

## COLORIZATION OF GRAYSCALE IMAGES USING DEEP LEARNING

Apurva Pavaskar<sup>1</sup>, Saurabh Dawkhar<sup>2</sup>, Diksha Kamble<sup>3</sup>, Praneet Mugdiya<sup>4</sup> & S. F. Sayyad<sup>5</sup>

<sup>1,2,3,4</sup> Research Scholar, Department of Computer Engineering, AISSMS College of Engineering, Pune, Maharashtra, India

<sup>5</sup> Professor, Department of Computer Engineering, AISSMS College of Engineering, Pune, Maharashtra, India

### ABSTRACT

Commercially viable photography began in the late 1820s, but the effective on-set of color photography started as late as 1970s. During this period of over a century, photographs captured were mostly black and white. Colorization plays a vital role in representing the true virtue of real-world manifestations. The human eye perceives color and often remembers information about an object based on its coloration. Colorization problem is difficult to solve without manual adjustment. The proposed systems develop colored versions of gray scale images that closely resemble the real-world versions. The use of Artificial Neural Networks in the form of Convolutional Neural Networks (CNN) and Generative Adversarial Networks (GAN) to learn about features and characteristics through training allows for assigning plausible color schemes without human intervention.

**KEYWORDS:** Colorization, Computer Vision, Artificial Intelligence, Deep Learning, Convolutional Neural Network, Generative Adversarial Network

---

### Article History

**Received: 04 Mar 2020 / Revised: 26 Mar 2020 / Accepted: 10 Apr 2020**

---

### INTRODUCTION

Colorization is tricky. Black and white images can be represented in grids of pixels. Each pixel has a value that corresponds to its brightness. The values span from 0–255, from black to white. One way to transform these images or to restore the old classics is to manually select a colour for every pixel partition, but its taxing and time consuming, and the amount of data is just overwhelming. The compelling need of the newer generations to view images from historical events, in colour, is undeniably growing, and to deliver highly accurate and extensive results is calling for faster and better autonomous colorization techniques.

### Importance of Colorization Techniques

Growing influence of social media and increasing references to past events has generated a need to make historically curated images widely available. Younger generation needs to be reminded of great things accomplished by their ancestors.

People immortal through thought will only live on, if they are known in the farthest outreaches of the community, otherwise they seem to wither away like memories. Early technologies could not capture color but their ability to capture emotion was just as fervent. To express these sentiments in a better light and to keep the interests at a high, colorization is becoming incumbent.

To preserve history and heritage and to make new discoveries in the lesser known domains of medical and space research, many individuals are using applications like Photoshop to manually add color to gray scale images. The volume of such data is enormous and it is a tedious job to colorize images manually. Automated systems that can learn from the robust available data can save a lot of effort and can process extensive volumes without human intervention.

## ASSOCIATED TECHNIQUES

### Artificial Intelligence and Deep Learning

Artificial Intelligence is a branch of Computer Science that focuses on design and implementation of algorithms that mimic human thinking. It strives to minimize human intervention in machine operations and make the systems intelligent enough to solve problems based on prior learning. Deep Learning is a subset of Artificial Intelligence which involves creation of Neural Networks with multiple hidden layers. Neural Networks are constructs that try to replicate human decision making by considering various possibilities, weights and biases of instances and giving appropriate outputs. They are used for training models to perform future tasks without human intervention.

Figure 1 shows a Convolutional Neural Network (CNN) is a Deep Learning algorithm which can assign measurable values to various aspects of the image in the form of weights and biases. CNN requires less pre-processing when compared to other classifiers. In most methods, constant tweaking of the filters is required to accommodate different instances, but in the case of a CNN, the model can learn to tweak the weights and biases by itself. With enough training, the CNN have the ability to understand the features and characteristics of the model.

### Convolutional Neural Networks

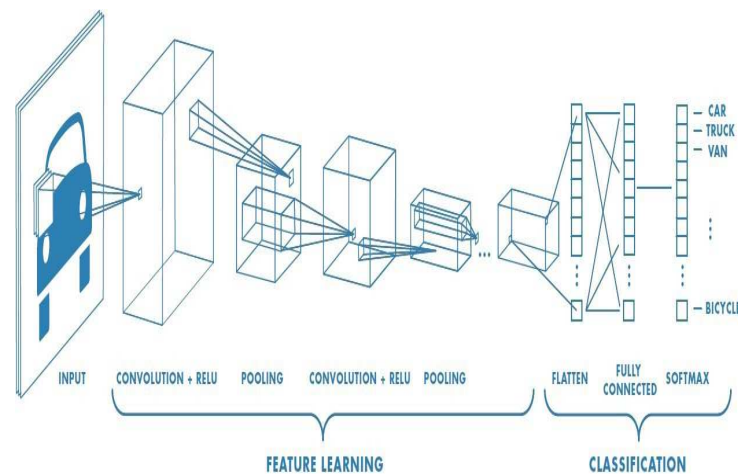
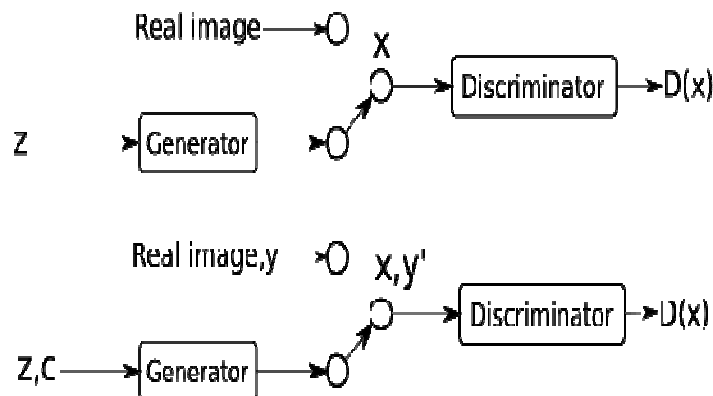


Figure 1: Example of Convolutional Neural Network (From [25]).

### Generative Adversarial Neural Networks

Generative Adversarial Networks (GANs) are algorithmic architectures that use two neural networks which compete against each other in order to generate instances of data that can be considered as real data. There are two main components: Generator and Discriminator. The generator is trained to produce outputs that can deceive an adversarially trained discriminator into believing that a generated image is a real one. The discriminator is trained to distinguish between a real image and an image produced by the generator. As they both get better at doing their tasks, the level of images that get produced and pass the discriminator, are very close to reality.

Conditional Generative Adversarial Network (cGAN) is a type of GAN in which certain conditions are predetermined. The network considers an additional parameter while performing various operations in cGANs.



**Figure 2: GAN Vs CGAN (From [29]).**

### Related Work

In [1] the authors have used Convolutional Neural Networks (CNN). The deep network uses a fusion layer that collaborates information from local features (or small patches) and global priors (or larger sections of the image) to colorize images. The model exploits semantic labels of the existing dataset to learn more discriminative global features. The network is formed by several sub components using Directed Acyclic Graphs. Four main components of the model are: Low level features network, mid level features network, Global features network, and colorization network. The output of the model is the chrominance of the image fused with its luminance. The model convolutionalizes the image in two parallel sub systems. One, divides the image in to small patches and derives features from the parts, and the other, analyzes the image as a whole and derive information about parts of the image that usually cover larger portions of the image (Eg: sky, sea, grass, trees, et cetera). Features from these two parallel operations are combined in the fusion layer and the intermediate output is scaled up and generated color values are added to it. The information is combined with the original black and white image and the final output is produced.

In [2] the authors have used Conditional Generative Adversarial Network (cGAN). Both the components- Generator and Discriminator- use modules of the from convolution-Batch Norm-ReLu. The model chooses features based on convolution and the generator and discriminator get trained at every stage. The model also provides the generator with paths to circumvent bottleneck layers in passing of the information. The input and the output data have similar structure and thus the low level data will be easier to handle if it makes its way from the down-sampling side to the up-sampling side of the network with a skip. The generator comes up with color values for parts of the image and the delivers an output image for the discriminator to validate. The discriminator checks the image for believability and either allows it to pass or rejects it. Gradually, the generator and the discriminator become better at their jobs. A persuasive generator tries to convince a fastidious discriminator. The images that are validated by the discriminator are given as the output by the model.

[3] The paper focuses on producing a plausible colour scheme for the input image based on its learning and providing vibrant results by re- balancing. They predict a distribution of possible colours for each pixel, to get the best possible approximations. They use Convolutional Neural Networks, CIE L\*a\*b\* colour space and an off the shelf VBB network. They focus on the design of the objective function, and a technique for inferring point estimates of colour from the predicted colour distribution, which is the annealed-mean of the distribution. They also re-weight the loss, at training

time, to emphasize rare colours to exploit the full diversity of the large-scale data on which it is trained. Another notable feature of the method is that they use a colorization Turing Test in training of the model.

[7] In this paper, they propose a novel colorization model which combines two Convolutional Neural Networks and uses multi-scale convolution kernels to get better spatial consistency. A multi-scale convolution kernels and combines low and middle features extracting from VGG-16. Experiments prove that this model is able to perform pretty good colorization on images from Chinese black and white films without any user interventions. In order to address the problem that the current training datasets for colorization do not applicable to historical old photographs, an image dataset is established by extracting frames from Chinese colour films of the last century.

[8] In this article, they have described a method for a more general form of colour correction that borrows one image's colour characteristics from another demonstrates that a colour space with decor related axes is a useful tool for manipulating colour images. Imposing mean and standard deviation onto the data points is a simple operation, which produces believable output images given suitable input images. They manipulated RGB images, which are often of unknown phosphor chromaticity. Applications of this work range from subtle post-processing on images to improve their appearance to more dramatic alterations, such as converting a daylight image into a night scene. We can use the range of colours measured in photographs to restrict the colour selection to ones that are likely to occur in nature.

[9] In the paper, they train the model to predict per-pixel color histograms. This intermediate output is used automatically to generate a colour image or further manipulate prior to image formation. In this, a greyscale image is processed through a deep convolutional architecture (VGG). Deep convolutional neural networks (CNNs) can serve as tools to incorporate semantic parsing and localization into a colorization system. A deep neural architecture that is g all kinds of objects is used for training the neural networks. Informative features like patch feature, DAISY, and a semantic feature are obtained and are fed into the neural network. And they have also compared their results and described the limitations of colorizing a grayscale image, trained end-to-end to incorporate semantically meaningful features of varying complexity into colorization. It uses color histogram prediction framework that handles uncertainty and ambiguities inherent in colorization while preventing jarring artefacts. Here, they propose fully automatic colorizer that produces strong results, improving upon previously leading methods by large margins on all datasets tested. They also propose a new large-scale benchmark for automatic image colorization, and establish a strong baseline with the method to facilitate future comparisons.

## Architecture

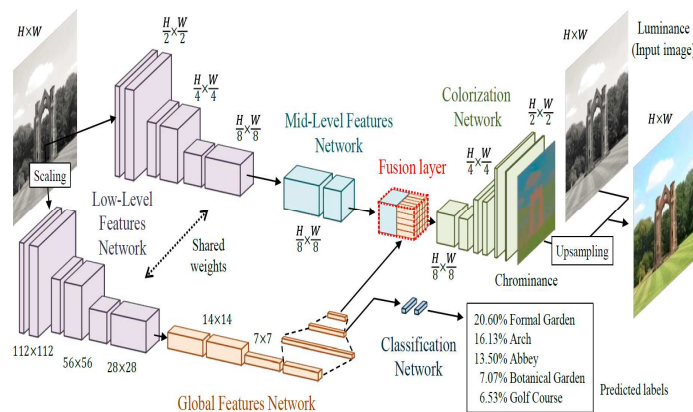


Figure 3: System Architecture (From [1]).

**Let there be Color!**

The authors have used Convolutional Neural Network (CNN) to design the colorization model. The CNN model consists of- 6 Low level features network layers, 2 Mid-level features network layers, 4 Global level features network layers, and 1 Colorization network layer.

The image is convulsed in to parts to study the low level features and the global level features concurrently. The global features network is used to predict class labels for parts of the image that cover more area and are common across nature with respect to shades and coloration. The image features are then combined in the colorization network by the fusion layer and scaled up to describe chrominance parameters and the mapped to the original black and white image to obtain output.

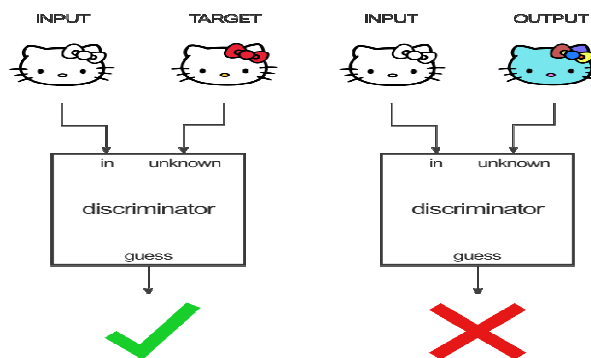
Figure 4 shows the authors have used Conditional Adversarial Network (cGAN) to design the colorization model. The cGAN model consists of- Generator Discriminator.

Figure 5 shows the Generator learns from the training data set and tries to produce images that can deceive the Discriminator in to thinking of it as a true image. The Discriminator is trained to discern true images from those produced by the generator. As the components learn more, they get better at their tasks. Since, the model is inherently competitive, the Generator starts developing highly indiscernible outputs which make their way past the Discriminator. Such outputs are often difficult for the human eye to tell apart, and thus a plausible coloured representation of the image is given as output.

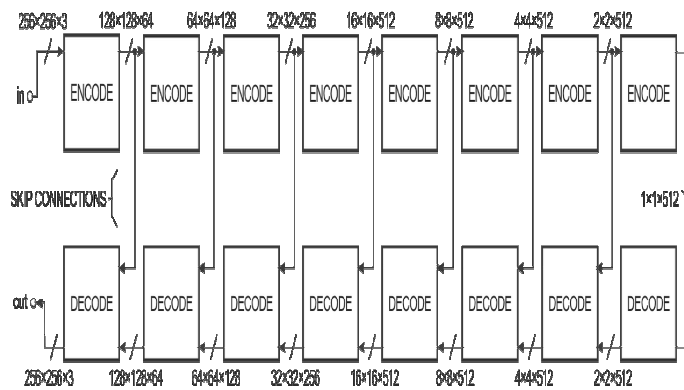
**Table 1: Details of the CNN Deep Layers (from [1])**

Type	Kernel	Stride	Outputs
fusion	-	-	256
conv.	3x3	1x1	128
up sample	-	-	128
conv.	3x3	1x1	64
conv.	3x3	1x1	64
up sample	-	-	64
conv.	3x3	1x1	32
output	3x3	1x1	2

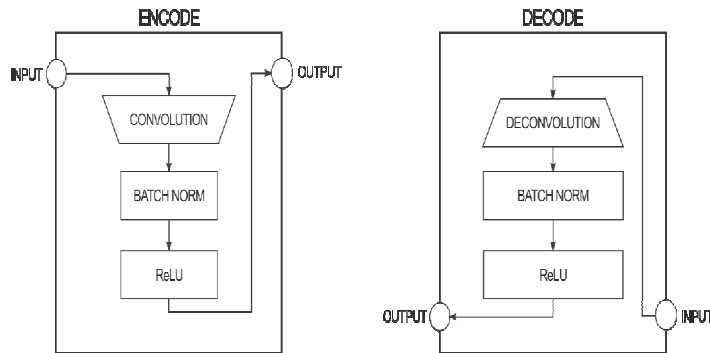
**Pix 2 Pix**



**Figure 4: Working of Generator and Discriminator in a Generative Adversarial Network (From [3]).**



**Figure 5: Block Diagram For Skipping Generator (From [3]).**



**Figure 6: Block Diagram for Convolution-Batchnorm-Relu (From [3]).**

**TOOLS AND LIBRARIES**

**Tensor Flow**

It is an open-source end-to-end platform that allows users to develop projects in the domain of Machine Learning. It houses a flexible ecosystem of tools, a rich community base and a robust collection of library and supporting resources. Tensor Flow offers necessary abstraction for beginners to understand and empowers them to create machine learning models.

**Keras**

It is a high level Application Programming Interface (API) designed in python for operations on Neural Networks. It offers pellucid and fluid prototyping and is compatible with Central Processing Unit as well as Graphical Processing Unit. It provides a user friendly experience that even beginners can develop sustainable systems. It offers modularity and extensibility, and combines with Tensor Flow to provide easy modelling, training and testing.

**APPLICATIONS**

**Legacy photos and films**

Images and films generated before exist in black and white or shades of gray. The modern generations ought to know about the history, but the demanding census also requires the footage to be in a form that they are familiar with. Thus colorization of images and films from the world wars or 1800s is in demand. Manually coloring millions of such instances will not be an easy task without automated colorizers.

### **Astronomical imaging**

The best telescopes in the world, other than optical telescopes, produce images using special electronic detectors to detect various rays emitted by different heavenly bodies in the cosmos. Optical telescopes are more expensive to produce and maintain. Telescopes based on technologies like Radio waves, X-ray, Ultraviolet rays, Gamma rays, produce gray scale images. Colorization of such valuable data can provide insightful knowledge and help us discover further secrets of the cosmos.

### **Medical Imaging**

Medical imaging and radiology develop images in grayscale. The medical sector is a place where even the smallest of mistakes can have lethal repercussions. Even the tiniest of advancements can reveal information that can save lives. Colorization of such important scans based on study of corpses or other sources could possibly help the sector make new discoveries and can lead to newer safer methods of operations.

### **ACKNOWLEDGEMENTS**

We would like to extend our sincere gratitude towards our project guide Dr S. F. Sayyad for her insightful guidance and amicable support.

### **REFERENCES**

1. Satoshi Iizuka, Edgar Simo-Serra, Hiroshi Ishikawa “Let there be Color!: Joint End- to-end Learning of Global and Local Image Priors for Automatic Image Colorization with Simultaneous Classification”, 2016 ACM
2. Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros “Image-to-Image Translation with Conditional Adversarial Networks” arXiv:1611.07004v3 [cs.CV] 26 Nov 2018
3. Richard Zhang, Phillip Isola, Alexei A. Efros, “Colorful Image Colorization”, IEEE
4. <https://affinelayer.com/pix2pix/> [online]
5. Zezhou Cheng, Qingxiong Yang, Bin Sheng, “Deep Colorization”, 2015 IEEE International Conference on Computer Vision (ICCV).
6. Yu Chen, Yeyun Luo, Youdong Ding, Bing Yu, “Automatic Colorization of Images from Chinese Black and White Films Based on CNN”, 2018 International Conference on Audio, Language and Image Processing (ICALIP).
7. Qingxiong Yang, Zezhou Cheng, Bin Sheng, “Deep Colorization”, 2016 IEEE Journal
8. Erik Reinhard, Michael Ashikhmin, Bruce Gooch, and Peter Shirley University of Utah” Color Transfer between Images” IEEE 2001
9. Gustav Larsson, Michael Maire, and Gregory Shakhnarovich, “Learning Representations for Automatic Colorization”, 2017 IEEE
10. Ayumu Shinya, Kazuki Mori, Tomohiro Harada, Ruck Thawonmas, “Potential improvement of CNN-based colorization for non-natural images”, 2018 International Workshop on Advanced Image Technology (IWAIT)

11. Carey Rappaport, "A Color Map for Effective Black-and-white Rendering of Color- Scale Images", *IEEE Antenna's and Propagation Magazine*, Vol. 44, No. 3, June 2002
12. Aditya Deshpande, Jason Rock and David Forsyth, "Learning Large-Scale Automatic Image Colorization", *International Conference on Computer Vision (ICCV)*, 2015, Santiago, Chile.
13. ArshiyaSayyed, ApekshaRahangdale, RutujaHasurkar, and KshitijaHande, "Automatic Colorization Of Gray-scale Images using Deep Learning ", *International Journal of Science, Engineering and Technology Research (IJSETR)*Volume 6, Issue 4, April 2017, ISSN: 2278 -7798532All Rights Reserved © 2017 IJSETR
14. FreeCodeCamp- Youtube [online]
15. [15] The Coding Train- Youtube [online]
16. Bhavesh Gupta, Stanford CS20SI- Youtube [online]
17. 3 Blue1Brown- Youtube [online]
18. <https://towardsdatascience.com/https-medium-com-piotr-skalski92-deep-dive-into-deep-networks-math-17660bc376ba> [online]
19. [19]<https://pathmind.com/wiki/generative-adversarial-network-gan> [online]
20. <https://lukemelas.github.io/image-colorization.html> [online]
21. [https://github.com/satoshiizuka/siggraph2016\\_colorization](https://github.com/satoshiizuka/siggraph2016_colorization) [online]
22. [https://github.com/2021rahul/colorization\\_tensorflow](https://github.com/2021rahul/colorization_tensorflow) [online]
23. <https://hackerstreak.com/colorizing-pokemon-with-deep-learning/> [online]
24. [https://github.com/Baakchsu/pix2pix\\_tensorflow](https://github.com/Baakchsu/pix2pix_tensorflow) [online]
25. <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53> [online]
26. [https://en.wikipedia.org/wiki/List\\_of\\_space\\_telescopes](https://en.wikipedia.org/wiki/List_of_space_telescopes) [online]
27. <https://www.tensorflow.org/> [online]
28. <https://keras.io/> [online]
29. [https://medium.com/@jonathan\\_hui/gan-cgan-infogan-using-labels-to-improve-gan-8ba4de5f9c3d](https://medium.com/@jonathan_hui/gan-cgan-infogan-using-labels-to-improve-gan-8ba4de5f9c3d)
30. [https://medium.com/@jonathan\\_hui/gan-a-comprehensive-review-into-the-gangsters-of-gans-part-1-95ff52455672](https://medium.com/@jonathan_hui/gan-a-comprehensive-review-into-the-gangsters-of-gans-part-1-95ff52455672)