

Automatic Colorization of Black and White Images Based on CNN

Varun D Trivedi¹, Hussain Saifuddin², Shubham Gudadinni³, Soyash S Sondhi⁴, M.A.R. Shabad⁵

B.E Students, Department of Computer Engineering, Sinhgad Academy of Engineering, Pune, India^{1,2,3,4}

Professor, Department Of Computer Engineering, Sinhgad Academy of Engineering, Pune, India⁵

ABSTRACT: We propose a new framework for automatic colorization of gray scale images by using the composition of features from multiple colour images to solve for the most likely colouring with machine learning. The intention of this process is to alleviate the amount of manual labour required to colorize a photograph. Given a target gray scale image and several context colour images, our methodology will colorize the target image using the data from the colour context images. The algorithm calculates custom user-defined features for each of the context images, and then uses these features to generate a local loss function for the colours at each pixel of the gray scale image. These loss functions are then used to generate a Markov Random Field, which is solved for the most likely colouring using graph cuts. Simple features have been tested such as the luminance or variance of 5x5 pixel neighbourhoods, and more complicated features have also been tested involving a combination of luminance values and texture statistics. In general, we have found that features carrying more texture-based information tend to result in better final colorizations. More complex and descriptive features could be tested, but at the cost of algorithm performance.

KEYWORD: Gray Scale Colorization, CNN, Machine Learning;

I. INTRODUCTION

A.BACKGROUND

Automatic image colorization is the act of hypothesizing the color data for a black and- white image without the need for user input or intervention. Clearly, there is not enough data within a single gray scale image to accurately recover the color data at each pixel. Due to the need for prior knowledge, and the lack of sufficient implicit data within the gray scale image itself, this problem is ill-posed and thus becomes a procedure of manufacturing accurate context from some reference as opposed to extracting it from the image itself. Put simply, image colorization can be formulated as a learning problem. There is currently much interest in the colorization of old photographs that were originally taken in black-and-white. From the perspective of professionals, there are entire companies that are dedicated to the re-touching and colorization of old photographs and movies. The process, by which this is done, especially in the case of movies, requires large amounts of skilled manual labor. From the perspective of hobbyists, entire internet communities have formed around users requesting for skilled individuals to colorize their old photographs (generally of distant relatives). The hobbyists who perform these colorizations generally bring their own time and expertise, often from a professional setting, to these communities.

B.PROBLEM DEFINITION

- Automatic colorization of black and white images system takes black and white images as input and converts it into a colourful image as output using convolutional neural network.

C.OBJECTIVES

- To identify different objects in the image.
- To predict the possible colour distribution for each pixel in the image.
- To calculate the loss and loss function.

II. LITERATURE SURVEY

Sr. No.	Paper Title/Author Details	Methods	Limitations
1.	Multi-Discriminator Generative Adversarial Network For High	In this paper, they introduced a novel generative adversarial	Complications arises due to multiple Discriminator used



	Resolution Gray-Scale Satellite Image Colorization Feimo Li, Lei Ma, And Jian Cai Institute Of Automation, Chinese Academy Of Science.	network with multiple discriminators for colorizing grey-scale satellite images with pseudo-natural Appearances.	which higher error rates.
2.	Automatic Example-based Image Colorizations using Location-Aware Cross-Scale Matching. Bo Li, Yu-Kun Lai, Matthew John, Paul L. Rosin.	This paper presents a novel automatic colorization algorithm that transfers color information from the reference image to the destination image.	Need for multiple reference images in order to produce more reliable results, due to the problem of having both reflectance and illumination to solve.
3.	Image Colorization with Deep Convolutional Neural Networks Jeff Hwang jhwang89@stanford.edu You Zhou youzhou@stanford.edu.	They present a convolutional-neural network-based system that faithfully colorizes black and white photographic images without direct human assistance.	Images are severely desaturated and generally unattractive.
4.	Automatic Image and Video Colorizationss using Deep Learning. Brian Sam Thomas Department of Computer Engineering St. John College Of Engineering and Management Village Vevoor, Manor Road, Palghar (East) 401404	They design and build a Convolutional Neural Network (CNN)-Long Short- Term Memory (LSTM) that accepts a black-and-white image as an input and generates a coloured version of the image as its output.	Once Relax activation ends up in this state, it is unlikely to recover.
5.	A Method for Colouring Lowresolution Black and White Old Movies Through Object Understanding. Ming Fang, Yu Song, Xiaoping Bai, Yushu Ren, Shuhua Liu.	This paper proposes a method for colouring low-resolution black and white old movies through object understanding. The approach first employs Microsoft's COCO dataset to train Conditional GAN (CGAN).	The trained model is not accurate due to lack of hardware required to obtain desirable image.

The most recent methods in this paradigm proposed by Iizuka et al. (2016), Larsson et al. (2016), and Zhang et al. (2016), use deep networks and are fully automatic. Although this makes colorizing a new photo cheap and easy, the results often contain incorrect colors and obvious artifacts. More fundamentally, the color of an object, such as a t-shirt, is often inherently ambiguous – it could be blue, red, or green. Current automatic methods aim to choose a single colorization, and do not allow a user to specify their preference for a plausible, or perhaps artistic, alternative. Might we be able to get the best of both worlds, leveraging large-scale data to learn priors about natural color imagery, while at the same time incorporating user control from traditional edit propagation frameworks? We propose to train a CNN to directly map gray scale images, along with sparse user inputs, to an output colorization. During training, we randomly simulate user inputs, allowing us to bypass the difficulty of collecting user interactions. Though our network is trained with ground truth natural images, the network can colorize objects with different or even unlikely colorizations, if desired. Most traditional tools in interactive graphics are defined either procedurally – e.g., as a designed image filter – or as constraints applied in a hand-engineered optimization framework. The behavior of the tool is therefore fully specified by human fiat. This approach is fundamentally limited by the skill of engineers to design complex operations/constraints that actually accomplish what is intended of them. Our approach differs in that the effect of interaction is learned. Through learning, the algorithm may come up with a more powerful procedure for translating user edits to colorized results than would be feasible by human design. Our contribution are as follows: (1) We end-to-end learn how to propagate sparse user points from large-scale data, by training a deep network to directly predict the mapping from grayscale image and user points to full color image. (2) To guide the user toward making informed decisions, we provide a data-driven color palette, which suggests the most probable colors at any given location. (3) We run a study, showing that even given minimal training with our interface and limited time to colorize an image (1 min), novice users can quickly learn to produce colorizations that can often fool real human judges in a real vs. fake test. (4) Though our system is trained on natural images, it can also generate unusual colorizations. (5) We demonstrate that this framework is not limited to user points, and can, in principle, be trained with any statistic of the output, for example, global color distribution or average image saturation.

III. PROPOSED SYSTEM ARCHITECTURE

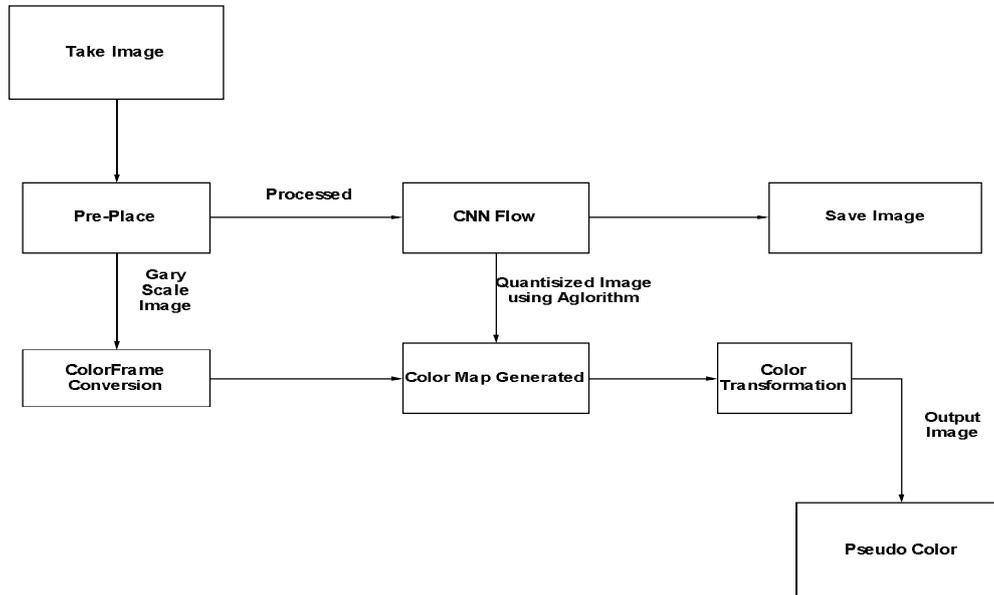


Fig.1: System Architecture

Colorization algorithms mostly differ in the ways they obtain and treat the data for modeling the correspondence between grayscale and color. Non-parametric methods, given an input grayscale image, first define one or more color reference images (provided by a user or retrieved automatically) to be used as source data. Then, following the Image Analogies framework color is transferred onto the input image from analogous regions of the reference image. Parametric methods, on the other hand, learn prediction functions from large datasets of color images at training time, posing the problem as either regression onto continuous color space or classification of quantized color values. Our method also learns to classify colors, but does so with a larger model, trained on more data, and with several innovations in the loss function and mapping to continuous output. Concurrent work on colorization. Concurrently with our paper, Larsson et al. and Iizuka et al. [24] have developed similar systems, which leverage large-scale data and CNNs. The methods differ in their CNN architectures and loss functions. While we use a classification loss, with rebalanced rare classes, Larsson et al. use an un-rebalanced classification loss, and Iizuka et al. use a regression loss. In Section 3.1, we compare the effect of each of these types of loss function in conjunction with our architecture. The CNN architectures are also somewhat different: Larsson et al. use hyper columns on a VGG network. Iizuka et al. use a two-stream architecture in which they fuse global and local features, and we use a single-stream, VGG-styled network with added depth and dilated convolutions.

IV. ALGORITHM USED

Convolutional Neural Network (CNN):-

In proposed work we are using CNN which takes images frames as a input. After getting frames from image it will be processed using image processing techniques for feature evaluation. We extract different features from those images regardless of their events in it consists. By using a series of mathematical functions we are going to identify the abnormal events. Every layer in CNN has capability to find out weights of images by using matrix evaluations which converts input to output with valuable functions. Layers of CNN used to identify fire events from extracted frames and give prediction by preserving high accuracy and less time.

- Step 1- Input Black & White image
- Step 2- Frame extraction from images
- Step 3- Image processing by using open-cv
- Step 4- Feature Extraction from images
- Step 5- Model generation
- Step 6- converted to Color Image.

Four main layer working approach of CNN explained below:-

a) Convolution Layer

We are going to extract different features of frames like pixel weight matrix calculations by using feature kernels. Perform mathematical convolutions on frames, where every function uses a unique filter. This outcome will be in different feature maps. At the end, we will collect all of these feature maps and draft them as the destination output matrix of the convolution layer.

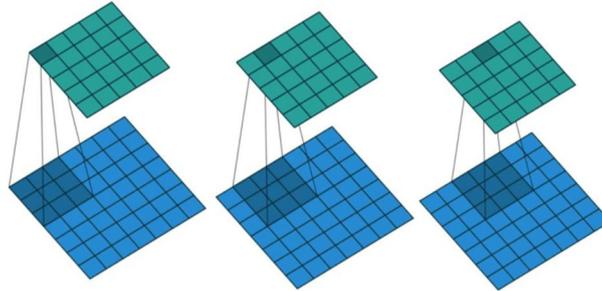


Fig.2 Convolution Layer

b) Pooling

The expression of pooling is to constantly decrease the dimensionality to limits the number of factors and calculation in the network. This limits the time of training and maintains over fitting problem. The max Pooling extracts out the largest pixel value out of a feature. While pooling average is calculated for the average pixel value that has to be evaluated.

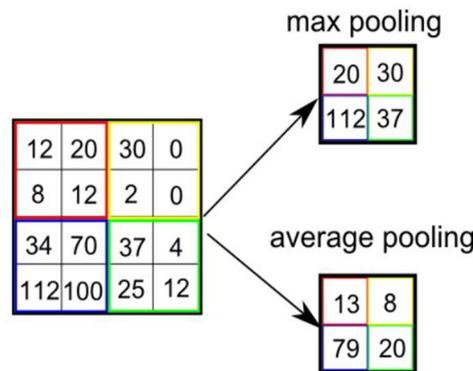


Fig.3 Pooling Layer

c) Flattening

Generally here we put the pooled feature into a single column as a sample input for further layer (transform the 3D matrix data to 1D matrix data)

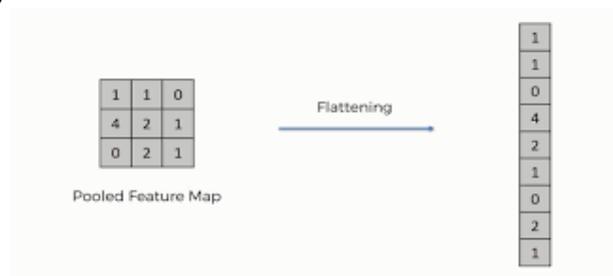


Fig.4 Flattening Matrix

d) Fully Connection

A fully connected layer has full connections of neurons to all the nodes in the previous layer. The fusion of more neurons to evaluates accurately.

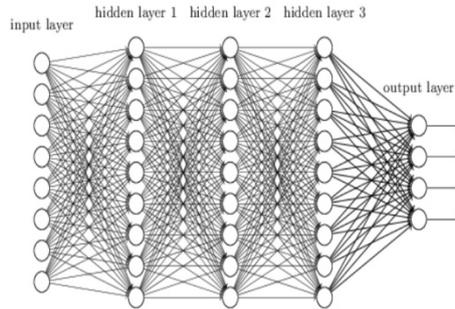
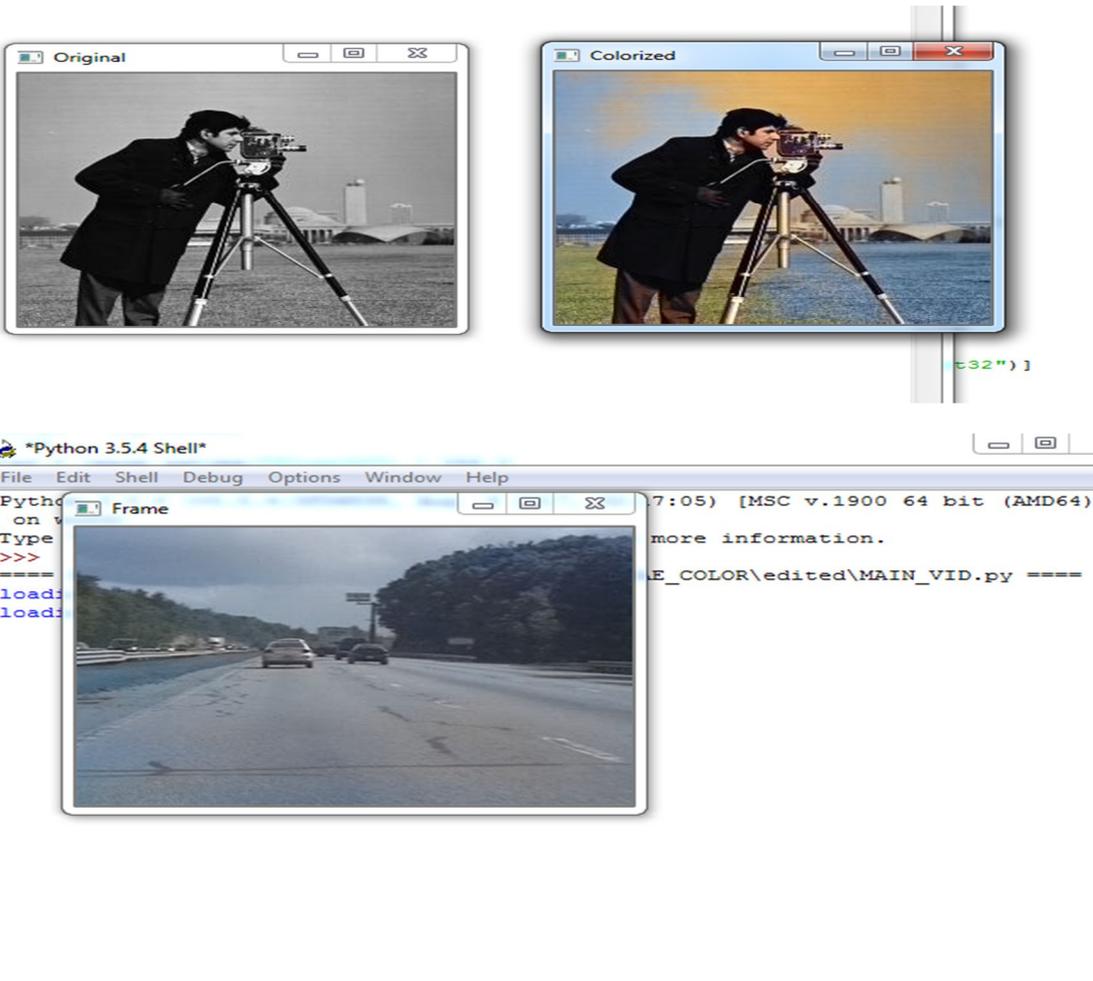


Fig.5 Fully Connected Layer

V. RESULT

```
0]
10]
20]
30]
40]
50]
60]
70]
110]
100]
-90]
-80]
-70]
-60]
-50]
-40]
-30]
-20]
-10]
0]
10]
20]
30]
40]
50]
--
```



The screenshot shows a Python 3.5.4 Shell window with a menu bar (File, Edit, Shell, Debug, Options, Window, Help). The shell contains the following code and output:

```
Python 3.5.4 Shell
File Edit Shell Debug Options Window Help
Python 3.5.4 Shell
>>>
====
load:
load:
```

The output shows two windows side-by-side. The left window, titled 'Original', shows a grayscale image of a man in a dark coat looking through a camera on a tripod. The right window, titled 'Colorized', shows the same image with color applied. Below these is a window titled 'Frame' showing a color image of a road with cars. The shell also displays some system information: '7:05) [MSC v.1900 64 bit (AMD64) more information. E_COLOR\edited\MAIN_VID.py ====='

VI. CONCLUSION

We have proposed a novel automatic colorization neural network, which uses multi-scale convolution kernels and combines low and middle features extracting from VGG-16. Although only trained to color, 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. 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 color 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 color, 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.

ACKNOWLEDGMENT

This work is supported in Pesticide detection on Banana field in India. Authors are thankful to Faculty of Engineering and Technology (FET), Savitribai Phule Pune University, Pune for providing the facility to carry out the research work.

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 _delity. ACM Transactions on Graphics (TOG) 26(3) (2007) 76
5. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
6. Bengio, Y., Courville, A., Vincent, P.: Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence 35(8) (2013) 1798{1828
7. Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., Ng, and A.Y.: Multimodal deeplearning. In Proceedings of the 28th international conference on machine learning (ICML-11). (2011) 689{696
8. Agrawal, P., Carreira, J., Malik, J.: Learning to see by moving. In: Proceedings of the IEEE International Conference on Computer Vision. (2015) 37{45
9. Jayaraman, D., Grauman, K.: Learning image representations tied to ego-motion. In: Proceedings of the IEEE International Conference on Computer Vision. (2015) 1413{1421