

Automatic Black & White Images colorization using Convolutional neural network

*¹Omar Abdulwahhab Othman, ¹Betül Uzbaş and ¹Sait Ali Uymaz

¹Faculty of Engineering and Natural Sciences, Department of Computer Engineering, Konya Technical University, Konya, Turkey

Abstract

In this paper, automatic black and white image colorization method has been proposed. The study is based on the best-known deep learning algorithm CNN (Convolutional neural network). The Model that developed taking the input in gray scale and predict the color of image based on the dataset that trained on it. The color space used in this work is Lab Color space the model takes the L channel as the input and the ab channels as the output. The Image Net dataset used and random selected image have been used to construct a mini dataset of images that contains 39,604 images splitted into 80% training and 20% testing. The proposed method has been tested and evaluated on samples images with Mean-squared error and peak signal to noise ratio and reached an average of MSE= 51.36 and PSNR= 31.

Key words: CNN, Image colorization, automatic colorization, Lab color space, deep learning.

1. Introduction

The automatic colorization is the task of adding the color to a given black and white image as input without any user effort. This is a problem because objects have many colors, for example the plastic can have random colors, trees can have various degree of green color making the automatic colorization without prior knowledge a challenging task. The prior knowledge is the user intervention. Most colorization approaches allow the user to select the color of the objects in the image and determine the color of some areas to extend the information to whole image. In the last years, researchers proposed many approaches such as defining a color flow function on neighboring pixels and estimates the local gray scale intensity variations function [1-3]. Another approach is by using a pre-defined threshold [4]. The user-based methods of image colorization have an advantage over the automatic colorization where the user select more color points to obtain the best results of the whole image by placing the points of color to give indirect information on the location of the color boundaries. With the rise of deep learning algorithms in recent years making solutions for various problems in the computer vision field, especially the Convolutional neural networks [5] which used in the image processing and features extraction. The convolutional neural networks can be used in the image colorization by utilizing model architecture for predicting the images colors by training the model with huge amount of data and

*Corresponding author: Address: ¹Faculty of Engineering and Natural Sciences, Department of Computer Engineering, Konya Technical University, Konya TURKEY. E-mail address: omer.salihi92@gmail.com, Phone: +905380609521

examples to learn how to reconstruct the color of an image with such objects and features. Predicting the color of image making it possible for the fully automatic colorization. In this paper, we construct convolutional neural network architecture that takes an input of gray scale image and an output of the image color. We used the Lab color space and convert all of the images to this color space then we split the image channels where the L channel is the input of the model and the ab is the output of it. When the model predicts the ab channel from an input of L channel (gray scale) the program will combine the channels to create a Lab color space image and then convert it back to RGB color space. For the training purpose, we used random selected images from ImageNet dataset the total number of images used is 39,604 splitted into 80% for training and 20% for testing.

2. Materials and Method

2.1. Convolutional Neural network(CNN):

CNN is the most popular deep learning algorithm for computer vision and image processing applications proposed firstly by YannLecun[6]. CNN is a regularized multi perceptron version, where each neuron in the layers is connected to all neurons the next layer which make it prone to over fitting data. The CNNs are inspired from the biology[7-10], especially the visual cortex of the animals. The CNN consists of many types of layers, the Convolution layer is a set of learnable filters, where each filter has a small width, height and same depth of the input volume. The convolution layer computes the dot product of all filters and image patch. The activation functions in the CNN apply element wise activation function to the convolution layer output, activation functions that most used is ReLU(Rectified linear units) , sigmoid and tanh. The other main layer of the CNN is the pooling layer which uses to reduce the size of the volume to make the computation faster and reduce memory usage and prevent over fitting in the network. The last layer in the CNN is the fully connected layer which is a regular neural network that takes inputs from the previous layer and computes the class scores. Figure 1 shows an example of CNN architecture.

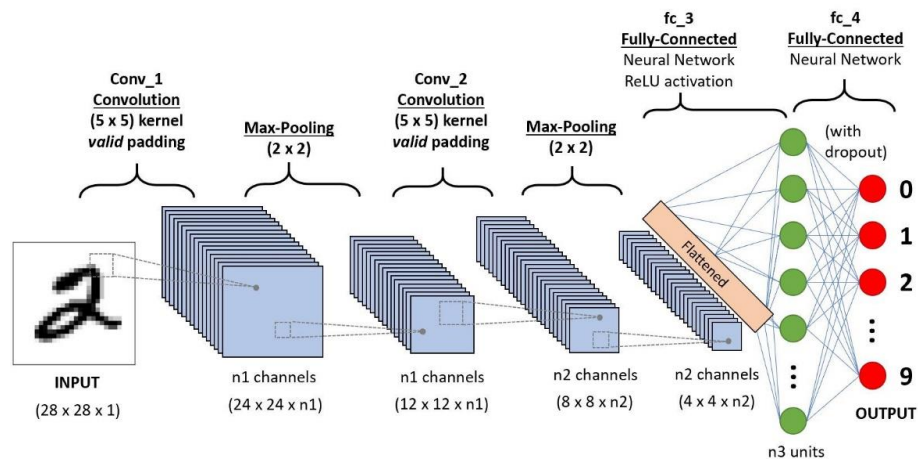


Figure 1. Example of CNN architecture.

2.2. Dataset

The dataset used in this study, is the ImageNet dataset [11]. ImageNet is an ongoing research project to provide an accessible images database for the researchers around the world. The ImageNet is organized according WordNet hierarchy. Each class in the ImageNet contains 1000 images. Figure 2 shows sample images from imagenet.

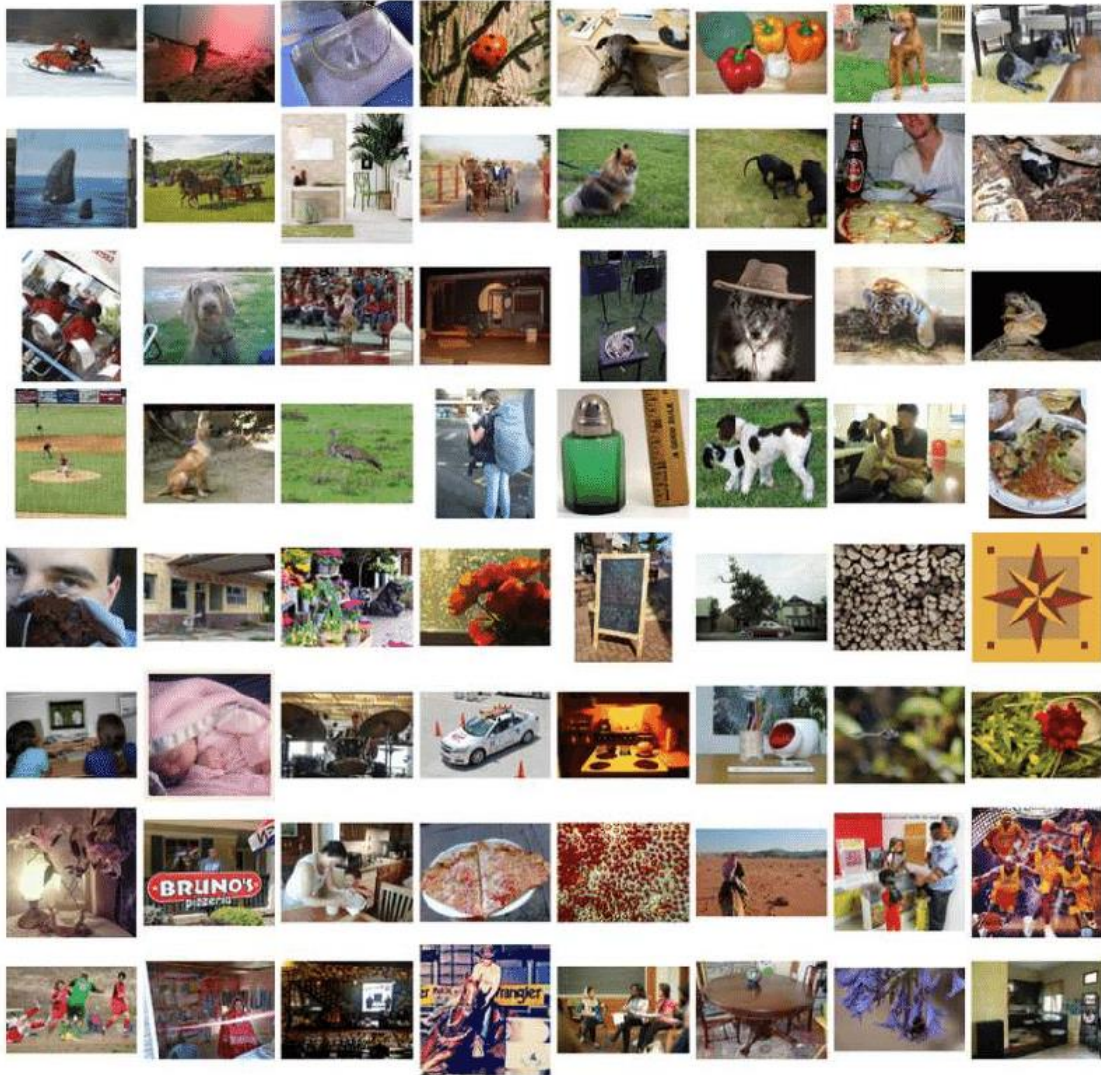


Figure 2. ImageNet sample images.

2.3. Proposed approach

We constructed a CNN model architecture for the purpose of Black and white images colorization. The Color space used is the Lab, all images used converted during the training and testing to Lab color space, the L channel of each image used as the input of the model, where the

output is the ab channel which define the image colors while the L channel is a grayscale. Table 1 shows the architecture of the model.

Table 1. Colorization model architecture.

Layer	Filters	Activation	Regularizer
Conv2D	64	ReLU	L2
Conv2D	64	ReLU	L2
Batch Normalization	None	None	None
Conv2D	128	ReLU	L2
Conv2D	128	ReLU	L2
Batch Normalization	None	None	None
Conv2D	256	ReLU	L2
Conv2D	256	ReLU	L2
Conv2D	256	ReLU	L2
Batch Normalization	None	None	None
Conv2D x 3	512	ReLU	L2
Conv2D x 3	512	ReLU	L2
Conv2D x 3	512	ReLU	L2
Batch Normalization x 3	None	None	None
Conv2D	256	ReLU	L2
Conv2D	256	ReLU	L2
Conv2D	256	ReLU	L2
Batch Normalization	None	None	None
Up Sampling 2D	None	None	None
Conv2D	128	ReLU	L2
Conv2D	128	ReLU	L2
Conv2D	128	ReLU	L2
Batch Normalization	None	None	None
Conv2D	313	Softmax	None

As shown in the Table 1 the model consists of Convolutional layers with different filters, the first block contains 2 convolutional layers with 64 filters regularized by L2 function and followed by Batch Normalization layer which used to avoid the overfitting in the network. The second block also consist 2 convolutional layers with 128 filters and also with L2 function used as regularizer and followed with batch normalization layer for the same reason. The Main activation function used in the Convolutional layers is ReLU function except the last layer which used SoftMax function. As shown the Up-sampling layer is used to up sample of the input volume to reach the needed size.

3. Results and Experiments

We trained the model described in section 4 on randomly selected images from ImageNet dataset, the total number of images is 39,604, splitted into 80% for the training and 20% for the testing. The performance of the model evaluated by using categorical cross entropy that optimized using SGD optimizer. The results of the final colorization is evaluated using Mean-squared error and peak signal to noise ratio. Figure 3 shows the model loss of 18 epochs.

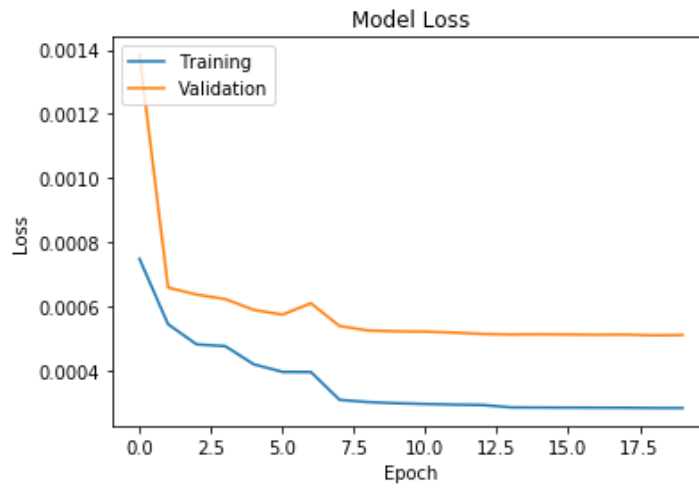













Figure 3. Model loss of training and validation.

The table 2 below shows the results of the automatic colorization for five images with different scenes.

Table 2. Results of automatic colorization using the proposed model.

Real image	GrayScale image	Output image	MSE	PSNR
			54.4	30.7
			51.8	30.9
			45.1	31.5

	61.9	30.2
	43.6	31.7

Conclusions

In this paper, we introduced an automatic image colorization based on deep learning approach. The using of the automatic colorization technique to add the color to black and white images will save the time and human effort because the using of user-based colorization is a time-consuming task when the images have more objects and complex scene. The development of this technique and using a huge amount of datasets and more complex images will improve the results of colorization and fill the gap between the real and output colors.

References

Reference to a journal publication:

- [1] Levin, A., D. Lischinski, and Y. Weiss. *Colorization using optimization*. in *ACM transactions on graphics (tog)*. 2004. ACM.
- [2] Yatziv, L. and G.J.I.t.o.i.p. Sapiro, *Fast image and video colorization using chrominance blending*. 2006. **15**(5): p. 1120-1129.
- [3] Horiuchi, T.J.I. and V. Computing, *Colorization algorithm using probabilistic relaxation*. 2004. **22**(3): p. 197-202.
- [4] Takahama, T., T. Horiuchi, and H. Kotera. *Improvement on colorization accuracy by partitioning algorithm in cielab color space*. in *Pacific-Rim Conference on Multimedia*. 2004. Springer.
- [5] LeCun, Y., Y.J.T.h.o.b.t. Bengio, and n. networks, *Convolutional networks for images, speech, and time series*. 1995. **3361**(10): p. 1995.

- [6] LeCun, Y.J.U.h.y.l.c.e.l., *LeNet-5, convolutional neural networks*. 2015. **20**: p. 5.
- [7] Hiley, L., A. Preece, and Y.J.a.p.a. Hicks, *Explainable Deep Learning for Video Recognition Tasks: A Framework & Recommendations*. 2019.
- [8] Hubel, D.H. and T.N.J.T.J.o.p. Wiesel, *Receptive fields and functional architecture of monkey striate cortex*. 1968. **195**(1): p. 215-243.
- [9] Fukushima, K.J.B.c., *Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position*. 1980. **36**(4): p. 193-202.
- [10] Matsugu, M., et al., *Subject independent facial expression recognition with robust face detection using a convolutional neural network*. 2003. **16**(5-6): p. 555-559.
- [11] Deng, J., et al. *Imagenet: A large-scale hierarchical image database*. in *2009 IEEE conference on computer vision and pattern recognition*. 2009. Ieee.