A Survey On Auto-Image Colorization Using Deep Learning Techniques With User Proposition

C.Santhanakrishnan¹, Neeraj Durgapal², Deepak Yadav³

¹,²,³Department of Computer Science and Engineering SRM Institute of Science and Technology
csksrm@gmail.com¹, durgpalneeraj@gmail.com², dydeepak97@gmail.com³

Abstract—An approach based on deep learning for automatic colorization of image with optional user-guided hints. The system maps a gray-scale image, along with, user hints” (selected colors) to an output colorization with a Convolution Neural Network (CNN). Previous approaches have relied heavily on user input which results in non-real-time desaturated outputs. The network takes user edits by fusing low-level information of source with high-level information, learned from large-scale data. Some networks are trained on a large data set to eliminate this dependency. The image colorization systems find their applications in astronomical photography, CCTV footage, electron microscopy, etc. The various approaches combine color data from large data sets and user inputs provide a model for accurate and efficient colorization of grey-scale images.

Keywords—image colorization; deep learning; convolutional neural network; image processing

1. INTRODUCTION

There are two broad approaches followed in computer graphics to image colorization: user-driven edit propagation and automatic colorization driven by data. In the first approach, a user provides hints by colored strokes over a gray-scale image. The enhancement strategy at that point produces a colorized picture it coordinates the client’s strokes, while additionally sticking to hand characterized picture priors, for example, piece-wise smoothness. These strategies can be utilized to yield extremely noteworthy outcomes however they regularly require serious client collaboration (once in a while in excess of fifty strokes), as each diversely hued picture area must be unequivocally shown by the client. Since the framework depends intensely on client contributions for hues, even locales with little shading vulnerability, for example, blue sky or green vegetation, should be determined.

An additional information driven colorization strategies is likewise attempted by the scientists to address these constraints. These techniques colorize a dim scale picture in one of two different ways: either by coordinating it with a model hued picture in a database and non-parametric duplicating hues from that photograph, or by taking in parametric mappings from dark scale to shading from substantial scale picture information.

The latest techniques in information driven worldview proposed, utilize profound neural systems and they are completely programmed. The aftereffects of the procedure more often than not contain incorrect hues and ancient rarities alongside the shaded yield. The precise shade of couple of nonexclusive articles, for example, a shirt, is regularly vague it’s shading could be orange, red, or pink.

The new methodology is to attempting to consolidate both of these strategies to outdo both, utilizing substantial scale information to learn priors about regular shading symbolism, while in the meantime joining client decisions. The thought is to prepare a CNN on an extensive informational collection to straightforwardly outline scale pictures, alongside client con- tributions, to create a colorized yield picture. Amid preparing, haphazardly recreating client inputs empowers us to counter the issue of
gathering an exceptionally substantial number of client cooperations.

2. LITERATURE SURVEY

Colorization basically involves assigning realistic colors to grey-scale image. Convolutional neural networks are specifically designed to deal with image data. Many authors have done promising work on this idea.

Domonkos Varga [1] proposed the idea of automatic coloring of cartoon images, since they are very different from natural images, they pose a difficulty as their colors depend on artist to artist. So, the data-set was specifically trained for cartoon images, about 100000 images, 70% of which were used in training and rest for validation. But unfortunately, the color uncertainty in cartoons is much higher than in natural images and evaluation is subjective and slow.

Shweta Salve [2] proposed another similar approach, employing the use of Google’s image classifier, Inception ResNet V2. The system model is divided into 4 parts, Encoder, Feature extractor, Fusion layer and Decoder. The system is able to produce acceptable outputs, given enough resources, CPU, Memory, and large data-set. This is mainly proof of concept implementation.

Yu Chen [3] proposed a approach to mainly address the problem of coloring Chinese films from past time. They used existing data-set with their data-set of Chinese images, fine-tuning the overall model. The network makes use of multi-scale convolution kernels, combining low and middle features extracted from VGG-16.

V.K. Putri [4] proposed a method to convert plain sketches into colorful images. It uses sketch inversion model and color prediction in CIELab color space. This approach is able to handle hand-drawn sketches including various geometric transformations. The limitation found was that, data-set is very limited but it works well for uncontrolled conditions. Richard Zhang [5] has proposed a optimized solution by using huge data-set and single feed-forward pass in CNN. Their main focus lies

![Fig. 1: Left: Source and Right: Fully-automatic predicted output image](image)

on training part. They used human subjects to test the results and were able to fool 32% of them, can have various number of neurons. The various attempts used various architectures. In some papers, generally number of neurons is same as the dimension of the feature descriptor extracted from each pixel coordinates in a gray-scale image.
2) *Global features*: Most of the methods utilize the global features to form an image filter, and then use this filter to select similar images from a large image set automatically. However, some models produced unnatural colorization result due to global similarity but semantic difference.

3) *Data Set size*: Parametric and Non-Parametric models use different sizes of data sets for training the CNN. The Parametric model use very large data set to train the CNN and produce more accurate result while the Non-Parametric models rely more on the input hints and reference image and use smaller data set for training the CNN.

4) *Semantic Information*: Semantic information has a significant and unavoidable role in deep image colorization. For effective colorization of images, the system must have information of the semantic composition of the image and its localization. For instance, leaves on a tree may be colored some kind of green in spring, but they should be colored brown for a scene set in autumn. VGG-16 CNN model was used by majority of approaches to extract semantic information about the image before applying colorization techniques.

5) *Feature Extraction*: Features of the image are obtained through by integrating pre-trained neural networks to extract information about objects, shapes and use this context to assign color values to the objects. Some approaches used using Inception ResNet V2 classifier or Tensorflow to serve this purpose.

**Fig. 2: Left: Source with user hint and Right: Predicted output image**

**A. Surveyed Techniques**

The methods for image colorization can be categorized into two major groups: Based on user inputs and automatic colorization based. The methods make use of CNN for the colorization. The non-parametric methods first define one or more color reference images using the input from either user or a source image as source data. Then, color is transferred onto the input image from matching regions of the reference data. On the other hand, Parametric methods learn from training on large data sets of colored images, using either regression onto continuous color space or classification of quantized color values.

**B. Survey details**
1) Architecture: The deep neural networks consists of an input layer, many hidden layers and an output layer. Each layer

3. CONCLUSION

Based on the methodologies studies in the survey, approaches involving optional user hints can be used improve the CNN prediction ability to achieve realistic and optimal results. Approach should include training of CNN on randomized data which stimulates user input to cover common human errors. An extra neural layer to take input of user hints and combine it with trained CNN. The training data set need to be very large to eliminate the dependency of the model on user inputs for colorization. To create more accurate and efficient model, large-scale data can be leveraged to learn predictions about natural color imagery, and also incorporate user choices at the same time. Randomly simulating user inputs during the training enables us to counter the problem of collecting a very large number of user interactions and the error in user input can be predicted while relying less on user inputs and more on the large-scale data from the training set.

4. REFERENCES

[1] C. A. S. Domonkos Varga and T. Szirff, "Automatic cartoon colorization based on convolutional neural network," https://core.ac.uk/download/pdf/94310076.pdf, 2017.

[2] S. Salve, T. Shah, V. Ranjane, and S. Sadhukhan, "Automatization of coloring grayscale images using convolutional neural network," Apr. 2018. DOI: 10.1109/ICICCT.2018.8473259.

[3] Automatic colorization of images from chinese black and white films based on cnn, 2018. DOI: 10.1109/ICALIP.2018.8455654.

[4] V. K. Putri and M. I. Fanany, “Sketch plus colorization deep convolutional neural networks for photos generation from sketches,” in 2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), Sep. 2017, pp. 1–6. DOI: 10.1109/EECSI.2017.8239116.

[5] R. Zhang, P. Isola, and A. A. Efros, “Colorful image colorization,” in ECCV, 2016.

[6] J. V. Matias Richart and J. Baliosian, Image colorization with neural networks, https://www.fing.edu.uy/sites/default/files/biblio/31582/wvc_paper.pdf.

[7] Z. Cheng, Q. Yang, and B. Sheng, “Deep colorization,” Apr. 2016.

[8] J. Hwang, “Image colorization with deep convolutional neural networks,” 2016.

[9] V. K. Putri, Image colorization with neural networks, https://www.fing.edu.uy/sites/default/files/biblio/31582/ wvc_paper.pdf.

[10] NHK K. ISMAIL*, "Estimation Of Reliability Of D Flip-Flops Using Mc Analysis”, Journal of VLSI Circuits And Systems 1 (01), 10-12,2019.

[11] Sulyukova,"Analysis of Low power and reliable XOR-XNOR circuit for high Speed Applications”,Journal of VLSI Circuits And Systems 1 (01), 23-26,2019.