NEURAL STYLE TRANSFER AND GAN APPLICATION

DOAN NGOC VINH - 20204935

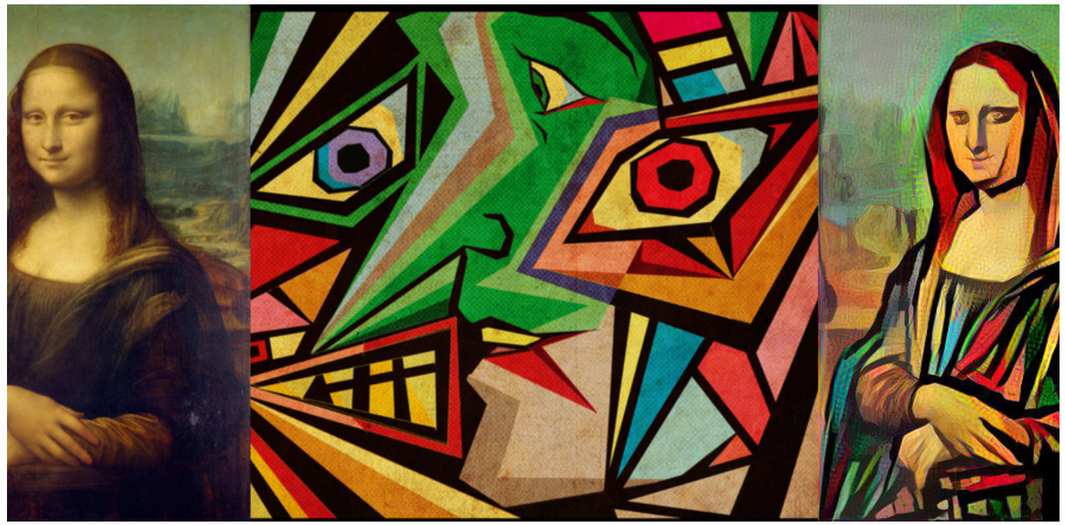
TABLE OF CONTENTS

INTRODUCTION

Computer vision is a branch of artificial intelligence (AI) that enables computers and systems to extract useful information from digital photos, videos, and other visual inputs and to execute actions or make suggestions based on that information.

Computer vision has became more and more popular, so do Neural style transfer and GAN. In the project, i will explain and give the illustration of using both.

NEURAL STYLE TRANSFER



Neural style transfer is a technique for transferring a certain look from one picture to another while preserving the original image’s information. The only modification is the way the image is styled to give it a more artistic feel.

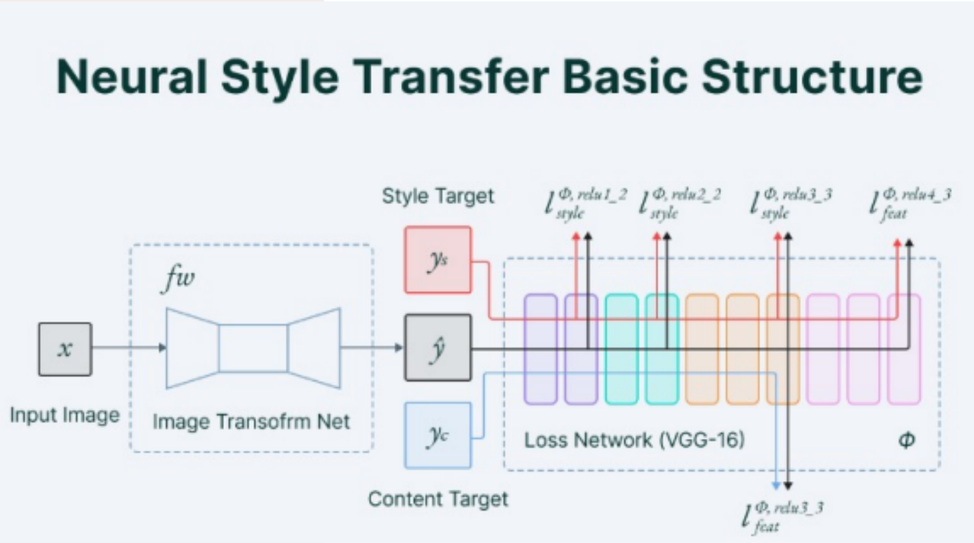
The layout or drawing is depicted in the content image, and the colors or painting used to represent the style are used. It is a computer vision application involving deep convolutional neural networks and image processing methods

NEURAL STYLE TRANSFER

A pre-trained feature extractor and a transfer network are required for training a style transfer model.

We use the correlation between several filter responses to reflect the style of a reference image: content loss and style loss then parameters are used to balance the total loss.

$$\mathcal{L}_{total}(\vec{p}, \vec{a}, \vec{x}) = \alpha \mathcal{L}_{content}(\vec{p}, \vec{x}) + \beta \mathcal{L}_{style}(\vec{a}, \vec{x})$$



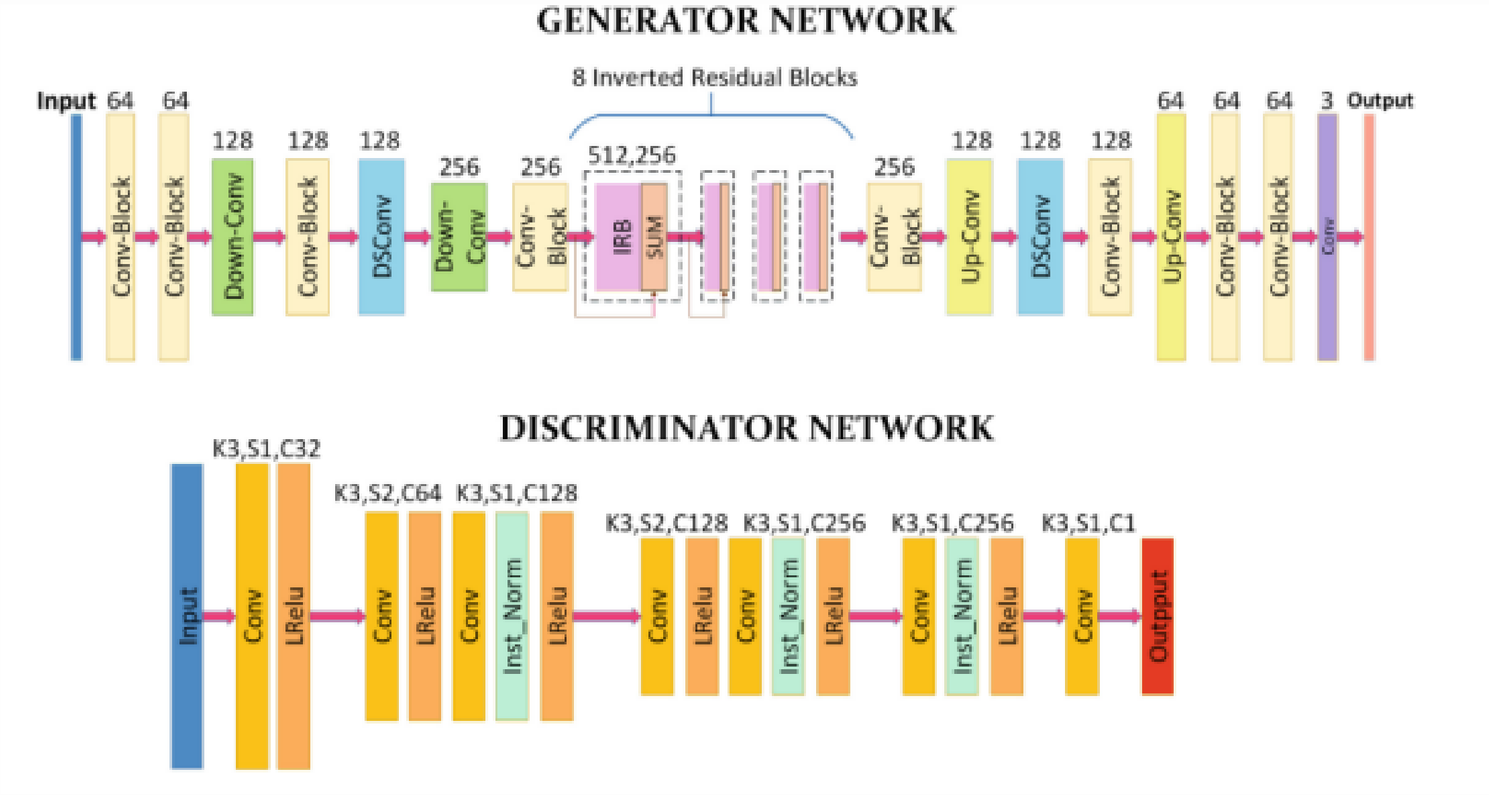
VGG16 IN NEURAL STYLE TRANSFER

Contains convolution layers and pooling layers.

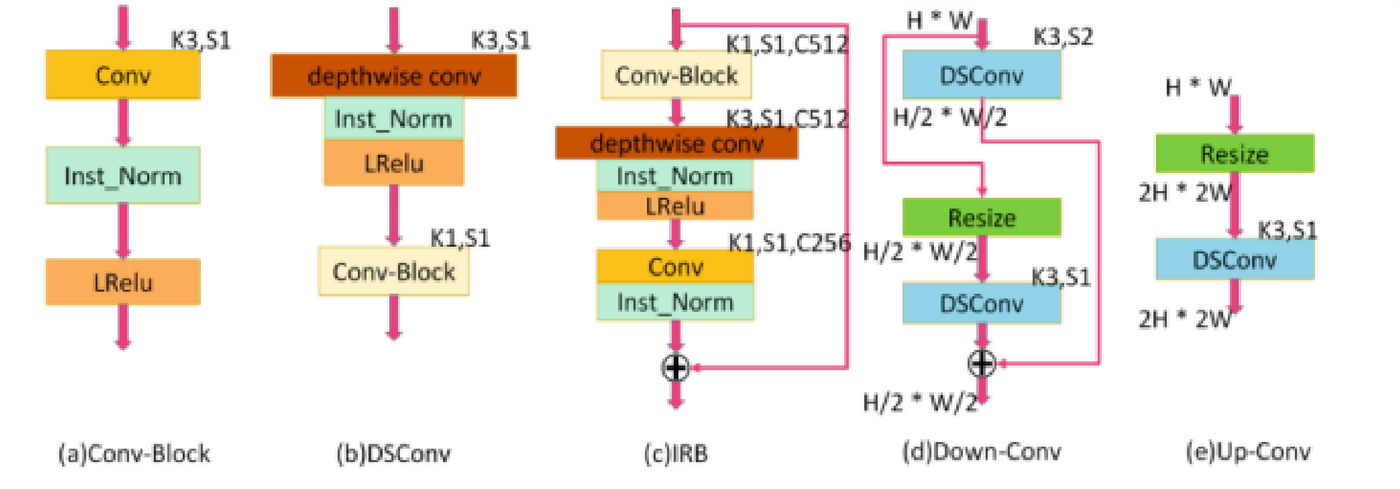
Conv 1-2, Conv 2-2. Conv 3-3. Conv 4-3 are extracted for style loss. Conv 4-3 is extracted for content loss.

GENERATIVE ADVERSARIAL NETWORK

In a generative adversarial network (GAN), two neural networks fight with one another to make predictions that are as accurate as possible. The majority of the time, GANs operate unsupervised and learn through cooperative zero-sum games



ARCHITECTURE



01 Conv-Block: This block's purpose is to obtain feature maps

02 DSConv: The block acts as a Conv-Block but less parameters, the Conv(K1-S1) is used to reduce the depth dimession.

03 IRB: This block is the core structure of the animeGANv2. The IRB is also used to extract features but significantly reduces the parameters. The skip connection can be illustrated as: a[l+2] = g(w[l+2] x a[l+2] + b[l+2] + a[l])

04 Down-Conv: The block is used as down-sampling module and resize it 2 times.

05 Up-Conv: The block is used as up-sampling module and resize it 2 times.

06 Discriminator network: Contrary to the generator network, discriminator seems less complicated.

LOSS FUNCTION

01 Total loss: L(G, D) = Wadv x Ladv(G, D) + Wcon x Lcon(G, D) + Wgra x Lgra(G, D) + Wcol x Lcol(G, D).

02 Adversarial loss: Ladv(G, D) is the adversarial loss that affects the animation transformation process in the generator.

03 Content loss: Lcon(G, D) is the content loss which helps to make the generated image retain the content of the input photo.

04 Grayscale style loss: Lgra(G, D) is the grayscale style loss which makes the generated images have the clear anime style on the textures and lines.

04 Color reconstruction loss: Lcol(G, D) is used as the color reconstruction loss to make the generated images have the color of the original photos.

05 Wadv = 300,Wcon = 1.5, Wgra = 3 and Wcol = 10 which gives balance result of style and style preservation.

06 Generator’s loss function: L(G) = Wadv(Epi~Sdata(p))[(G(pi) - 1)2] + Wcon x Lcon(G, D) + Wgra x Lgra(G, D) + Wcol x Lcol(G, D)

07 Discriminator’s loss function: L(D) = Wadv[Eai~Sdata(a) [(D(ai) - 1)2] + Epi~Sdata(p) x [(D(G(pi)))2] + Exi~Sdata(x) [(D(xi))2] + 0.1Eyi~Sdata(y)[(D(yi))2]]

APPLICATION

HAYAO STYLE

APPLICATION

SHINKAI STYLE

APPLICATION

PAPRIKA STYLE

APPLICATION

APPLICATION

CONCLUSION

In conclusion, the project is first start evolving deep learning technique: Neural style transfer and GAN. The project is a fine combination of knowing the architecture and real application.

.Moreover, we can develop the project to a more specific task like creating a transformer to transform a long movie to a anime-like comic. The idea is best suit for person who prefer reading comics to watch a whole movie.