Final Project Report

Image Colorization using GANs

GitHub Link

Pravallika Srinivas Naidu

Cognitive Science, University of Osnabrück

Implementing Artificial Neural Networks with Tensorflow (8.3304)

15 April 2021

1.Introduction

1. 1 Image Colorization

It can be defined as the process of transforming the grayscale images into colored images by adding plausible color information to monochrome counterparts.

1. 2 General Adversarial Networks

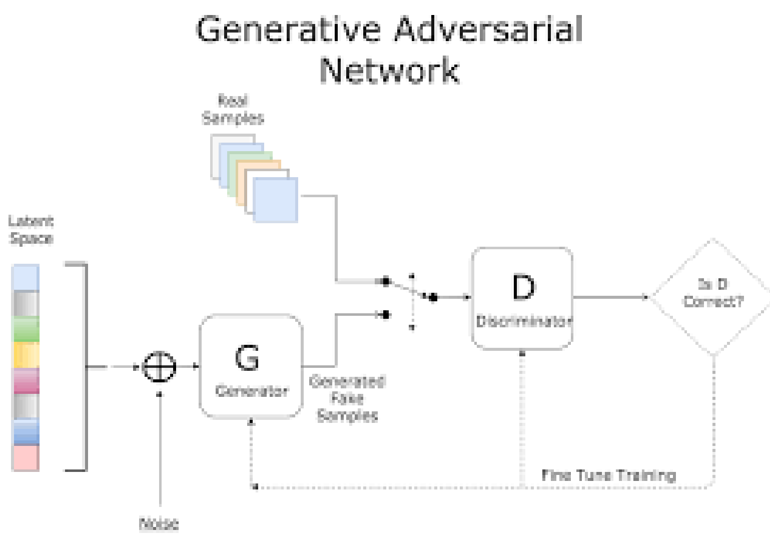


Image from (5)

General Adversarial Networks or simply GANs is the type of neural network used for generative Modelling, where the model is used to generate new samples from an existing distribution. They are trained using two models, namely Generator or Generative network that learns to generate new samples from a latent space/ noise and a Discriminator or Discriminative Network whose sole job is to identify the fake samples from the real ones. However, it’s work is most notable in the image to image translation problems like season transfer, style transfer, colorizing black and white images, it being the main focus in our project.

1. 3 Analysis of previous work in image processing

As accurately described by Dhir. R at el (2021) (1), there have been several studies in comparison for image processing. Below is the short description of such works.

No .	Reference paper	Methodology	Limitations
1.	Generative adversarial nets by Ian Goodfellow in 2014	Internal working and concept of GANs	Generator and discriminator must be well synchronized during training in order to avoid "Helvetica scenario" There is no explicit representation of p(x) in the model
2.	Emotional image color transfer via deep learning by Da Liu, Yaxi Jiang, Min Pei, Shiguang Liu in 2018	Mitigates the problem of unnatural coloring taking place in colorizing problems	Training is time consuming
3.	Thermal infrared colorization via conditional generative adversarial network by Xiaodong Kuang, Jianfei Zhu, Xiubao Sui, Yuan Liu, Chengwei in 2019	The method uses a composite objective function to produce finely detailed and	Encounters poor results with blurry or distorted image details
4.	Context-aware colorization of gray-scale images utilizing a cycle-consistent generative adversarial network architecture by Mohammad Mahdi Johari, Hamid Behroozi in 2020	Parallel colorization models introduced as opposed to traditional single models	Does not deal with pixel-to-pixel mapping
5.	Optimization based grayscale image colorization, Nie Dongdong, Ma Qinyong, Ma Lizhuang, Xiao Shuangjiu, 2007	Optimization approach to colorization that reduces computational time and reduces color diffusions	Spatial-temporal approach to be developed for maintaining temporal coherence

2. Methods and Models

We use a special type of GAN which will take in grayscale images, and output their RGB counterparts. The generator will have an encoder-decoder structure with layers placed symmetrically, similar to a UNet and we primarily work on the Places 365 dataset consisting of 5000 images of aquariums. The goal is to build a successful GAN model using Google Collab with limited dataset and resources to colorize images.

Here, the generator model takes a grayscale image consisting of 1-channel, namely the L channel which represents the lightness of the image and produces a 2-channel image. The discriminator takes these two produced channels and concatenates them with the input grayscale image and decides whether this new 3-channel image is fake or real. The discriminator is also given some real images (3-channel images) that are not produced by the generator and learn that they are real.

The encoder in the generator will take a grayscale image and produce a latent representation of it. The decoder will then produce an RGB image by enlarging this latent representation and uses MSE loss function. The discriminator used is a standard CNN used for classification. It takes an image and outputs a probability of whether the given image is an original or not (generated image). Binary cross-entropy loss is used for both the outputs of the discriminator which is then summed to obtain the final loss function.

3. Training, Results, & Ablation Studies

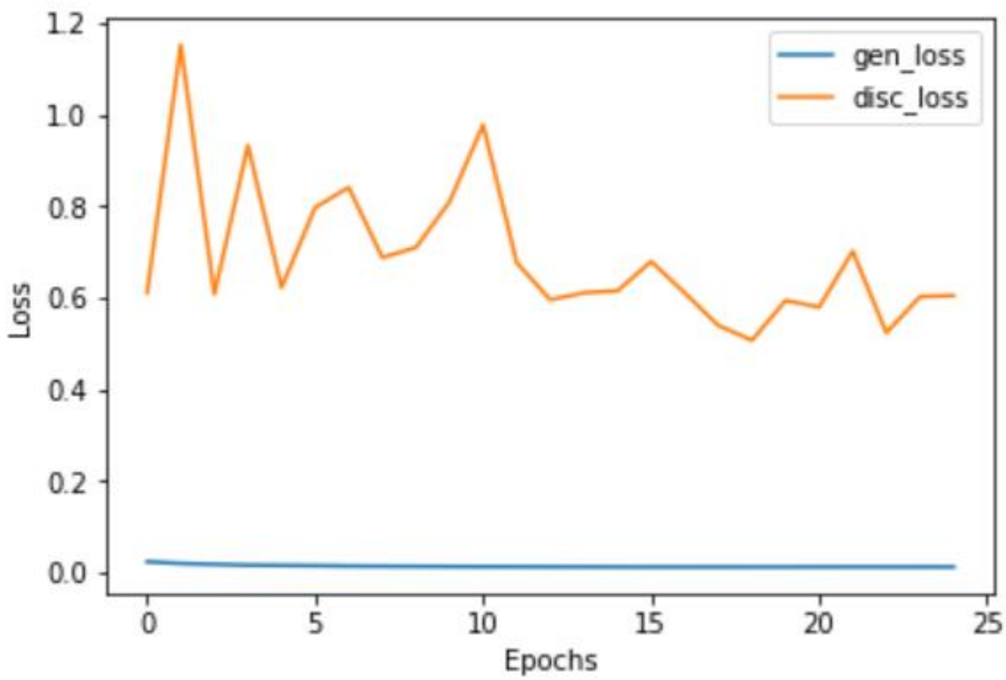
3. 1 Training

The training loop in general performs a train step in which the model generates predictions from Generator and Discriminator, calculates the losses and optimizes the model using Adam Optimizer with learning rate 0.001 which was chosen after trying out several trials on different learning rates. The model is trained for 50 epochs on the dataset with batch size being 32.

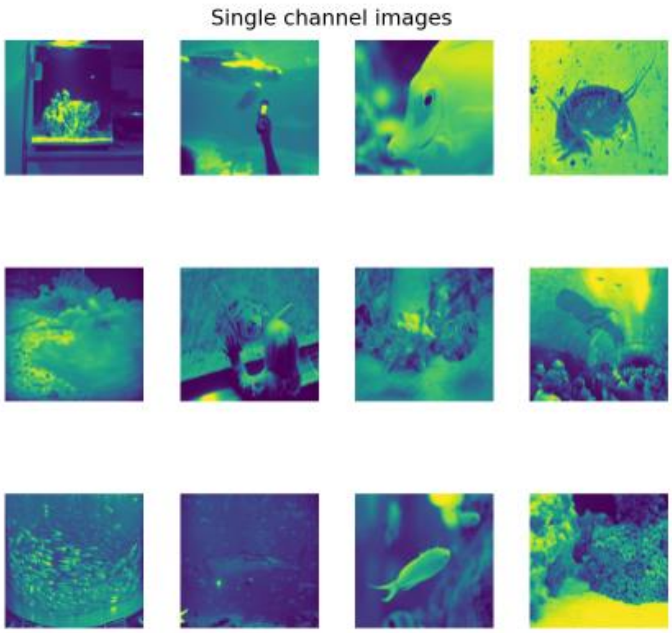
3. 2 Results and Analysis

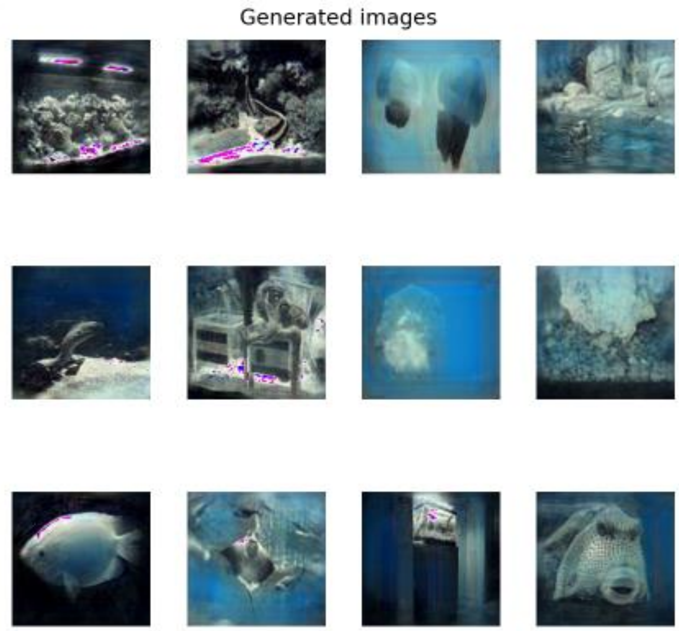
The results were successful to an extent, where a clear distinction between the grayscale and the colorized images generated could be differentiated which is acceptable. It was successful in majority of blue coloring since our chosen dataset consisted of mostly aquariums explaining the blue. However, there exists some interference which is quite distinct from the red/pink patches in the background.

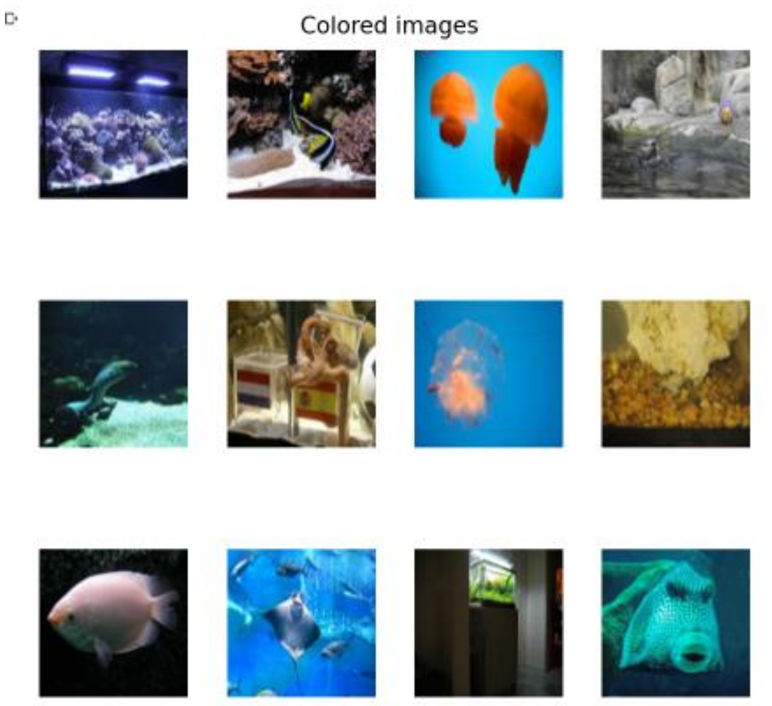
Below is the plot showing Generator and Discriminator losses:



Plots for: Single channeled images vs Colored images by Generator vs Ground truth color images;







3. 3 Ablation Studies

Using the Lab color space instead of RGB, we were able to isolate the grayscale channel which is the L channel used by the generator as input to predict A and B which reduced the complexity of the color channels from three to two(4).

The generator model also employed dilation rates which in turn uses dilated convolution, helping us in detection of fine-details by processing inputs in higher resolutions, broader view of the input to capture more contextual information (2)(3).

Skip connection used in the generator model concatenating the output directly from some parts of the encoder to the decoder which helped the reconstruction of the images from latent representation.

References:

1. Dhir R., Ashok M., Gite S., Kotecha K. (2021) Automatic Image Colorization Using GANs. In: Patel K.K., Garg D., Patel A., Lingras P. (eds) Soft Computing and its Engineering Applications. icSoftComp 2020. Communications in Computer and Information Science, vol 1374. Springer, Singapore. https://doi.org/10.1007/978-981-16-0708-0_2
2. Seo, J. D. (2018, July 8). Understanding 2D Dilated Convolution Operation with Examples in Numpy and Tensorflow with Interactive Code. Medium. https://towardsdatascience.com/understanding-2d-dilated-convolution-operation-with-examples-in-numpy-and-tensorflow-with-d376b3972b25
3. E. (2018, April 5). Dilated Convolution. A Blog From Human-Engineer-Being. https://erogol.com/dilated-convolution/
4. Shariatnia, M. (2020, November 19). Colorizing black & white images with U-Net and conditional GAN — A Tutorial | Towards Data Science. Medium. https://towardsdatascience.com/colorizing-black-white-images-with-u-net-and-conditional-gan-a-tutorial-81b2df111cd8
5. Mosquera, D. G. (2020, May 19). GANs from Scratch 1: A deep introduction. With code in PyTorch and TensorFlow. Medium. https://medium.com/ai-society/gans-from-scratch-1-a-deep-introduction-with-code-in-pytorch-and-tensorflow-cb03cdcdba0f