Auto Colorization of Photographic Pictures

Image Colorization using Deep Learning

Tridib Chakraborty[1], Sudeep Ghosh[2], MD Mizan Chowdhury[3], Suparna Das[4], Debayan Roy[5] , Sriparna Ghosh[6] ,Gourav Nandy[7], Sayani Sengupta[8] , Roumita Singha[9] ,Pallab Nath[10]

[1,2,3] *Assistant Professor,* [4] *Technical Assistant,* [5,6,7,8,9,10] *Student,* Dept. of Information Technology
Guru Nanak Institute of Technology
Panihati, West Bengal, India
Tridib.chakraborty@gnit.ac.in, Sudeep.ghosh@gnitac.in, Chowdhury.mizan@gnit.ac.in,
dassuparna0031@gmail.com,roydebayan9875@gmail.com

***Abstract*—We present a convolutional-neural-network-based system that faithfully colorizes black and white photographic images without direct human assistance. We explore various network architectures, objectives, color spaces and problem formulations. The final classification-based model we build generates colorized images that are significantly more aesthetically-pleasing than those created by the baseline regression-based model, demonstrating the viability of our methodology and revealing promising avenues for future work. Automated colorization of black and white images has been subject to extensive research through computer vision and machine learning techniques. Our research addresses the problem of generating a plausible colored photograph of ancient, historically black and white images using deep learning techniques without direct human intervention. In order to formulate the process , a wide range of datasets have been set up . A CNN model will be set up which will convert all the training images from RGB color space to the Lab color space. The L channel will be used as the input to the network and train the network to predict the ab channels and the next input L channel will be combined with the predicted ab channels. The model is tested and run to generate the result.**

Keywords — Convolutional; neural; networks; Lab-Color; Channel

I. INTRODUCTION

Automated Colorization of Black and white images has been subject to much research within the computer vision and machine learning communities. Beyond simply being fascinating from an aesthetics and artificial intelligence perspective, such capability has broad practical applications ranging from video restoration to image enhancement for improved interpretability.

Here, we take a statistical – learning – driven approach towards solving this problem. We design and build a convolution neural network (CNN) that accepts a black and white image as an input and generates a colorized version of the image as its output; Figure 1 shows an example of such a pair of input and output images. The system generates its output based solely on images it has "learned from" in the past, with no further human intervention.

In recent years, CNNs have emerged as the de facto standard for solving image classification problems, achieving error rates lower than 4% in the Image Net Challenge. CNNs owe much of their success to their ability to learn and discern colors, patterns and shapes within images and associate them with object classes. We believe that these characteristics naturally lend themselves well to colorizing images since object classes, patterns and shapes generally correlate with color choice.

II. SCOPE OF THE PROJECT

In this Project , The scope of our work is to –

- Detect and recolor a black and white image
- Image Restoration

Classic deep learning methods are based on the assumption that the data are vectors to exploit basic operations such as convolutions. While this suffices for many signals' classification problems such as speech, image, and video

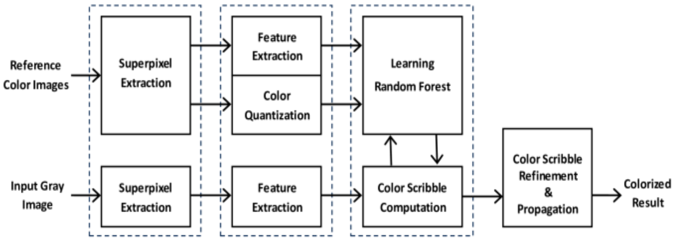
Fig. 1– Sample Input Image (Left)
Fig. 2– Output image (Right)

classification/segmentation, in various applications the data have other structures. Image Colorization plays a key role in many fields like Multimedia Retrieval and Human-Machine Interaction. Vocabulary used to address people changes according to the age groups very often.

The objectives of the project are mainly to input a black and white image and then detect it and convert it into a colored image and then display it. The dataset which is used for the project is OpenCV. OpenCV is the open source computer which enables the system to recognize images and patterns to give desired results. The system supports deep learning frameworks such as Tensor Flow , Keras etc.

III . PROBLEM DEFINITION

The problem of the project states that to automate the colorization of black & white pictures through a deep neural network. While it is probably difficult to predict exactly the correct colors due to the loss of information (a tulip for example can be of many different colors), the recognition of items within



a picture (tree, leaf, type of flower, sky, insect, etc.) should lead to a realistic color. Reconstructing the missing color of a grayscale pixel is here viewed as the problem of automatically selecting the best color among a set of color candidates.

IV . METHODOLOGY

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

- The L channel encodes lightness intensity only
- The channel encodes green-red.
- And the b channel encodes blue-yellow

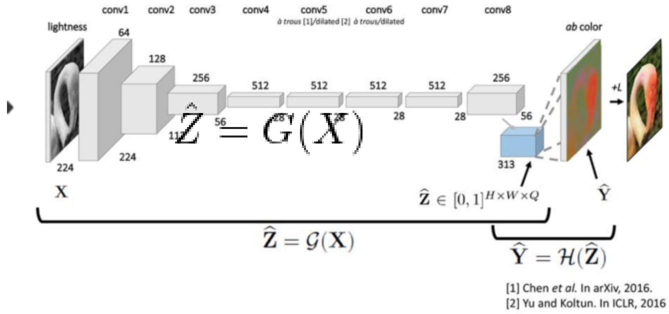
A full review of the Lab color space is outside the scope of this post (see this guide for more information on Lab), but the gist here is that Lab does a better job representing how humans see color.

Since the L channel encodes only the intensity, we can use the L channel as our grayscale input to the network.

From there the network must learn to predict the a and b channels. Given the input L channel and the predicted ab channels we can then form our final output image.

The entire (simplified) process can be summarized as:

- Convert all training images from the RGB color space to the Lab color space.
- Use the L channel as the input to the network and train the network to predict the ab channels.
- Combine the input L channel with the predicted ab channels.
- Convert the Lab image back to RGB.



$$\hat{Z} = G(X)$$

The input image is rescaled to 224×224. Let us represent this rescaled grayscale input image by .When it passes through the neural network shown above, it gets transformed to $\hat{Z}$ by the neural network. Mathematically, this transformation G by the network can be written as

The dimensions of $\hat{Z}$ is HxWxQ, where $H(=56)$ and $W(=56)$ are the height and width of the output of the last convolution layer. For each of the HxW pixels, $\hat{Z}$ contains a vector of $Q(=313)$ values where each value represents the probability of the pixel belonging to that class. Our goal is to find a single pair of ab channel values for each probability distribution $\hat{Z}_{h,w}$.

V . TECHNOLOGY USED

Software Used –

- Spyder(Anaconda)
- Tensorflow
- OpenCV
- Keras

Hardware Used –

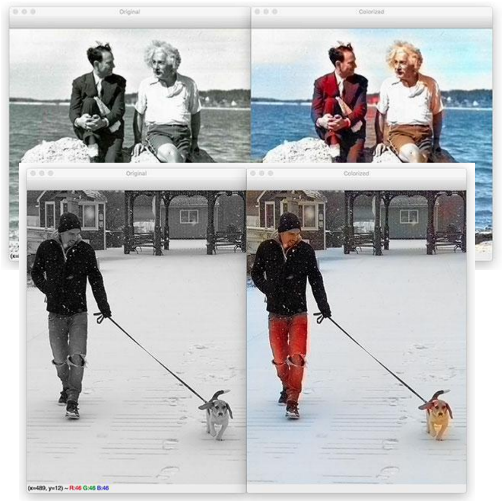
In all Kinds of Hardware that supports recent technology

VI . BLOCK DIAGRAM

(Process Flow Diagram of Image Colorization)

(Simplified Step Diagram of Image Colorization)

VII . OUTPUT SCREEN



■ These images of output Screen shows the given input image in monochrome in the left window and the given output image in color in the right window. Colorization can be a powerful pretext task for a self supervised feature for learning, acting as a cross-channel encoder.

■ Consider a grayscale image, if we look it seems less graceful because the picture is not appealing and the color features which are possessed by the objects in it are lost and it seems very hard to digest. If we pay a little close attention to it, we know that certain semantics possess the same features like: the sky is typically blue, and the grass is typically green.

■ As we know the prediction of color is free, and we can use any color photo to train the model. The prediction of the colors is multimodal, which means several objects can take on several colors.

VIII . BENEFITS OF THE PROJECT

■

These projects will provide benefits to these applications in the following sectors –

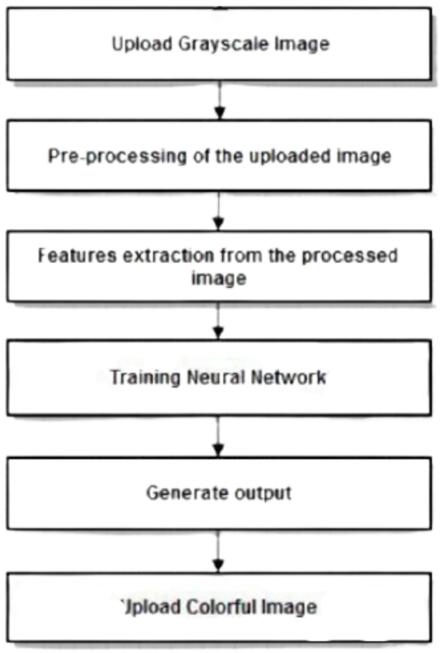
Astronomy **-** The Hubble telescope doesn't use color film. Its cameras record light from the universe with special electronic detectors. These detectors produce images of the cosmos not in color, but in shades of black and white. The matrix corresponding to the image is sent back to the earth station, different filters are applied (red, green and blue) and the equivalent matrices are sent back to the earth station where these are combined to form a color image. This process is tedious and incurs communication cost, hence our system makes this process of colorization automatic and more effective.

Archeology - Archeology Images captured a few decades ago when there were no color cameras are still in grayscale format and the process to convert them to color requires manual effort. Hence this application is chosen so that the process of colorization can be atomized.

Electron Microscopy - Electron microscope is used to observe objects that cannot be seen with naked eye, such as viruses and bacteria. The electron microscope functions by passing an electron beam and focusing it through electromagnets; this beam is either passed through the object or reflected off the surface of the object on a fluorescent plate.

This electron beam does not transfer color information, and hence the images captured thereby are gray. Colors are essential in electron microscopy as it represents vital information. Hence our system would make this colorization process simpler.

IX . CONCLUSION AND FUTURE WORK



Through our experiments, we have demonstrated the efficacy and potential of using deep convolutional neural networks to colorize black and white images.

In particular, we have empirically shown that formulating the task as a classification problem can yield colorized images that are arguably much more aesthetically-pleasing than those generated by a baseline regression-based model, and thus shows much promise for further development. Our work therefore lays a solid foundation for future work.

Moving forward, we have identified several avenues for improving our current system. To address the issue of color inconsistency, we can consider incorporating segmentation to enforce uniformity in color within segments. We can also utilize post-processing schemes such as total variation minimization and conditional random fields to achieve a similar end.

Finally, redesigning the system around an adversarial network may yield improved results, since instead of focusing on minimizing the cross-entropy loss on a per pixel basis, the system would learn to generate pictures that compare well with real-world images.

Based on the quality of results we have produced, the network we have designed and built would be a prime candidate for being the generator in such an adversarial network.

■ ACKNOWLEDGMENT

We have great pleasure in presenting the report on Image Colorization. We take this opportunity to express our sincere thanks towards our project mentor Prof. Tridip Chakraborty, Assistant Professor Department of Information Technology, GNIT for providing the technical guidelines and suggestions regarding the line of work. We would like to express our gratitude towards her constant encouragement, support and guidelines through the development of the project.

Words are inadequate in offering our thanks to the other mates, teachers and other members at GNIT, Kolkata for their encouragement and cooperation in carrying out this project work. The guidance and support received from all the members who are contributing to this project was vital for the success of this project.

REFERENCES

[1] Lasagne. https://github.com/Lasagne, 2015.

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

[3] X. Glorot and Y. Bengio. Understanding the difficulty of training deep feed forward neural networks. In International conference on artificial intelligence and statistics, pages 249–256, 2010.

[4] B. Hariharan, P. Arbelaez, R. Girshick, and J. Malik. Hypercolumns for object segmentation and fine-grained localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 447–456, 2015.

[5] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. arXiv preprint arXiv:1512.03385, 2015.

[6] M. J. Huiskes and M. S. Lew. The mir flickr retrieval evaluation. In Proceedings of the 1st ACM international conference on Multimedia information retrieval, pages 39–43. ACM, 2008.

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

[8] D. Kingma and J. Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097–1105, 2012.

[10] A. Oliva and A. Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. International journal of computer vision, 42(3):145–175, 2001.

[11] A. Olmos et al. A biologically inspired algorithm for the recovery of shading and reflectance images. Perception, 33(12):1463–1473, 2004.

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

[13] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[14] I. Sutskever, J. Martens, G. Dahl, and G. Hinton. On the importance of initialization and momentum in deep learning. In Proceedings of the 30th international conference on machine

[1] learning (ICML-13), pages 1139–1147, 2013. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.