IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

COLORIZATION OF BLACK AND WHITE IMAGES: A SURVEY

Angelica Deigratia Pyngrope
PG Student,
Department of Master of Computer Applications
School of CS & IT
JAIN (Deemed-to-be University)

Pawan Kumar
Assistant Professor,
Department of Master of Computer Applications
School of CS & IT
JAIN (Deemed-to-be University)

Abstract: One of the most exciting Deep Learning applications is color grading black and white pictures. Coloring a grayscale image is a simple exercise for the human mind in general; we learn to fill in missing shades in colouring books from a young age. This task used to necessitate a lot of human involvement and hardcoding, but now, thanks to AI and Deep Learning, the entire process can be automated from start to finish. In this review, I looked at a number of publications that presented various AI and Deep Learning methodologies. I'll go over each approach and tactic that has been utilised to investigate different Neural Networks in Deep Learning and how they may be used to achieve incredible outcomes.

***Index Terms* - Colorizing black and white images, Deep Learning, Images, AI.**

I. Introduction

The act of taking an input grayscale (black and white) image and turning it into an output colourized image that represents the input's conceptual colours and tones. For example, an object in the image with the colour yellow must be identified; otherwise, the model will colour it blue.

Deep learning (sometimes called deep structured learning) is a machine learning technique that uses artificial neural networks and representation learning. Learning can occur in supervised, semi-supervised, or unsupervised conditions.

When photography was first established, only black and white photos were available due to technological restrictions. Color photography, on the other hand, is now commonplace. There are numerous recollections and links between the present and the past when it comes to historical photography. Converting them to coloured versions would be more fascinating in terms of enhancing hidden meanings and making them more visually appealing. Manual or Photoshop colorization was used, which took a long time. In recent years, many Deep Learning-based colorization techniques have been proposed. When a colour image is transformed to a grayscale version, information is lost across dimensions, which causes the colorization problem. Some solutions took a classification approach to the problem, while others took a regression method. In order to evaluate these approaches, I present a review of what the authors have done, then incorporate a whole generator model and tweak their training strategy.

II. Related Work

In recent years, a variety of deep learning-based colorization techniques have been developed. The Colorful Image Colorization paper took a classification approach to the problem, and they took into account the problem's uncertainty (for example, a vehicle in the image can take on many distinct and genuine colours, and we can't be certain about any of them); however, another paper took a regression approach to the problem, with a few system modifications. Each approach has advantages and disadvantages, and I will describe how each differs from the other in terms of the strategies and techniques used and the accuracy of their model.

Richard Zhang et al., [1] approached the problem by using a feed forward pass in a CNN and has stated it as a multinomial classification. He utilized the ImageNet dataset and extracted only the first few 10,000 images for training and testing. A "colorization Turing test" was being used to evaluate the algorithm, which required human volunteers. To signify predictions using a natural objective function, the objective function is optimised using the CIE Lab colour space model in the Euclidean Loss L2 between predicted and ground truth colors.

Loss function:

$$L_{cl}(\widehat{\mathbf{Z}}, \mathbf{Z}) = -\sum_{h,w} v(\mathbf{Z}_{h,w}) \sum_{q} \mathbf{Z}_{h,w,q} \log(\widehat{\mathbf{Z}}_{h,w,q})$$

A mapping from a given input to a probability distribution on all feasible colours, where Q is the number of quantized ab values. A function that transforms ground truth Y to vector Z using a gentle encoding approach is defined to compare predictions against ground truth. Finally, a multinomial cross entropy Loss L_{cl} is defined as:

$$L_{cl}(\widehat{\mathbf{Z}}, \mathbf{Z}) = -\sum_{h,w} v(\mathbf{Z}_{h,w}) \sum_{q} \mathbf{Z}_{h,w,q} \log(\widehat{\mathbf{Z}}_{h,w,q})$$

where *v (·)* is a re-balancing concept based on color-class scarcity that can be used to re-balance the loss.

Phillip Isola, et al., [2] Phillip applied conditional adversarial networks to image-to-image translation challenges as a general-purpose solution. Their system comprises of a few other techniques such as synthesizing photos from label maps, reconstructing objects from edge maps, and colorizing images, among other tasks. He mentions some of the drawbacks of utilising the CNN method for colorization, such as the need to minimise the Euclidean distance in[1], between predicted and ground truth colors, The issue is that the Euclidean distance can only be decreased if all of the outputs have an average value, which will further blur the image. As a result, he chose a conditional GAN that will learn a conditional generative model, distinguishing it from ordinary GAN models. Various researchers have already experimented with GANs in the past., [3] illustrates the usage of the model to learn a multi-modal model as well as generation of image tags. However, in [2], the usage of the GAN network is solely for the purpose of image translations. The system architecture includes a generator-discriminator in which both have different roles such as, a U-net architecture is used for the generator and as for the discriminator, a PatchGAN classifier is used which will penalize structure at the scale of image patches. The modules used by the generator-discriminator is of the form convolution-BatchNorm-ReLu. Another important aspect is the level of information between the input and output, resulting in a bottleneck as well. The addition of a skip connection to the U-Net architecture in the generator has been proven to tackle this. The configuration of a GAN neural network can be illustrated through the figure below which will describe how the paper utilized the algorithm:

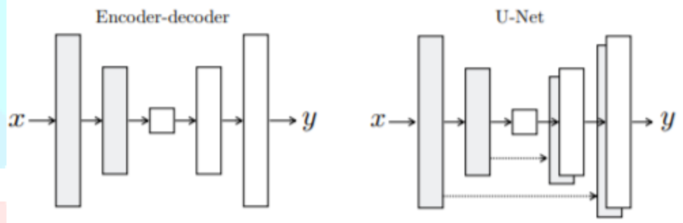

Fig 1.1 two choices for the architecture of the generator. U-net architecture [2].

In Fig. 1.1, the GAN learns a mapping from observed image x and random noise vector *z*, to *y*, *G: {x, z}* → *y*. During training and test time, noise is only delivered in the form of dropouts, which is distributed to numerous layers of the generator. Considering the dropout noise, the output of the nets has relatively moderate randomness.

The goal of Madhab et a.,[4] colorization was to protect Nepal's historical culture by retaining its uniqueness. In addition, the study recommended employing a CNN in conjunction with an Inception-RestnetV2 and the RGB colour model for pattern recognition using the backpropagation method. To extract low-level information from the input image, the network also employs an encoder-decoder architecture. A self-generated dataset of 1,200 ancient and historical photographs of Nepal with a resolution of 256x256 pixels was employed. The MSE (Mean Squared Error) and the PSNR (Percentage of Squared Error) are the two loss functions that have been applied (Peak-Signal-to-Noise-Ratio). The model's validation is combined with a subjective value as the MOS to assess colorization accuracy (Mean Opinion Score).

Y. Morimoto, et al., [5] proposed an automatic colorization system based on information from a scene structure that used one million photos. A gist scene descriptor, which is a feature vector to characterise the global scene in lower dimension, was utilised to discover the source image. In [6], through examples, the approach learns to colourize. A LEARCH framework is utilised to train a quadratic objective function that is analogous to a Gaussian random field in the chromaticity maps. The goal function coefficients are trained on image features using a random forest.

Convolutional neural network approaches have recently been applied to the problem of colorization by a number of academics. As noted in the study, the CNN(Convolutional Neural Networks) has been shown to be the most commonly utilised algorithm There are various aspects to their proposed methodology that must be examined:

- The network is trained using a multinomial cross entropy loss function weighted by colour rarity.
- The approach works well on a variety of photos and captures the multimodal aspect of pixel colour; nonetheless, colour bleeding defects occur occasionally, as they do with many other colorization algorithms.
- Image semantics are well captured by convolutional neural networks. The weights for each cost element, however, are dependent on the input photographs, making the model difficult to generalise to a large number of images. Even though the computation is done on GPU, it is difficult to achieve a real-time colorization experience when pixels are generated with reasonable based optimization.

III. Conclusion

Image colorization is a specialised computer graphics activity, but it is also an example of a tough computer vision pixel prediction problem. When we look at several research publications, we notice that the diverse methodologies and strategies used can give results that are indistinguishable from true colour images. Various methodologies, as well as a well-chosen loss function, significantly improve the results. I'd want to express my gratitude to the authors of the excellent works for their contributions.

References

[1] R. Zhang *et al.*, “Real-time user-guided image colorization with learned deep priors,” *ACM Trans. Graph.*, vol. 36, no. 4, pp. 1–11, Jul. 2017, doi: 10.1145/3072959.3073703.

[2] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-Image Translation with Conditional Adversarial Networks,” *arXiv:1611.07004 [cs]*, Nov. 2018, Accessed: Mar. 05, 2022. [Online]. Available: http://arxiv.org/abs/1611.07004

[3] M. Mirza and S. Osindero, “Conditional Generative Adversarial Nets,” *arXiv:1411.1784 [cs, stat]*, Nov. 2014, Accessed: Mar. 05, 2022. [Online]. Available: http://arxiv.org/abs/1411.1784

[4] M. R. Joshi, L. Nkenyereye, G. P. Joshi, S. M. R. Islam, M. Abdullah-Al-Wadud, and S. Shrestha, “Auto-Colorization of Historical Images Using Deep Convolutional Neural Networks,” *Mathematics*, vol. 8, no. 12, p. 2258, Dec. 2020, doi: 10.3390/math8122258.

[5] Y. Morimoto, Y. Taguchi, and T. Naemura, “Automatic colorization of grayscale images using multiple images on the web,” in *SIGGRAPH 2009: Talks on - SIGGRAPH ’09*, New Orleans, Louisiana, 2009, pp. 1–1. doi: 10.1145/1597990.1598049.

[6] A. Deshpande, J. Rock, and D. Forsyth, “Learning Large-Scale Automatic Image Colorization,” in *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, Dec. 2015, pp. 567–575. doi: 10.1109/ICCV.2015.72.

[7] F. Pierre and J.-F. Aujol, “Recent Approaches for Image Colorization,” in *Handbook of Mathematical Models and Algorithms in Computer Vision and Imaging*, K. Chen, C.-B. Schönlieb, X.-C. Tai, and L. Younces, Eds. Cham: Springer International Publishing, 2021, pp. 1–38. doi: 10.1007/978-3-030-03009-4_55-1.

[8] E. H. Land and J. J. McCann, “Lightness and Retinex Theory,” *J. Opt. Soc. Am.*, vol. 61, no. 1, p. 1, Jan. 1971, doi: 10.1364/JOSA.61.000001.

[9] S. Iizuka, E. Simo-Serra, and H. Ishikawa, “Let there be color!: joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification,” *ACM Trans. Graph.*, vol. 35, no. 4, Art. no. 4, Jul. 2016, doi: 10.1145/2897824.2925974.

[10] G. Larsson, M. Maire, and G. Shakhnarovich, “Learning Representations for Automatic Colorization,” *arXiv:1603.06668 [cs]*, Aug. 2017, Accessed: Mar. 05, 2022. [Online]. Available: http://arxiv.org/abs/1603.06668