Discriminative Models on Image Colorization

Xinyu Sun *
Department of Computer Science
University of Toronto
`xinyusun@cs.toronto.edu`

Fan Shen *
Department of Computer Science
University of Toronto
`van12@cs.toronto.edu`

Abstract

In this project, we plan to investigate the performance of discriminative models on colorization of grey-scale images. We will conduct a systematic analysis from network architectures, hyperparameters, regression approach and classification approach. We also want to devise a new model to merge the advantages based on our analysis. If our idea is feasible, we hope to get lower per-pixel chrominance loss result on CIFAR-10 [1] dataset than the original models.

1 Introduction

Automated colorization is a task with growing interest in computer vision and machine learning field. Given a grey-scale image, humans are possible to fill in the realistic colors with the existing clues in image. This shows that the grey-scale image is sufficient for a human to colorize. However, it is very difficult to retrieve the exact colors from grey-scale image for a machine learning model since there is a large amount of information loss from a colored image to grey-scale one. Thus, the accuracy will not be very high, especially for those contained items with various possible colors like car.

In this project, we would like to learn and compare some existing discriminative models trained by colored image datasets. After analysing, we desire to build a new model for automated colorization which merge the advantages of the existing models. Our model will take a grey-scale image as input and produces a colorized image with no additional human interpretation. To get the impact of various possible colors to the loss, we will use CIFAR-10 dataset, which contains natural images like bird, horse and images with various possible colors like truck, automobile. Hopefully, we will get lower per-pixel chrominance loss by our new model compared to the models we analysed.

2 Related Work

Some existing models like UNet [2], VGG [3] and ResNet [4] have already shown the possibility that they could be applied to colorizing grey-scale images. UNet [2] has shown that connections between hidden layers of a bottleneck neural network could enhance the performance greatly. Later, ColorUNet [5] was proposed and proved that a rather lightweight architecture like UNet could also achieve satisfactory results image colorization. Similarly, as stated in the original UNet paper, the authors of ColorUNet found that data augmentation could also significantly improve the performance and robustness of the model.

With the amount of image data increases, deeper networks started to be applied for image colorization. Larsson et al. [6] trained VGG-styled models on ImageNet and made the model to predict a color histogram, instead of a single color, which provides a new evaluation method. Guadarrama et al. [7] proposed PixelCNN, which used ResNet to extract features and then trained a second CNN to generate a high-resolution colorization of an image. This approach produces a lot of diverse and plausible colorizations than existing methods. However, compared to a lightweight network like

*These authors contributed equally to this work.

UNet, even though deeper networks can lead to better results, they are usually quite complex, and training it to a satisfactory degree requires significant amount of computation resources.

3 Methods

3.1 Problem Definition

In this project, we decide to use the YUV colorspace to represent color images instead of the RGB colorspace, since the YUV colorspace can minimize the correlation between the color channels. In these three channels, Y channel encodes the luminance component of the image, which can be regarded as the grayscale of the image; U,V channels represent the chrominance, which encode the colors of the corresponding pixels.

Therefore, we formalize this problem as: Given a image set X, take $x \in X$, then our model $F(Y)$ will take the Y channel of x and output the correct U and V channels, which is $F(Y) = (U, V)$.

3.2 Experiment Setup

In this project, we will use CIFAR-10, which consists of 60,000, 32×32 RGB color images in 10 classes. The RGB images will be converted to the YUV colorspace. We will randomly divide the dataset into 50,000 training set, 5,000 validation set and 5,000 test set.

3.3 Loss Functions

Since we also want to conduct experiments between classification and regression approaches, we have the following two loss functions:

For classification networks, we will take a similar method as Programming Assignment 2, which divides the color into 20 bins, and then use cross entropy function to calculate the loss for each pixel.

$$L_{cls} = \frac{1}{N} \sum_{pixel} (L_{CE}(U_{pred}, U_{gt}) + L_{CE}(V_{pred}, V_{gt})) \tag{1}$$

For regression, we calculate MSE of the predicted image pixels and actual image pixels in the U and V channels:

$$L_{reg} = \frac{1}{N} \sum_{pixel} (\|U_{pred} - U_{gt}\|^2 + \|V_{pred} - V_{gt}\|^2) \tag{2}$$

References

[1] A. Krizhevsky *et al.*, "Learning multiple layers of features from tiny images," 2009.

[2] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.

[3] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[5] V. Billaut, M. de Rochemonteix, and M. Thibault, "Colorunet: A convolutional classification approach to colorization," 2018.

[6] G. Larsson, M. Maire, and G. Shakhnarovich, "Learning representations for automatic colorization," in *European Conference on Computer Vision*, pp. 577–593, Springer, 2016.

[7] S. Guadarrama, R. Dahl, D. Bieber, M. Norouzi, J. Shlens, and K. Murphy, "Pixcolor: Pixel recursive colorization," *arXiv preprint arXiv:1705.07208*, 2017.