Image Colourization (using Machine Learning)

Akash
*Student, **Department of Computer Science and Engineering, Raj Kumar Goel Institute of Technology**, Ghaziabad, Uttar Pradesh 201003, India*

Amit Kumar Prajapati
*Student, **Department of Computer Science and Engineering, Raj Kumar Goel Institute of Technology**, Ghaziabad, Uttar Pradesh 201003, India*

Manish Kumar Pandey
*Student, **Department of Computer Science and Engineering, Raj Kumar Goel Institute of Technology**, Ghaziabad, Uttar Pradesh 201003, India*

Project Guide: Ms. Rinki Tyagi
*Assistant Professor, **Department of Computer Science and Engineering, Raj Kumar Goel Institute of Technology**, Ghaziabad, Uttar Pradesh 201003, India*

ABSTRACT

Colorization of grayscale images has become a more studied area in recent years, due to the advent of Deep convolutional neural networks. We are going to apply this concept to colouring black & white images through deep learning. While Previous similar studies have focused primarily on natural image colouring, cartoons colouring has traditionally been done by using hand drawing Techniques. The proposed method is a Self-Supervised Learning or Fully Automated Approach.

The main goal of the project report is to provide an overview of how to convert a grayscale image to a colourful image using colorization problems. achieving artifact-free quality usually requires manual matching which is considered as a very difficult problem. This process usually requires careful selection of colourful suggestive images.

Given a grayscale photo as input, this article addresses the problem of hallucinating a photorealistic colour version of the photo. This issue is obviously not limited, so previous approaches have either relied on significant user interaction or resulted in desaturated results. We offer fully automatic approach to providing vivid and realistic colorizations. We accept the underlying uncertainty of the problem by posing it as a classification task and use class rebalancing during training time to increase the diversity of colours in our result. The system is implemented as a feed-forward pass on a CNN during test time and is trained on over a million colour images. Evaluate the algorithm with a "Turing test", by asking participants to choose between the generated colour image and the original colour image.

Keyword: - *CNNs, Turing Test, Colourization , Automated, Rebalancing, Artificial Intelligence, Machine Learning, Deep Learning.*

---

1. INTRODUCTION

Image colouring is the process of assigning colour to a grayscale image to make it more aesthetic and meaningful. This is known to be a complex task that often requires prior knowledge about the image content and manual adjustments to achieve artifact free quality. Also, objects can have different colours, there are many possible ways to assign colours to pixels in an image, which means there is no unique solution to this problem. There are two main approaches for colorization of image: one is let user assign colours to some regions and extends such information to the whole image, and another one that tries to learn the colour of each pixel from a colour image with similar content. In article, we use the latter approach; we extract the information about colour from an image and transfer it to another image. Recently, deep learning has gained increasing attention among researchers in the field of computer vision and image processing. As a typical technique, convolutional neural networks (CNNs) have been well-studied and successfully applied to several tasks such as image recognition, image reconstruction, image generation, etc. A CNN consists of multiple layers of small computational units that only process portions of the input image in a feed-forward fashion. Each layer is the result of applying various image filters, each of which extracts a certain feature of the input image, to the previous layer. Thus, each layer may contain useful information about the input image at different levels of abstraction. With the evolution of computational resources, especially the computing power of GPUs, it has become possible to train very deep CNNs, and they have achieved some remarkable results recently. For example, a deep CNN (He et al., 2015)[3] has surpassed human-level performance on ImageNet classification, or an adversarial network (Radford et al., 2015), in which two CNNs are trained simultaneously, is capable of generating plausible-looking images of many kinds of objects.

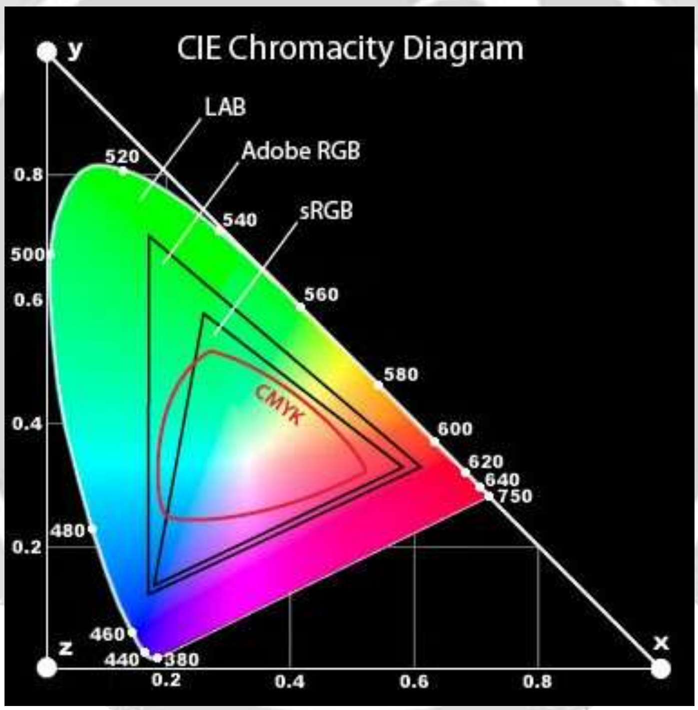

Figure -1: CIE Chromatic Diagram

2. LITERATURE REVIEW

Though convolutional neural networks (often shortened to ConvNets, CNNs, dCNNs) was not entirely new technology, it had taken the world by storm since [6]. CNNs now developed which are utilized in different digital applications were actually first observed in living organisms. In the 1950s & 1960s, Hubel & Wiesel have worked on cat and monkey which showed that their visual cortexes contained neurons that separately reciprocate to small areas of the visual field. The visual stimuli in the area of visual space affects the firing of a single neuron provided their eyes are not moving. This is its receptive field.

Likely and intercepting receptive regions have been observed in the neighbouring cells. Each hemisphere in the cortex represents the contralateral visual field. This lead to the introduction of neocognitron, delay in the time of neural networks and trainable weights. All these formed basics to the first ever documented commercial use of CNNs which dates back to 1998, with LeNet-5. LeNet-5 was designed by LeCun et al [7]. Years of researching CNNs, made it possible to recognise text character based on 32x32 pixel images.

Model consists of four main components: a low-level features network, a mid-level features network, a global features network, and a colorization network. Conceptually, these networks function as follows: First, a common set of shared low-level features are extracted from the image. Using these features, a set of global image features and mid-level image features are computed. Then, the midlevel and the global features are both fused by our proposed "fusion layer" and used as the input to a colorization network that outputs.

3. METHODOLOGY

We are working in Lab color space so the L channel has a grayscale information and serves as input to our system we are looking to predict the ab channel or the color information & we learn the mapping from L to ab using the CNN. We can take the predicted ab channel concatenate them with input and hopefully get a plausible colorization of the input grayscale image channel encodes lightness intensity only, a channel encodes green-red, b channel encodes blue-yellow. To produce more plausible black and white image colorization we also utilized a few additional techniques including mean annealing and a specialized loss function for color rebalancing. The image colorization software made, is an automation System/ software which colorizes the images based on the Image Colorization Deep Learning Algorithm.

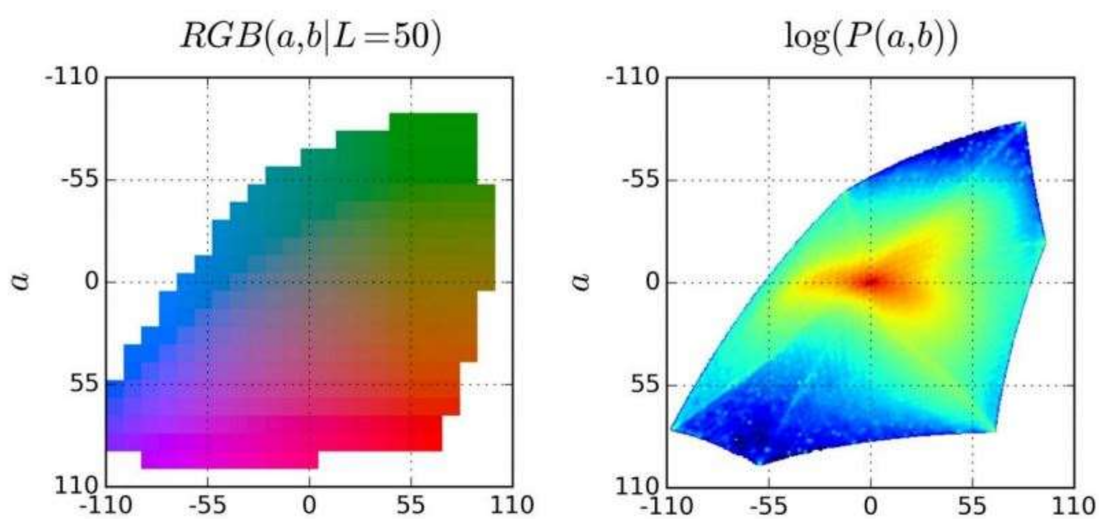

Figure 2: a*b* Space and Rebalancing

We train a CNN to map from a grayscale input to a distribution over quantized colour value outputs using the architecture shown in Figure 2. Architectural details are described in the supplementary materials on our project webpage1, and the model is publicly available. In the following, we focus on the design of the objective function, and our technique for inferring point estimates of colour from the predicted colour distribution.

our system compared to previous work well many previous techniques have focused on using a non-parametric framework so in the nonparametric framework a reference images first obtained and the colours from the reference image are then transferred over to the grayscale input image now this technique can actually work very well but it can also fail to generalize and also getting the reference image itself can be slow or requires some user intervention now parametric techniques some previous parametric techniques have also used l2 aggression[1] both with hand engineered features as well as deep networks we're also not the first of course to cast the colorization problem as multinomial classification.

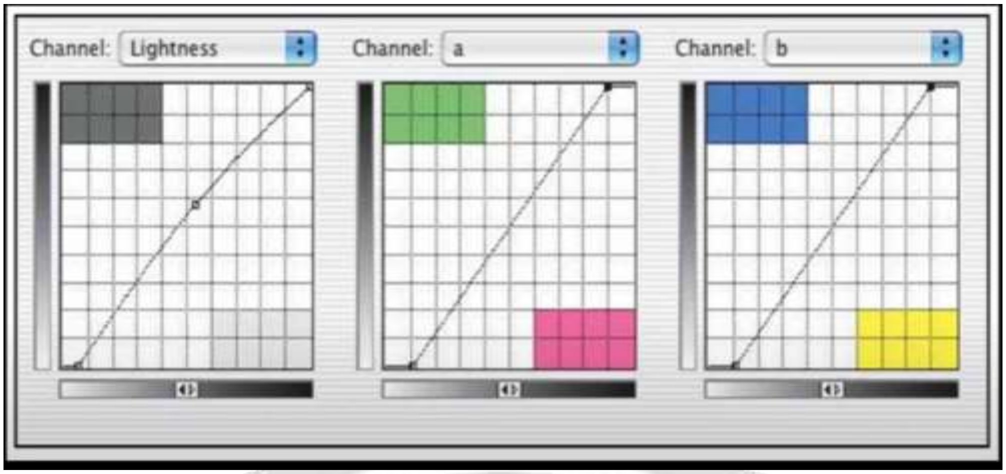

Figure- 3: L*a*b* Space

4. EXPERIMENTAL RESULTS

Figure 4: Output

4.1 Strength

We propose a fully automatic approach that produces vibrant and realistic colorizations. We embrace the underlying uncertainty of the problem by posing it as a classification task and use class-rebalancing at training time to increase the diversity of colours in the result. The system is implemented as a feed-forward pass in a CNN at test time and is trained on over a million colour images. This approach results in state-of-the-art performance on several feature learning benchmarks. This problem is clearly under constrained, so previous approaches have either relied on significant user interaction or resulted in desaturated colorizations.

4.2 Drawback

Like many systems built on Convolutional Neural Networks, Colourful Image Colorization produces some remarkable results, but it struggles with edge cases. For images with different compositions, novel objects, and unconventional colours, though, the algorithm struggles. If your image looks pretty good after colorization by the algorithm, it's probably a fairly "average" image in terms of composition and colour — it doesn't deviate much from the millions of images on which the system was trained. Despite its power, Colourful Image Colorization does have blind spots. While the colours it hallucinates are plausible, they're not always historically accurate — the Pan Am logo [3] in the image at the top of this article looks great in red, for example, but it was actually blue.

Lot of research is going on in this field to obtain colorized images that are as realistic as possible colour of objects such as scars clothes etc can be different from the ground truth but can still be considered accurate however the challenge lies in accurately colorizing standard components like skin Tone eyes hair nature and sky our goal with this project is to produce colorized images from grey scale ones that seem natural to the human eye in our implementation we propose a method for image colorization we pre-process coloured images to create grayscale images which are used as an input for our model we then train our model with these grayscale images and as input and the original coloured images as output upon training the model we would be able To generate rgb images for with a reasonable understanding of colour within its inherit texture we have used three datasets for training and validation of our model we first used the dataset to train our baseline model he model weights were initialized with pre-drain weights from bgg16 trained on the ImageNet Dataset we later used 50000 images from the sepha dataset to train our final model as it contained images of smaller size once we obtained decent results from this we decided to train our model using 20000 images from the ImageNet dataset for implementing this we tried two different approaches we first used a transfer learning approach as a baseline that uses features obtained from the Image after providing it to the pre-trained model to make predictions using a cnn[2] we later used an order encoder that initially through a series of cnns and down samplings learns a reduced dimensional representation of the data and then using cnn's and up samplings is able to regenerate coloured images.

5. CONCLUSION

Colorization Results on ImageNet							
Method	Model Params (MB)	Model Feats (MB)	Model Runtime (ms)	AuC non-rebal (%)	AuC rebal (%)	VGG Top-1 Class Acc (%)	AMT Labeled Real (%)
Ground Truth	–	–	–	100	100	68.3	50
Gray	–	–	–	89.1	58.0	52.7	–
Random	–	–	–	84.2	57.3	41.0	13.0±4.4
Dahl [2]	–	–	–	90.4	58.9	48.7	18.3±2.8
Larsson et al. [23]	588	495	122.1	**91.7**	65.9	**59.4**	**27.2±2.7**
Ours (L2)	129	127	17.8	91.2	64.4	54.9	21.2±2.5
Ours (L2, ft)	129	127	17.8	91.5	66.2	56.5	23.9±2.8
Ours (class)	129	142	22.1	91.6	65.1	56.6	25.2±2.7
Ours (full)	129	142	22.1	89.5	**67.3**	56.0	**32.3±2.2**

Figure 5: Colorization results on 10k images in the ImageNet validation set [4], as used in [5]

While image colorization is a boutique computer graphics task, it is also an instance of a difficult pixel prediction problem in computer vision. Here we have shown that colorization with a deep CNN and a well-chosen objective function can come closer to producing results indistinguishable from real colour photos. Our method not only provides a useful graphics output, but can also be viewed as a pretext task for representation learning. Although only trained to colour, our network learns a representation that is surprisingly useful for object classification, detection, and segmentation, performing strongly compared to other self-supervised pre-training methods.

6. ACKNOWLEDGMENTS

This work was supported by Raj Kumar Goel Institute of Technology (RKGIT) with Basic Technology for Extracting High-level Information from Multiple Sources Data base on Intelligent Analysis.

7. REFERENCES

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

[2] Kirkpatrick, S., Vecchi, M.P., et al.: Optimization by simmulated annealing. science 220(4598) (1983) 671–680

[3] Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)

[4] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., et al.: Imagenet large scale visual recognition challenge. International Journal of Computer Vision 115(3) (2015) 211–252

[5] Larsson, G., Maire, M., Shakhnarovich, G.: Learning representations for automatic colorization. European Conference on Computer Vision (2016).

[6] Owens, A., Isola, P., McDermott, J., Torralba, A., Adelson, E.H., Freeman, W.T.: Visually indicated sounds. CVPR (2016)

[7] Welsh, T., Ashikhmin, M., Mueller, K.: Transferring color to greyscale images. ACM Transactions on Graphics (TOG) 21(3) (2002) 277–280