Comparing Regression, Autoencoder and Generative Adversarial Network (GAN) for Image Colorization

Vikram Bharadwaj, Sudhanva Narayana
CS6140 Machine Learning
Professor: Dr. Ehsan Elhamifar

May 3, 2022

Abstract

Every color image is a combination of 3 channels namely: R, G, and B. Another way used to represent images is called the L, *a, *b (CIELAB) colorspace.

The three channels are described in detail below: L - Lightness channel, which represents the illuminatoion or the color temperature of the image. *a *b - Chromaticity channels, represents the color of the image, that is a combination of red, green, blue, and yellow, the four primary colors. Traditional image colorization methods often use regression to predict the *a and *b channel values, where often times, the obtained color images do not reflect the actual image.

In this study, we aim to compare the performance of regression, an autoencoder and GANs for image colorization.

1 Introduction

The idea is to use grayscale images for complex image processing. The goal of this study is to develop a method for colorizing grayscale images using Regression and compare how the model performs against the use of an autoencoder and a GAN.

For image colorization, we will use three methods.

- **Convolutional Neural Netwrok + Regression:** Loss between actual values of *a and *b channels and predicted *a and *b channels using vanilla CNN, followed by regression.
- **Convnet Autoencoder:** Uses CNN in an encoder-decoder fashion, where the encoder represents the grayscale image as a latent representation, which the decoder then converts to a 3-channel color image.
- **GAN:** Uses a custom minimax log-loss function with a generator and a discriminator. The generator takes a grayscale image as an input, which it tries to convert to a 3-channel color image. The discriminator tries to tell 2 sets of images apart, i.e., (grayscale, original-color) and (grayscale, generator-color-image).

1.1 Related Work

The idea is to use grayscale images for complex image processing. The goal of this study is to develop a method for colorizing grayscale images using Regression and compare how the model performs against the use of an autoencoder and a GAN.

2 Technical Details

The idea is to use grayscale images for complex image processing. The goal of this study is to develop a method for colorizing grayscale images using Regression and compare how the model performs against the use of an autoencoder and a GAN.

2.1 Method 1: Regression

The idea is to use grayscale images for complex image processing. The goal of this study is to develop a method for colorizing grayscale images using Regression and compare how the model performs against the use of an autoencoder and a GAN.

2.2 Method 2: Autoencoder

The idea is to use grayscale images for complex image processing. The goal of this study is to develop a method for colorizing grayscale images using Regression and compare how the model performs against the use of an autoencoder and a GAN.

2.3 Method 3: Generative Adversarial Network

First you have to upload the image file from your computer using the upload link in the file-tree menu. Then use the includegraphics command to include it in your document. Use the figure environment and the caption command to add a number and a caption to your figure. See the code for Figure **??** in this section for an example.

Note that your figure will automatically be placed in the most appropriate place for it, given the surrounding text and taking into account other figures or tables that may be close by. You can find out more about adding images to your documents in this help article on including images on Overleaf.

3 Technical Details

The idea is to use grayscale images for complex image processing. The goal of this study is to develop a method for colorizing grayscale images using Regression and compare how the model performs against the use of an autoencoder and a GAN.

3.1 Dataset: MIT Places

We will be using a color-filled landscape dataset called Places 365 from MIT CSAILVision Lab. The dataset contains images of natural scenes and their corresponding colorizations. The dataset is available at https://places.csail.mit.edu/places365/. Places contains more than 10 million images comprising 400+ unique scene categories. The dataset features 5000 to 30,000 training images per class, consistent with real-world frequencies of occurrence.

3.2 Dataset: Preprocessing

The idea is to use grayscale images for complex image processing. The goal of this study is to develop a method for colorizing grayscale images using Regression and compare how the model performs against the use of an autoencoder and a GAN.

3.2.1 Pre-processing for Regression

For the first method, where we leverage pure CNN with regression, the following are the pre-processing steps:
- All images are first converted to L*a*b channel image from RGB channel.
- The L channel first extracted from the L*a*b channel image and this serves as the input to the regression model.
- The *a and *b channels are used as the tragets for the regression model.
- The images are then resized to a fixed size of 256x256.
- The images are then normalized to the range of [0, 1] for the L channel and [-1, 1] for *a and *b channels.
- The images are then divided into training and test sets, with a train split of 75%
- The training set is then used to train the model.

3.2.2 Pre-processing for Convnet based encoder-decoder and GAN

Here, we will use the same pre-processing steps as for the regression method. The only difference is that we will use a CNN in an encoder-decoder fashion for the second method and a GAN for the third, as shown below:

3.3 Implementation

The idea is to use grayscale images for complex image processing. The goal of this study is to develop a method for colorizing grayscale images using Regression and compare how the model performs against the use of an autoencoder and a GAN.

3.4 Results

The idea is to use grayscale images for complex image processing. The goal of this study is to develop a method for colorizing grayscale images using Regression and compare how the model performs against the use of an autoencoder and a GAN.

4 Limitations and Further Work

The idea is to use grayscale images for complex image processing. The goal of this study is to develop a method for colorizing grayscale images using Regression and compare how the model performs against the use of an autoencoder and a GAN.

5 Contribution

The idea is to use grayscale images for complex image processing. The goal of this study is to develop a method for colorizing grayscale images using Regression and compare how the model performs against the use of an autoencoder and a GAN. [BMR17] [CZP+18a] [CZP+18b] [LMS16] [RDB16] [ZIE16]

References

[BMR17] Federico Baldassarre, Diego González Morín, and Lucas Rodés-Guirao. Deep koalarization: Image colorization using cnns and inception-resnet-v2. *CoRR*, abs/1712.03400, 2017.

[CZP+18a] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. *CoRR*, abs/1802.02611, 2018.

[CZP+18b] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. *CoRR*, abs/1802.02611, 2018.

[LMS16] Gustav Larsson, Michael Maire, and Gregory Shakhnarovich. Learning representations for automatic colorization. *CoRR*, abs/1603.06668, 2016.

[RDB16] Manuel Ruder, Alexey Dosovitskiy, and Thomas Brox. Artistic style transfer for videos. *CoRR*, abs/1604.08610, 2016.

[ZIE16] Richard Zhang, Phillip Isola, and Alexei A. Efros. Colorful image colorization. *CoRR*, abs/1603.08511, 2016.