International Journal for Research in Engineering and Emerging Trends (IJREET), Volume 4, Issue 2, Dec, 2020
ISSN: 2545-4523 (Online)

# Black and White Image Color Restoration using Caffe Model

**Rishabh Kapoor[1], Bhavya Manocha[2], Abhishek Dobhal[3], Kunal Chandel[4], Princy Jain[5]**

[1,2,3,4] *Student, Dept. of information technology, ADGITM, Delhi, India*
[5] *Assistant professor, Dept. of information technology, ADGITM, Delhi, India*

## ABSTRACT

Image color restoration or colorization is the process of taking a black and white image as an input and then producing an output colorized image as a result that shows the semantic colors and tones of the input image, the goal is to make the output image as realistic as the input image. The authors of this paper have used OpenCV library, Caffe model, NumPy, Argparse to produce the output image, But it doesn't mean that the output image will match the ground reality every time. We accept the underlying problem as a classification task and have used class rebalancing at the time of training the image set to increase the range of colors in the resultant image.

***Keywords***: Open CV, Caffe model, Image restoration, CNN, Self-supervised learning.

## INTRODUCTION

Image colorization can add color to photographs that were originally taken in black and white. This can be used to provide a best-guess as to the context of the picture and help bridge the gap between the past and the present. The goal of our model is to produce realistic colorized photos with minimum desaturation and minimum human intervention. As a start one could build up over the semantics, the sky is typically blue, the grass is green. But these kinds of semantics cannot define the picture with utmost precision.

We start with a simple grayscale image as an input. We then use a neural network to output a predicted colorized image. Given the Lightness channel, a and b color channels are predicted by our system of the image which is being previously converted into a CIELAB color space.

Having freely available access to an abundance of colored images, we were in luck. The easy availability of such ImageNet and the previous works have trained CNN's to predict color on large datasets[1,2].

But the results that CNN produced tend to be desaturated and a little off, furthermore the loss function that was used in the CNN was derived by the conservative predictions where the losses were inherited from the standard regression problems where the main aim was to minimize the Euclidean error between the ground truth reality and the estimate.

We on the other hand have used a loss function more specific to the colorization problem.
As we can see, the model suggested by showed that the color predictions are multimodal, which means that many different objects can take up many different colorizations.
To appropriately model our multimodal nature of the problem that we are tackling we have to tailor our loss function accordingly. We came to an understanding that by being able to predict a distribution of possible colors of each pixel we can significantly enhance our model in predicting a more realistic image. By re-weighting the loss at training time we were able to emphasize the rare colors. That's where the availability of large datasets comes in handy. To further enhance our model to produce more realistic imagery, we produced the final colorization by taking the *annealed mean* of the produced distribution.
Although being able to create an image's-colored self, we cannot assure that it will match its ground reality instead the model is producing an image so realistic that the viewers could not spot the fake while looking at it the true image and the image that is produced by our model.
So to sum up, we created an appropriate objective function that used a more appropriate loss function that is capable of handling multimodal uncertainty and thus producing more realistic, non-distinguishable from real one's images thus setting a benchmark by being able to train a model over such a large dataset.

## RELATED WORKS

In this section, we are going to highlight the existing work that has already been done in this field. Before 2005 colorization was a painstaking task done by hand. The digital grayscale image would be touched up pixel by pixel. Some changes were made in this area by adding color leaking and color continuity. This allowed for batch work to be done, but this was still done manually. In 2005, a technique was described how human scribbles could suggest to the computer which colors to fill in. Although this did automate the process to some extent it still required heavy user (human) annotation.
Most of our research involved the use of convolutional neural networks to colorize images. Early models (mainly after 2005) use a simple mean squared error for the loss function. It leads

to desaturation of images as it stimulates the model to make safe predictions, which can lead to browning Newer models, such as Colorful Image Colorization, tailor their loss function to the colorization problem reweighting to more rare colors given that frequency of colors is not distributed evenly. This helps encourage bolder pixel choices rather than conservative ones.

Colorization algorithms are mostly different in ways such that they obtain and treat the data for modeling the correlation between grayscale and color. Non-parametric methods, given an input grayscale photo, first define one or more color reference images either given by a user or fetched automatically to be used as source data. Then, computing the Image Analogies framework, color is transferred onto the input image from different regions of the reference images. On the other hand, parametric methods, learn prediction functions from large datasets of color images at the time of training, posing the problem as a regression into continuous color space or classification of quantized color values. Our method also learns to classify colors, but it does so with a larger model, trained on more data, and with a lot of transformation in the loss function and mapping to a final output that is continuous.

## DATASET AND FEATURES

### IMAGENET

ImageNet is an image database organized according to the WordNet hierarchy, in which each node of the hierarchy is depicted by hundreds and thousands of images. Currently, it has an average of over five hundred images per node.

### CIELAB

Similar to the RGB color space, the Lab color space has three channels. But unlike the RGB color space, Lab encodes color information differently.

1.) The ‘L’ channel encodes lightness intensity.

2.) The ‘a’ channel encodes green-red.

3.) The ‘b’ channel encodes blue-yellow.

## METHODOLOGY

### 1.) Lab Colour Space

In previous methods, we used to code a color photo using the RGB model. The RGB color model is an additive color model in which red, green, and blue light are added together in various ways to reproduce a wide array of colors. But the model that will be used on this project is the "Lab" color space. The CIELAB color space represents color as three numerical values, L stands for the lightness and a and b stands for the green–red and blue-yellow color components. Unlike the RGB color model, Lab color is designed to approximate human vision. It aims for perceptual consistency, and its L component closely matches human perception of lightness. The L component is exactly what is used as input of the AI model, that was trained to estimate the remaining components, "a" and "b".

### 2.) The AI Process

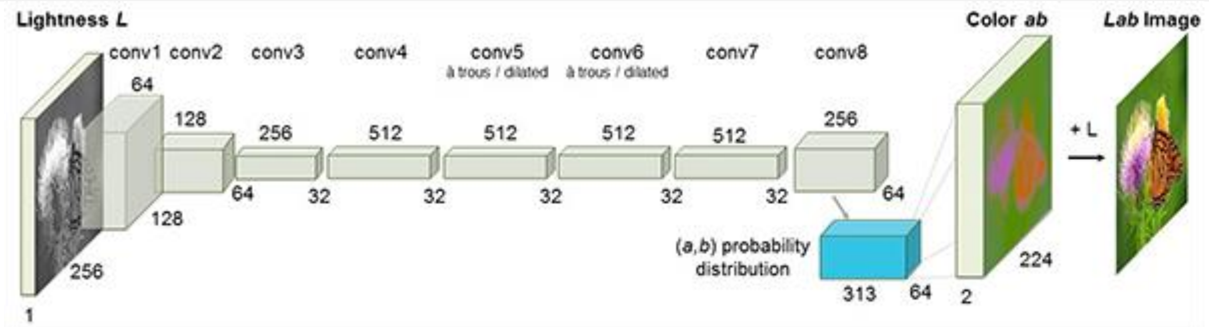
The Artificial Intelligence (AI) approach is implemented as a feed-forward pass in a CNN ("Convolutional Neural Network") at time of testing and dataset is trained on over a million coloured images. In other words Over a million coloured images were disintegrated using CIELAB model and lightness was used as an input feature (L) and 'a' and 'b' were used as a classification label. In short, using a broad and wide range of objects and scenes from a dataset of 1.3 Million photos from ImageNet and applying a Deep Learning algorithm (Feed-Forward CNN), final models for the image were generated.

### 3.) Neural network

A neural network is a series of algorithms that attempts to find existing relationships in a set of data through a process that mimics the way the human brain functions. Neural networks are adaptive and can adapt to changing input; so the network generates the best possible result without needing to redesign the output criteria.

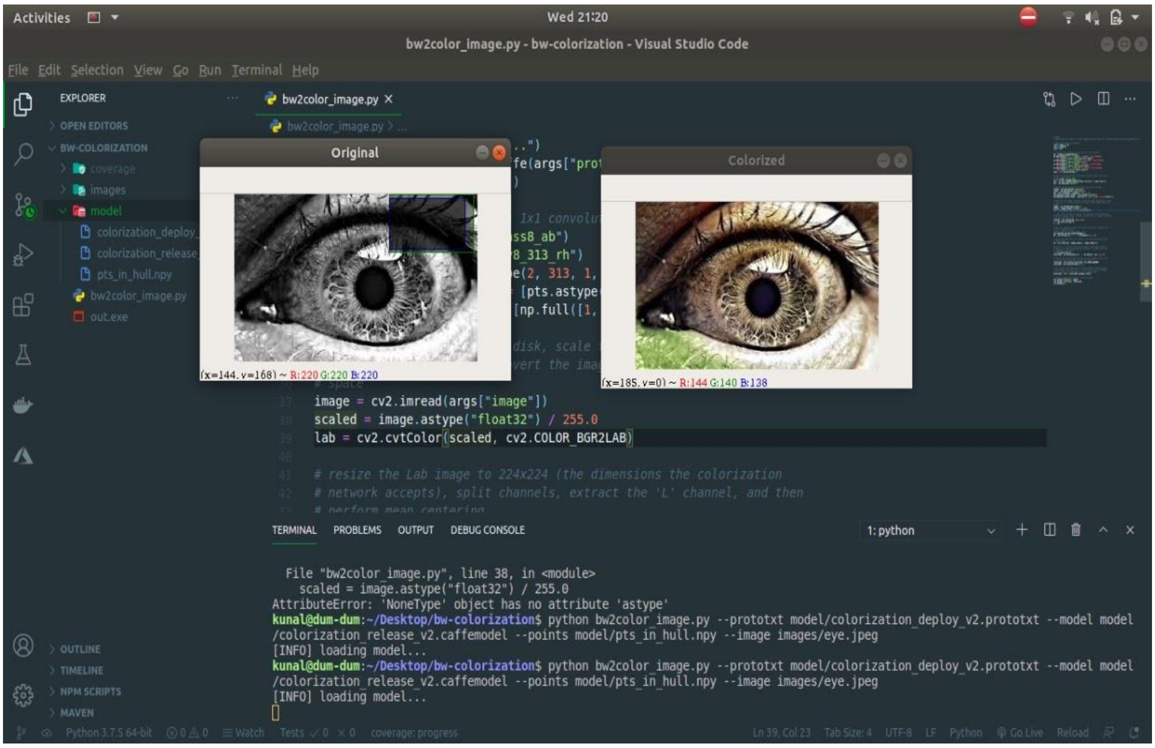
### 4.) Convolutional Neural Network

CNN's are regularized versions of multilayer perceptrons. Multilayer perceptron’s usually mean fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. The "fully-connectedness" of these networks makes them prone to overfitting data. Typical ways of regularization include adding some form of magnitude measurements of weights to the loss function.




## RESULTS/DISCUSSIONS

We received pretty astonishing results, we obtained an astounding 32% more accuracy than the previous models in delivering the ground true colors of the image, and also desaturation was comparatively less than it’s been in the older models. Moreover, we have shown that colorization can be a powerful task of self-supervised feature learning, acting as a *cross-channel encoder*. This approach can be used in several feature learning benchmarks. We also believe that our work can be improved with larger data sets of the images and more training time which can result in more accurate colors altogether.




## CONCLUSION

We created an appropriate objective function that used a more accurate loss function that is capable of handling multimodal uncertainty and thus producing more realistic, non-distinguishable from real one's images thus setting a benchmark by being able to train a model over such a large dataset.

Also, the Lab color space seems like the best way to work with grayscale and color images. The series of convolutional neural networks is a good simple model to colorize photos and adding a parallel classifier to learn more about these grayscale photos to help in the colorization process makes intuitive sense.

## FUTURE SCOPE

Due to time constraints, we weren’t able to expand the scope of our model. We could have improved our model, perusal with more time and more compute.

Furthermore, the concept of image restoration and colorization can be extended to the scope of black and white video colorization, and those might be very helpful in the restoration of many older documentaries as a colorized frame would provide us with more details about the events that happened in the past.

## REFERENCES

[1]. Cheng, Z., Yang, Q., Sheng , B.: Deep colorization. In: Proceedings of the IEEE International Conference on Computer Vision. (2015) 415–423

[2]. Dahl, R.: Automatic colorization. In: http://tinyclouds.org/colorize/. (2016)

[3]. Charpiat, G., Hofmann, M., Scholkopf, B.: Automatic image colorization via multimodal predictions. In: Computer Vision–ECCV 2008. Springer (2008) 126–139 4. Ramanarayanan, G., Ferwerda, J., Walter, B., Bala, K.: Visual equivalence: towards a new standard for image fidelity. ACM Transactions on Graphics (TOG) 26(3) (2007) 76

[4].https://github.com/jantic/DeOldify

[5].http://richzhang.github.io/colorization

[6]. Wallner, Emil https://www.freecodecamp.org/news/colorize-b-w-photos-with-a-100-line-neural-network-53d9b4449f8d/

[7].https://en.wikipedia.org/wiki/Convolutional_neural_network

[8].https://caffe.berkeleyvision.org/

[9]. Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A.: The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results. http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html

[10].Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (VOC) challenge. International journal of computer vision 88(2) (2010) 303–338

[11]. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM international conference on Multimedia, ACM (2014) 675–678 36.

[12].Baldassarre, Federico, lonzález Morin, and Lucas Rodes-Guirao. "Deep Koalarization: Image Colorization using CNN. and Inception ResNet-v2. arXiv preprint arXiv 1712.03400 (2017)