Old Movies Colorization using Deep Learning Techniques

Chandan Agrawal (170102077), Mukul Ranjan (170102084)
Supervisors – Dr. P. K. Bora, Dr. A. Rajesh

Introduction

Colors carry a lot of information about a photograph. They can express and influence the viewer's mood, they are able to draw viewers attention to certain parts of the image and overall we can say they add realism to the scene in the photograph. This is the reason why a lot of old movies are colorized at the huge expense of money and manual labor. In this project we aim to build a generalized but robust model for automatic colorization of old hindi movies of indian cinema. We mainly make use of generative adversarial networks to execute this task.

Milestones

- First objective is to build an initial basic colorization model for individual images.
  - Literature review of different colorization models for images.
  - Reproduction of GAN (Pix2Pix and Self Attention GAN) models.
- Improve the model by multi-task learning with tasks such as scene identification which can help better colorization.
- Fine tune the model for old indian movies settings.
  - Create a movies dataset of movies whose domain closely matches to the domain of old black and white movies.
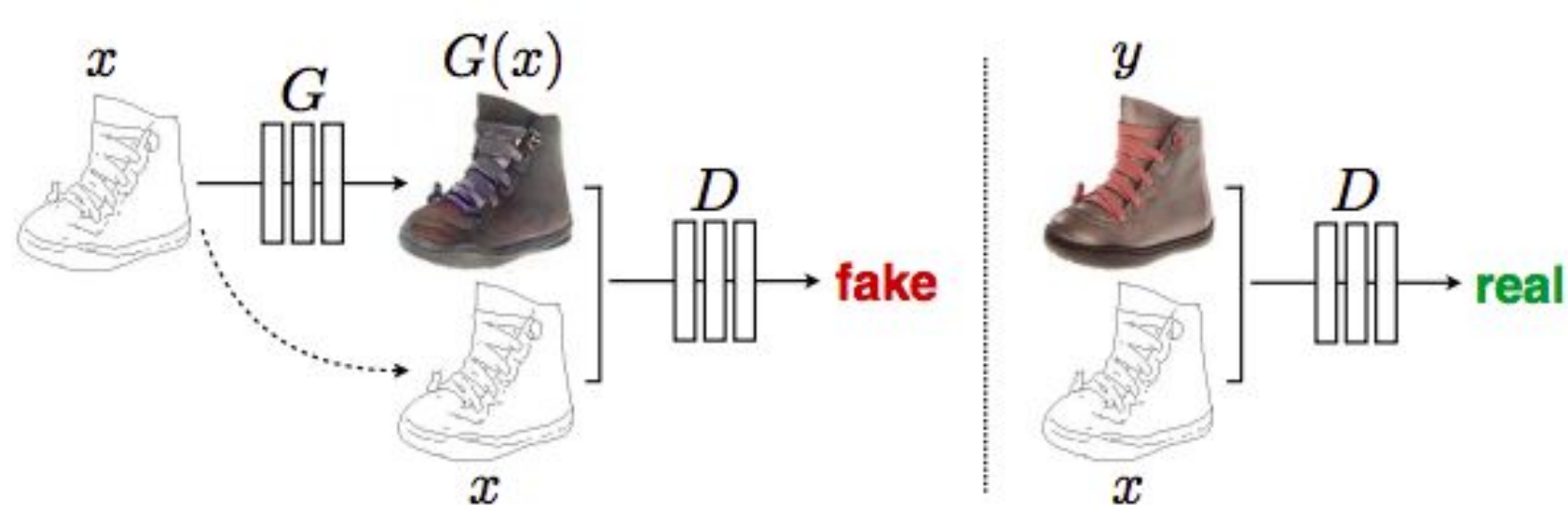- Ensure a flicker free video generation.
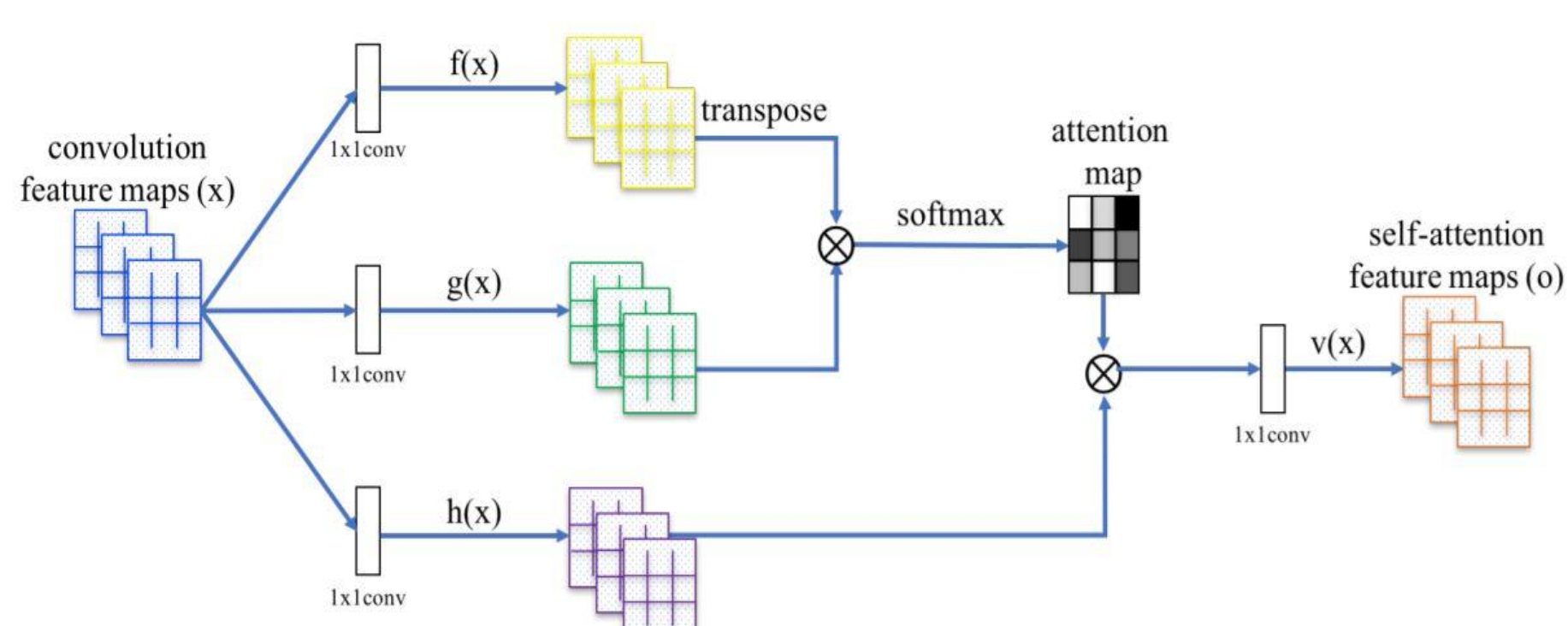
Design

fig. 1: Schematic of Pix2Pix model

fig. 2: Attention Module

Method

The input of our model is a grayscale image and its corresponding ground truth is the original color image. As every available color image can converted into a grayscale image, we actually have infinite data for this task. We first transform a RGB image into Lab color space image without any loss. This way, we separate the image based on lightness(L) and chrominance(ab). Here the L channel image is nothing but a grayscale version of the color image and ab is 2-dimensional linear color space. We set L channel image as our input and ab channels as the ground truths.

We use image to image translation model of GAN, i.e. Pix2Pix. We have selected UNet architecture as the generator. It is designed for square images of size 64. The discriminator is a simple convolutional network whose input is pair of input image and ground truth or output of the generator (Generated image). We input the L channel grayscale image to the generator whose expected output is a 2 channel image of same height and width. We send this pair of L channel image and 2 channel generator output into the discriminator. The discriminator outputs a probability of this pair being real or not. We use cross entropy loss as the objective function. For the generator along with cross entropy loss, we use L_1 loss. Other hyperparameters such as batch size, learning rate etc. are set for optimal performance. We also plan to add the self attention module into both the networks as it helps to model long range dependencies.

Results
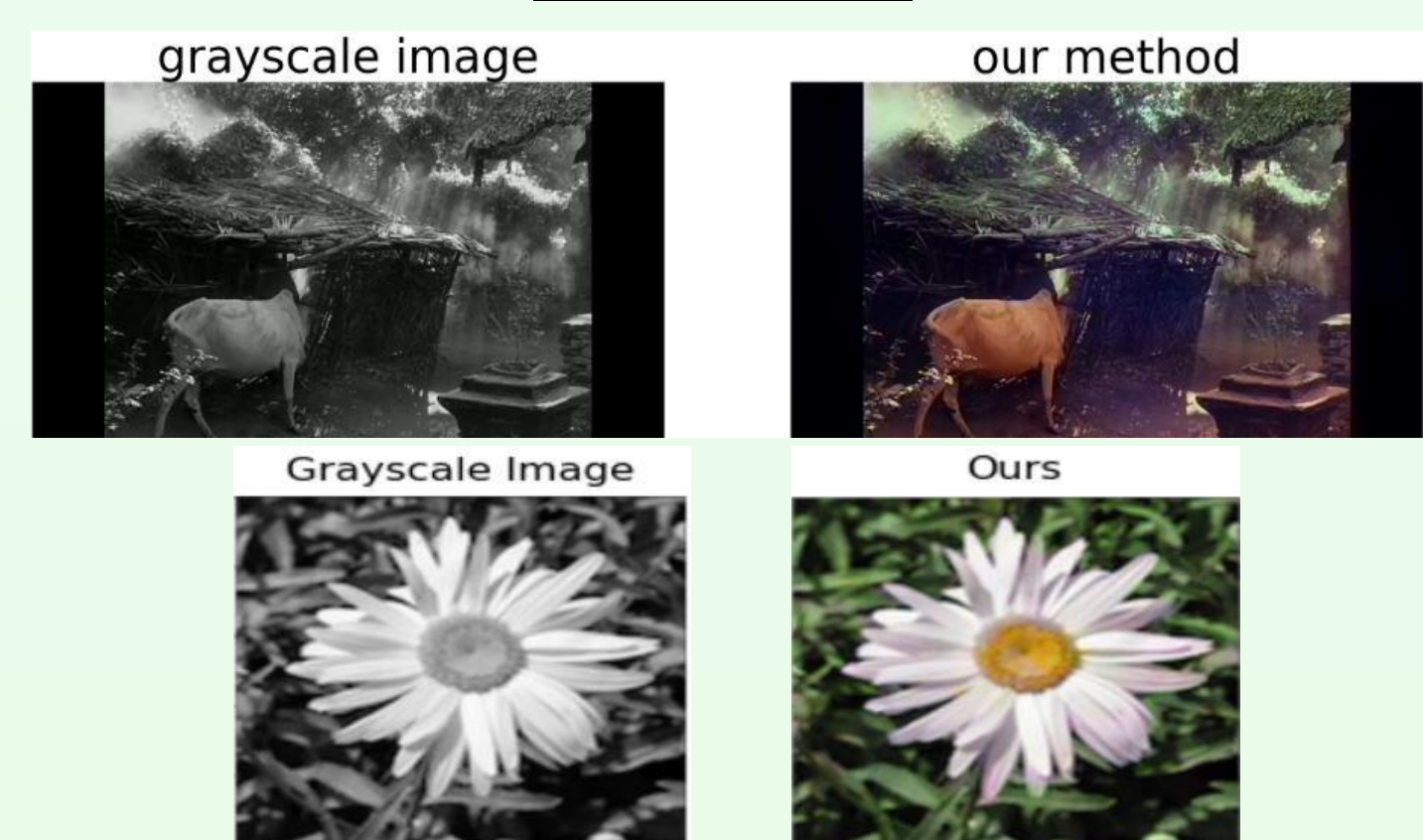
Future Work

- Colorization is a complex task which requires a model to detect and segment each and every object in the image to select appropriate colors. Hence we plan to build a multi task learning model with tasks such as object classification, segmentation and colorization.
- As observed by previous works, colorization is a domain specific task as it has an element of uncertainty attached as to what color to choose for a particular object. So we plan to create a separate dataset of movies with domain as close as possible to the old indian movies.

References

1. R. Zhang, P. Isola, and A. A. Efros. Colorful image colorization. In European conference on computer vision, pages 649-666. Springer, 2016.
2. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1125{1134, 2017.
3. H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena. Self-attention generative adversarial networks. In International Conference on Machine Learning, pages 7354-7363. PMLR, 2019.