INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 6.078
(Volume 6, Issue 2)
Available online at: www.ijariit.com

Colorization of black and white images using generative adversarial networks

Ranjit Mahadik
ran.mahadik14@gmail.com
Vidyalankar Institute of Technology, Mumbai, Maharashtra

Rashi Ryapak
rashiryapak1@gmail.com
Vidyalankar Institute of Technology, Mumbai, Maharashtra

Yash Panchal
yashpanchal@gmail.com
Vidyalankar Institute of Technology, Mumbai, Maharashtra

Ankita Korde
ankita149korde@gmail.com
Vidyalankar Institute of Technology, Mumbai, Maharashtra

ABSTRACT

We present a generative adversarial network-based system that faithfully colorizes black and white images without human intervention. Recent methods for such problems typically use per-pixel loss between the output and ground-truth images and the images generated using such loss lacks details. where as in "Perceptual losses for real-time style transfer and super-resolution" [2] paper suggests the use of the perceptual loss function that generates high-quality images. We combine the benefits of both approaches. we have replaced pixel-loss function with perceptual loss function which gives visually pleasing results and used discriminative learning technique where we train first half of generator with lower learning rate as we are using pre-trained model and last half with higher learning rate which reduces the training time of generator. It's easier and faster to train a large number of samples of smaller images initially and then scale up the network by improving images to 220px by 220px from 64px by 64px. This is called progressive resizing [6]. it also helps the model to generalize better as is sees many more images of different shapes and less likely to be overfitting.

Keywords— *Generative Adversarial Networks, Transfer Learning, Progressive Resizing, Overfitting, Perceptual Loss Function*

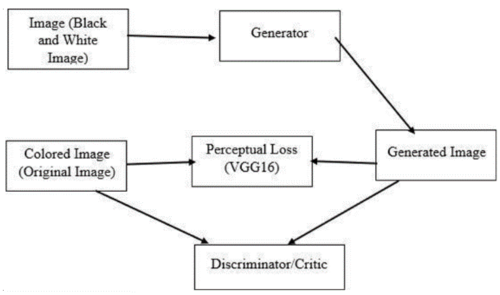
1. INTRODUCTION

Generative Adversarial Networks are the most active area of machine learning research. Every day 100-200 papers are published describing how they can be improved to generate high-quality images. Generative Adversarial Networks (GANs) are developed and introduced by Ian J. Goodfellow in 2014 [1]. It consists of two parts: a generator network and a discriminator network. The function of the generator is to generate data that resembles the training data. The Discriminator, on the other hand, is a Binary classifier that takes all the pairs of generated images and actual images and tries to classify which is generated and which is actual. In short generator tries fool discriminator. After a few iterations, the generator and discriminator get better and better in their respective job. In this paper we attempt to explore how generative adversarial networks can be improved and used to colorize black and white images. To achieve this, we are using state of the art techniques such as discriminative learning rate [8] which allows model to train faster and also using progressive resizing [6] to avoid overfitting and perceptual loss function [2] which allows model to generate high-quality images. The input to our system is a black and white image. We then use generative adversarial networks to output a prediction of a realistic colorization of the image.

2. PROPOSED SYSTEM

The problem this project attempts to explore is one where the black and white image is available but colour-image is not known and must be derived automatically. For each pixel of the grayscale target image, we look for a colour value to assign to this pixel. The main goal of the paper is to determine whether it is possible to use GANS for fully automated colorization of black-and-white (grayscale) images plausibly, as standalone images. To achieve this, we use GANS with customized loss function (perceptual loss [2]) and with other few tweaks like discriminative learning rates [8], transfer learning and progressive resizing [6].

2.1 Flowchart



2.2 Flowchart description

- Initially we started with an ImageNet dataset [7] and converted all the images (i.e. turn all images into black and white images). then trained a model to restore them to the original state.

- We are using perceptual loss function from “perceptual losses for real-time style transfer and super-resolution” [2] paper.
- Initially Generator and Discriminator are not trained simultaneously. Generator is trained first with perceptual loss function for few epochs. We try to make the Generator model as good as possible in this stage before moving to the discriminator training phase.
- Images Generated by Generated are stored. The discriminator model is a pre-trained ResNet-34 model trained on the ImageNet dataset [7] that takes all the pairs of generated images and actual images and tries to classify which is actual and which is generated.
- Now train both models according to literature [1].

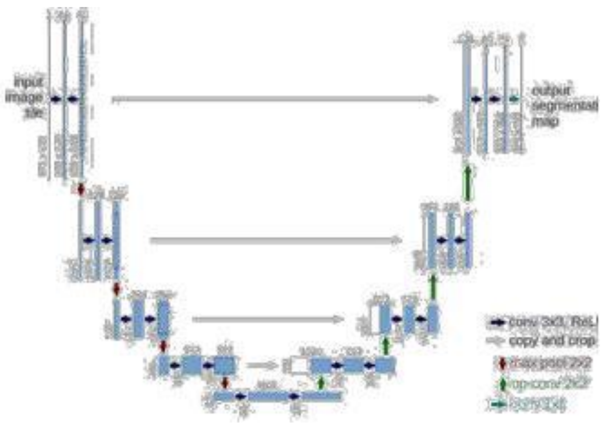
3. IMPLEMENTATION

3.1 Dataset

Initially the dataset for “Black and white images" was not available thus we downloaded ImageNet dataset from ImageNet [7] competition website and converted all the images to black and white. It’s easier and faster to train a large number of samples of smaller images initially and then scale up the network by improving images to 220px by 220px from 64px by 64px. This is called progressive resizing [6]. it also helps the model to generalize better as is sees many more images of different shapes and less likely to be overfitting.

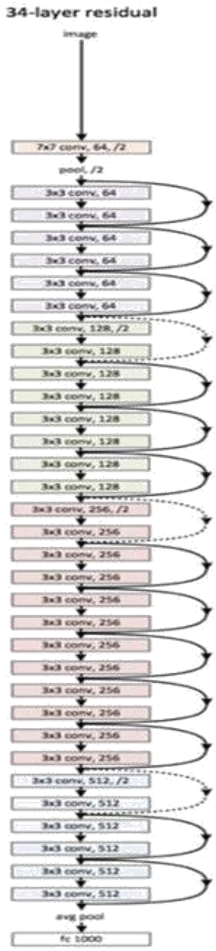
3.2 Generator Architecture

A U-Net [3] is a convolutional neural network architecture that was developed for biomedical image segmentation. U-Nets are very effective for tasks where the output is of similar size as the input. This makes them very good for creating segmentation masks and for image processing/generation such as super-resolution and image colorization. The down-sampling is the left-hand section of the U-Net network based on ResNet-34 [9]. The up-sampling part is the right-hand section of the U-Net of the network. The down-sampling has pre-trained weights based on ImageNet competition [7], i.e. Transfer Learning. During training we freeze or train down-sampling part with lower learning rate and the up-sampling part with a higher learning rate called a discriminative learning rate [8]. Which helps us to reduce the training time of generator.

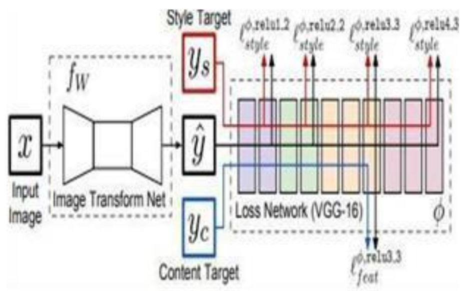


3.3 Discriminator/Critic Architecture

The Discriminator is a pre-trained ResNet-34 model which is trained in ImageNet Competition [7]. Discriminator is a Binary classifier that takes all the pairs of generated images and actual images and tries to classify which is generated and which is actual. In short generator tries fool discriminator.



3.4 Loss Function



The loss function is based upon the research paper “Perceptual Losses for Real-Time Style Transfer and Super-Resolution” [2]. This paper focuses on perceptual loss i.e. Feature Loss. The researchers did not use a U-Net architecture [3] as the deep learning community were not aware of them at that time.

The model used here is trained with a similar loss function to the paper, using VGG-16 model trained in ImageNet competition [7] but also combined with pixel mean squared error loss.

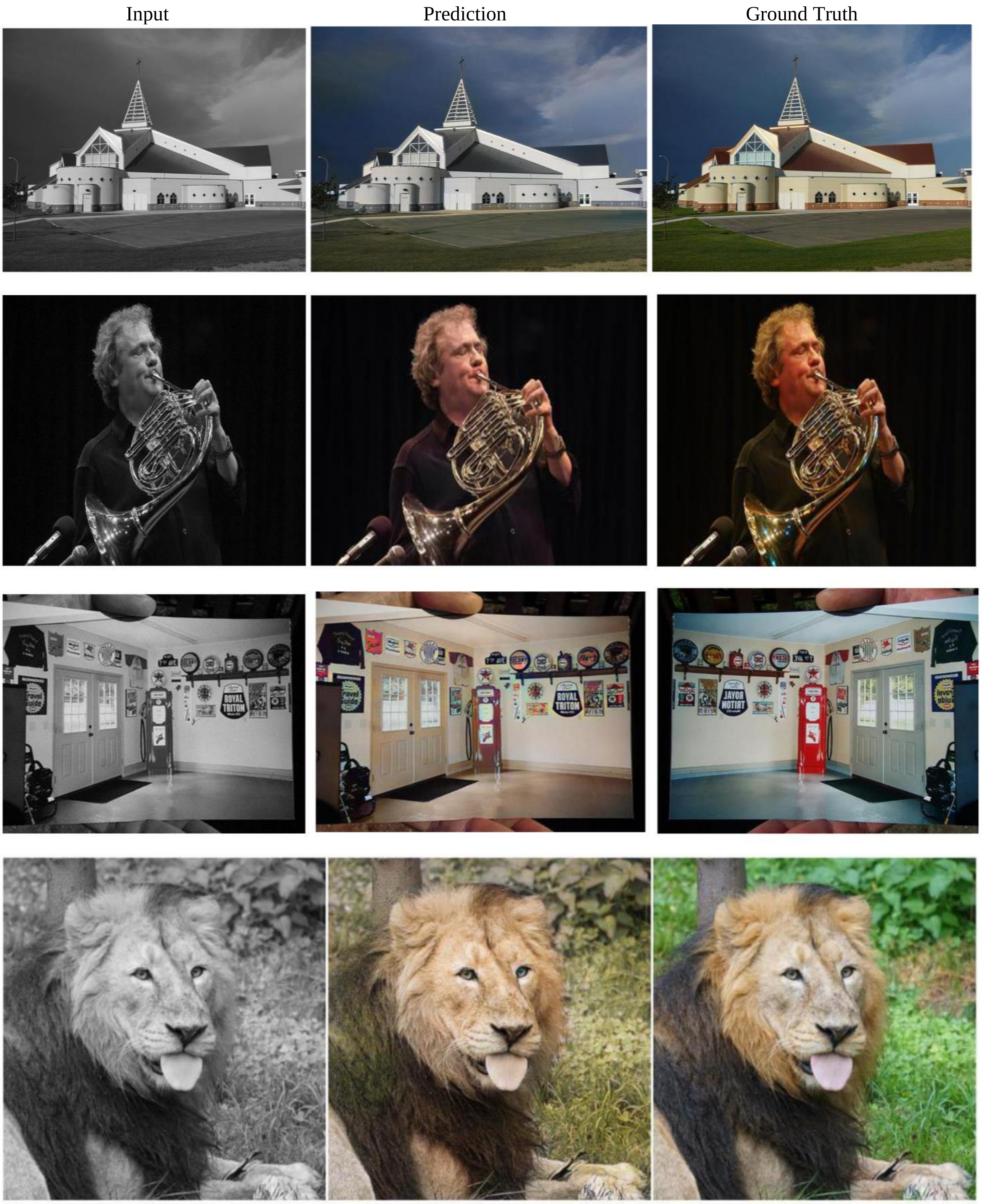
3.5 Training Details

Two methods are used here in the training process. These are progressive resizing [6] and discriminative learning rates [8]. The model’s architecture is split into two parts, the down-sampling part and the up-sampling part. The down-sampling has pre-trained weights based on ResNet34 trained on ImageNet Competition [7]. The down-sampling needs its weights training as these layers’ weights are randomly initialized to produce the desired end output. We train down-sampling part with a lower learning rate or freeze it and train up-sampling part with a higher learning rate.

It's easier and faster to train a large number of samples of smaller images initially and then scale up the network by improving images to 220px by 220px from 64px by 64px. This is called progressive resizing [6]. It also helps the model to generalize better as is sees many more images of different shapes and less likely to be overfitting.

Training Steps [9]:
- Pre-train the Generator: Initially Generator and Discriminator are not trained simultaneously. Generator is trained first with perceptual loss function [2] for a few epochs. We try to make the Generator model as good as possible in this stage before moving to the discriminator training phase.
- Save images generated by the generator.
- Training Critic/Discriminator: The Discriminator is a pre-trained ResNet-34 model trained in ImageNet Competition [7]. It is Binary Classifier. The model architecture is split into two parts, the second part is randomly initialized hence it needs training here we use discriminative learning rates [8]. We train the first part with a lower learning rate and the second part with a higher learning rate. i.e. we need to finetune discriminator.
- Train both generator and discriminator in the traditional manner.

4. RESULTS
The following images shows the results.



5. CONCLUSION

In this study, we were successful in automatically color the black and white images using generative adversarial networks, to visual degree which is acceptable. The images of ImageNet dataset with synthetic colors by GAN looked reasonably well and similar to the original images. There were some incidences where the model misunderstood the seawater for grass during the training process, but with further training, it was successful in coloring green color for grasses. We observed that the model faced an unusual problem with human faces, which it learned after many epochs as compared to other colors.

6. FUTURE WORK

A betterment in the image colorization could still be made with large datasets that are not available now. The model works very well with the image colorization. This method can be applied to colorize black and white videos as videois a continuous stream of images.

7. REFERENCES

[1] Ian. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In Advances in neural information processing systems, pages 2672– 2680, 2014.

[2] Justin Johnson, Alexandre Alahi, Li Fei-Fei Perceptual Losses for Real-Time Style Transfer and Super-Resolution.

[3] Olaf Ronneberger, Philipp Fischer, Thomas Brox U-Net: Convolutional Networks for Biomedical Image Segmentation.

[4] Han Zhang, Ian Goodfellow, Dimitris Metaxas, Augustus Odena Self-Attention Generative Adversarial Networks.

[5] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, Yuichi Yoshida Spectral Normalization for Generative Adversarial Networks.

[6] Tero Karras, Timo Aila, Samuli Laine, Jaakko Lehtinen Progressive Growing of GANs for Improved Quality, Stability, and Variation.

[7] Imagenet Dataset http://www.image-net.org

[8] Universal Language Model Fine-tuning for Text Classification Jeremy Howard, Sebastian Ruder 2018.

[9] Fastai Gan https://docs.fast.ai/vision.gan.html

