows that the representations of style and content are well separable using CNN but there is always some trade-off between content and style matching.

I think this paper could be useful to me as it provides detail steps and methods to perform image style transfer with CNN. Furthermore, it provides lots of functions which honestly kind of hard to understand but still could be useful to take and modify into my own project.

**Third Paper**

The third paper is called “Deep Generative Adversarial Networks for Image-to-Image Translation: A Review”. [3] This is also a review type paper, but with a completely new topic from the first one. The paper introduced the Generative Adversarial Networks (GAN), and how it can be used to perform image to image translation. While CNN uses feature extraction and image classification for style transfer, Deep Generative Model uses data distribution to discover the underlying features from large amounts of data in an unsupervised setting. They are able to generate new images by learning the distribution and predicting the results in contexts. The related applications include image super-resolution text-to-image generation and image-to-image translation. “A GAN is composed of two competing neural networks inspired by the two-player minmax game: a generative network, called a generator and denoted G, and a discriminative network, called a discriminator and denoted D. The generator network tries to generate realistic samples to fool the discriminator, while the discriminator tries to distinguish real samples from fake samples.” [3] These two models work against each other to create the best undistinguishable result. The paper also introduces different types of GANs including fully connected GAN, conditional GAN, informative GAN and big GAN. After done with introduction to all the GANs and their structure it starts the concept of image-to-image translation and then transited into how we can use GAN for this operation. It further introduced three supervised translation techniques and three unsupervised translation techniques.

This is a very long and detailed review paper which covered a lot of different GANs and their structures, functionalities and applications. It is very hard to summarize its full content without writing a whole page here. However, I found it very informative and useful to learn about GAN in general. I found that GAN could be use for image style transfer and it might actually suit my project better than CNN. Only if I don’t consider the difficulty, data samples needed and computational power.

**Fourth Paper**

The fourth paper is called “Training Generative Adversarial Networks with Limited Data”. [4] This paper introduces a new technique to achieve a satisfying result with GAN with limited data. It starts with explaining why overfitting is a problem in GAN and how does it happen. After that, it introduces Stochastic discriminator augmentation, which enforce discriminator consistency for both real and generated images. Basically, it’s a set of augmentations to all images shown to the discriminator. It evaluates the discriminator only using augmented images, and do this also when training the generator. From there, there is a further developed which turns it into adaptive discriminator augmentation, which avoids manual tuning of the augmentation strength and instead control it dynamically based on the degree of overfitting. Finally, “the method was tested against a number of alternatives in FFHQ and LSUN CAT, first in a setting where a GAN is trained from scratch, then by applying transfer learning on a pre-trained GAN.” [4]

From the last paper, I learned that I have to use a lot of data and computational power if I want to use GAN for my project, which could be a big problem as I probably don’t have enough times to find and collects tons of images on the internet and trained them with my potato laptops (non Pro version Google Collab doesn’t do a better job either). With all that being said this paper could be very helpful to me to limit the data use for my project and save the times.

**Fifth Paper**

The last paper is called “CartoonGAN: Generative Adversarial Networks for Photo Cartoonization”. [5] In this paper, they proposed a solution to transforming photos of real-world scenes into cartoon style images. They used learning-based methods for the project, and generative adversarial network (GAN) frame-work for cartoon stylization. They took unpaired photos and cartoon images for training, with two different losses designed for their cartoonization. First one is a semantic content loss. It’s said in the paper to be “formulated as a sparse regularization in the high-level feature maps of the VGG network”.[5] It would then cope with the substantial style variation between images of real-life scenes and cartoons. Second, they used an edge-promoting adversarial loss for preserving clear edges. They further introduced an initialization phase, to improve the convergence of the network to the target manifold. Their results show that their method is able to generate high quality cartoon images from real world scenes.

I don’t think I need to further comment on how this paper related to my own project, because their project is pretty much similar to mine even with what they were training. This is probably the most useful and important paper that I can learn and use for my own project.

**References**

[1] Long Liu, Zhixuan Xi, RuiRui Ji, Weigang Ma. Advanced Deep Learning Techniques for Image Style Transfer: A Survey. Signal Processing: Image Communication 78 (2019) pg. 465-470.

[2] Leon A. Gatys, Alexander S. Ecker, Matthias Bethge. Image Style Transfer Using Convolutional Neural Networks. CVPR 2016, pg. 2414-2423.

[3] Aziz Alotaibi. Deep Generative Adversarial Networks for Image-to-Image Translation: A Review. Symmetry 2020, 12, 1705.

[4] Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, Timo Aila. Training Generative Adversarial Networks with Limited Data. NeurIPS-2020.

[5] Yang Chen, Yu-Kun Lai, Yong-Jin Liu. CartoonGAN: Generative Adversarial Networks for Photo Cartoonization. CVPR 2018, pg. 9465-9474.