An Automatic Broad Colorization Approach Using Deep Learning

Dr. Subedha. V, Sangavi R, Snega S, Swetha G

[1] Dr.V.Subedha Head of Department(CSE), Panimalar Institute of Technology, India
[2] Sangavi R, Panimalar Institute of Technology, India
[3] Snega S, Panimalar Institute of Technology, India
[4] Swetha G, Panimalar Institute of Technology, India

Abstract

The basic objective of colorization, which transforms grayscale images into aesthetically appealing colour images, is to persuade the user with reliable results. In this project, we apply deep learning methodology because user- guided methods requires more user interaction to produce a colorized image. Deep learning is used since it is a totally autonomous model and creates photos with accurate colors and contrast. Convolutional neural networks are used as a tool to automate the colouring process because we are aware of how laborious conventional methods are and because deep learning has many advantages (CNNs). A CNN is a particular kind of network architecture used by deep learning algorithms to process pixel input and perform tasks like image identification. Although there are many various types of neural networks used in deep learning, CNNs are the most commonly used network design for identifying and recognising objects. The neural network model automatically creates a function that converts a grayscale image's pixel properties to color values.

Keywords: Automatic methods, black-and-white image, colorization, deep learning methods, image quality assessment, Grayscale images.

1. Introduction

The perspective of the observer is significantly altered when black-and-white photographs are coloured. Instead of precisely reconstructing the colour, colorization aims to fool the observer and have him assume that the colourized image is real. Common applications for colorization include restoring old black-and-white films, colouring astronomy photographs, and reviving historical black-and-white photos. The human visual system's subjective reaction to electromagnetic energy with visible spectrum wavelengths between 380 nm and 780 nm is called colour. The ability to characterise an object using hue, brightness, and saturation. Understanding colour vision requires knowledge of physics, physiology, and psychology. Color perception is influenced by vision, light, and personal interpretation.

Deep learning models have recently experienced astounding success in a variety of application domains (including image classification, pedestrian detection and tracking, face detection, handwritten character classification, image super- resolution, photo adjustment, photo enhancement, sketch simplification, style transfer, in painting, image blending, denoising, etc.), and as a result, they hold out the promise of further innovative advancements in the near future. In this research, a thorough analysis of the currently used colorization techniques was carried out, along with a quantitative assessment of the outcomes. The automatic and interactive versions of the algorithms of Iizuka et al, Zhang et al, Levin et al, Su et al, Vitoria et al, and Zhang et al were analysed and evaluated.

As a result, the three categories of colorization methodologies include learning-based (or deep learning) methods, example-based methods, and scribble-based methods. The deep learning approaches are the most appealing.

2. Related Work

1) R. K. Gupta, A. Y.-S. Chia, D. Rajan, E. S. Ng, and Z. Huang, "Image colorization using similar images,"

[1] They suggest a brand-new automatic colorization technique that transfers colour information from a reference colour image to an input gray image by utilizing a variety of image attributes.

Disadvantages:

Their method may not work if adequate colour exemplars are not readily available because it depends on the availability of colour exemplars that are semantically comparable to the gray image.

Advantages:

Even with a fixed set of parameter settings, their approach produces perceptually pleasing colorization. Comparisons with current state-of-the-art techniques also show that our method is superior in colorization.

2) E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation,"

[2] Recognition technology is developing thanks to convolutional networks. Convnets are getting better at local tasks with structured output in addition to whole-image classification. They demonstrate that fully convolutional networks (FCNs) trained on semantic segmentation end-to- end and pixel-by-pixel without the use of additional equipment.

Benefits:

Fully convolutional networks are a diverse class of models with several applications to pixel-level problems.

3) R. Zhang, P. Isola, and A. A. Efros, "Colorful image colorization,"

[3] This system, which has been trained on more than a million colour images, is implemented as a feed- forward pass in a CNN during testing. Also, they test an algorithm using a colorization Turing test, which yields results that are almost photorealistic. They emphasize the objective function's construction as well as their method for extrapolating point estimates of colour from the projected colour distribution.

Advantage:

Colorization with a deep CNN and a carefully selected objective function can produce results that are nearly identical to true colour photographs. This method might be seen as a pretext challenge for representation learning in addition to producing meaningful graphics.

4) X. Liu et al., "Intrinsic colorization"

[4] An example-based colorization method that is resistant to lighting changes between grayscale target and colour reference images was presented by X. Liu et al. The approach carries out colour transfer in an illumination- independent domain that is largely devoid of shadows and highlights in order to do this. It first extracts an intrinsic reflectance image of the target scene that is independent of illumination from a variety of colour references found online.

5) G. Charpiat, M. Hofmann, and B. Schölkopf, "Automatic image colorization via multimodal predictions,"

[5] They suggested a technique for automatically and manually colouring greyscale pictures. If necessary, userprovided colour landmarks might then be used to interactively correct the colour proposition. They employ machine learning algorithms to extract as much data as they can from a dataset of coloured examples, working in the L- a-b colour space to simulate how humans perceive distances between hues. Unlike to their prior methods, the final algorithm is quick, made to be more robust to texture noise, and most importantly, it can deal with ambiguity. They will concentrate on more global approaches in their endeavour rather than relying heavily on extremely local texture-based classification or segmentation. The fact that they directly address the issue at hand, using the assistance of graph cuts, makes the framework more robust to noise and local prediction errors.

6) Yuxi Jin , Bin Sheng , Ping Li and C. L. Philip Chen , "Broad colorization"

[6] They suggested an automatic colorization method that integrates the local and global properties of the input gray-scale photos and is independent of user input and does not require a lengthy training period. Low-level, mid- level, and high-level features are combined as local features that signify cues that were present in the grayscale image. Before directing the colorization process, the global feature is viewed as data. The local broad learning system is taught to extract the chrominance value of each pixel from the local features, which might then be expressed as a chrominance map depending on the positioning of the pixels. Following that, the chrominance map is educated using the global broad learning system.

There are no user prerequisites for this strategy, and our framework's training time is orders of magnitude quicker than that of conventional deep learning techniques. The technology enables users to enhance training data without retraining the system in order to increase the user's subjective initiative.

7) M. Richart, J. Visca, and J. Baliosian, "Image colorization with neural networks,"

[7] The process is entirely automatic and does not rely on human input. It is independent of segmentation, writing, or complex image processing methods. It is based on back propagation training of a straightforward classifier over a set of training colour and corresponding grayscale images. Based on the grayscale of the pixels around it, the classifier guesses the colour of a pixel. This tiny patch records a regional texture. Self Organizing Maps are used to decrease the colors in order to keep the predictor's domain limited. A small number of chroma values with sufficient variance are produced by this reduction, and they can be used to create accurate approximations for each colour in the training set.

8) G. Larsson, M. Maire, and G. Shakhnarovich, "Learning representations for automatic colorization,"

[8] The method makes use of both low-level and semantic representations, leveraging current developments in deep networks. The model is trained to predict per-pixel colour histograms by simulating the appearance of as many scene features as possible using multi-modal colour distributions. This intermediate output can be further processed before an image is formed or utilised to automatically create a colour image. The model performs better than current techniques on colorization jobs that are entirely and partially automatic. Moreover, colorization is explored as a tool for learning self-supervised visual representations.

9) L. Yatziv and G. Sapiro, "Fast image and video colorization using chrominance blending,"

[9] Colorization is ambiguous in nature, requires some degree of human intervention, and is not unique in the mapping between intensity and colour. This research presents a computationally basic yet efficient colorization method. By delivering a condensed set of chrominance scribbles, the method is quick and easy to apply "on the fly," enabling the user to interactively obtain the necessary results right away. Due to the principles of luminance-weighted chrominance blending and rapid intrinsic distance computations, high-quality colorization results for still images and video are obtained at a fraction of the complexity and processing cost of previously published techniques. Using the algorithm described here as a potential expansion, or video, in addition to changing the underlying brightness, and many other unique effects presented here, the colours of an existing colour image might be altered.

10) Fuya Luo , Yunhan Li, Guang Zeng, Peng Peng, Gang Wang and Yongjie Li,"Thermal Infrared Image Colorization for Nighttime Driving Scenes With Top-Down Guided Attention"

In this paper, PearlGAN, a generative adversarial network based on top-down attention and gradient alignment, was proposed. The first step in reducing the semantic encoding uncertainty during translation is to build a top-down guided attention module and an elaborate attentional loss. The edge consistency between the translated and input pictures is then encouraged by the introduction of a controlled gradient alignment loss. Moreover, pixel-level annotation is done on datasets to assess how well different translation techniques preserve semantics.

3. DEEP LEARNING METHOD AND CNN NETWORK

The visual system of living beings has an impact on a certain type of multilayer neural network or deep learning architecture called a convolutional neural network (or CNN). Each colour image may be converted to grayscale and then matched with its corresponding colour version to create an easy training example. This makes obtaining training data simple. The algorithm employs multiple feed-forward passes to eventually take in a grayscale image and, in the words of the system's developers, "hallucinate" a reasonable (but not always accurate) set of colours to fill in the image.

Deep learning, a kind of machine learning, has made rapid advancements during the past 20 years. It is currently utilised in a wide variety of machines and devices and is more pervasive in our daily life. Neurophysiologists David Hubel and Torsten Wiesel discovered in 1959 that the neurons in a cat's brain are organised in layers. Later, their research was published under the title "Receptive fields of single neurons in cat's striate cortex." By initially extracting the local features and then integrating the features acquired for higher level representation, these layers learn to recognise visual patterns. In the end, this idea evolved into one of the pillars of deep learning. Kunihiko Fukushima proposed Neocog-nitron in 1980, a self-organizing neural network with multiple layers that can recognise visual patterns hierarchically through learning. In the end, CNN's initial theoretical model had this layout. Inspiration came from Hubel and Wiesel's books. Kunihiko Fukushima proposed Neocog-nitron in 1980, a self-organizing neural network with many layers that can recognise visual patterns hierarchically through learning. Its layout then evolved into the original theoretical model of CNN.

LeCun et. made a significant advancement over Neocognitron's architecture in 1989 by creating LeNet-5, a cutting-edge CNN framework that recognised the MNIST collection of handwritten digits. LeNet-5 was trained using the error back-propagation method, which eliminates the need for a separate feature engineering procedure, and it can recognise visual patterns straight from the original input photos. After the creation of LeNet-5, CNN encountered a number of challenges that hindered its ability to perform well in a range of challenging tasks, including as a shortage of large training data, a lack of method innovation, and insufficient computing power. Nevertheless, in the age of big data, we now have enormous labelled datasets, more sophisticated algorithms, and especially powerful GPU processors. AexNet, which won the ImageNet Large Scale Visual Recognition Competition in 2012, was developed by Krizhevsky et al. using this form of enhancement. The development of multiple CNN models and their application in diverse computer vision and natural language processing domains on the ImageNet Large Scale Visual Recognition Challenge were both made possible by the success of AlexNet. The success of AlexNet paved the door for the creation of several CNN models as well as their use in various computer vision and natural language processing fields.

Concepts of convolutional neural networks The Convolutional Neural Network (CNN), commonly known as ConvNet, is a form of Artificial Neural Network (ANN) that has a deep feed-forward architecture and exceptional generalising powers, in contrast to other networks with FC layers. It can more effectively recognise objects by learning highly abstracted features of them, notably spatial data. A deep CNN model consists of a small number of processing layers that are capable of learning various input data features at various abstraction levels. While the introductory layers learn and extract the high level features (with less abstraction), the deeper layers learn and extract the low level features (with higher abstraction). When compared to the CIE-recommended color-difference metrics, several metrics based on weighted CIELAB dl*, dC*, and dH* color-difference components performed much better.

4.WORKING OF CNN LAYERS

A) The convolutional layer receives the image's pixel input and performs the convolution operation. The outcome is a perplexing map.

B) A corrected feature map is produced by applying the convolved map with a ReLU (rectified linear unit) function.

C) To identify the features, the image is processed using several convolutions and ReLU layers.

D) The outcome is produced by flattening the pooled feature map and sending it to a fully connected layer.

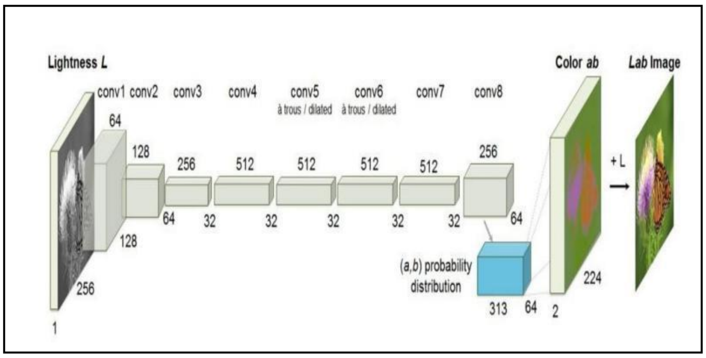


Figure 1: Architecture of neural network for colorization.

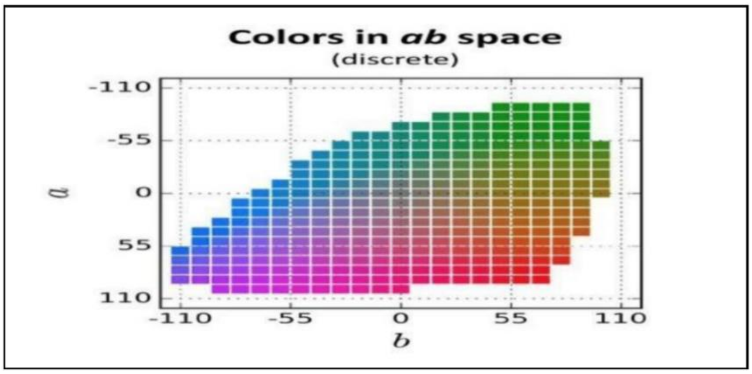


Figure 2: Quantized colors in ab space.

5. ALGORITHM FOR COLORIZATION

A) Using OpenCV, convert every training image from RGB to Lab colour space.

B) Use the L channel as the network's input, then train it to predict the combination of chrominance and luminance known as the ab channels. Convolution layer, pooling, flattening, and completely connected are the four fundamental CNN layers that we used to create the 8 layers of the caffemodel, which was pretrained using the imagenet dataset and prototxt file.

C) Use the fully connected layer to combine the predicted ab channels and the input L channel.

D) Use opencv to convert the Lab picture back to RGB.

6. RESULTS OBTAINED

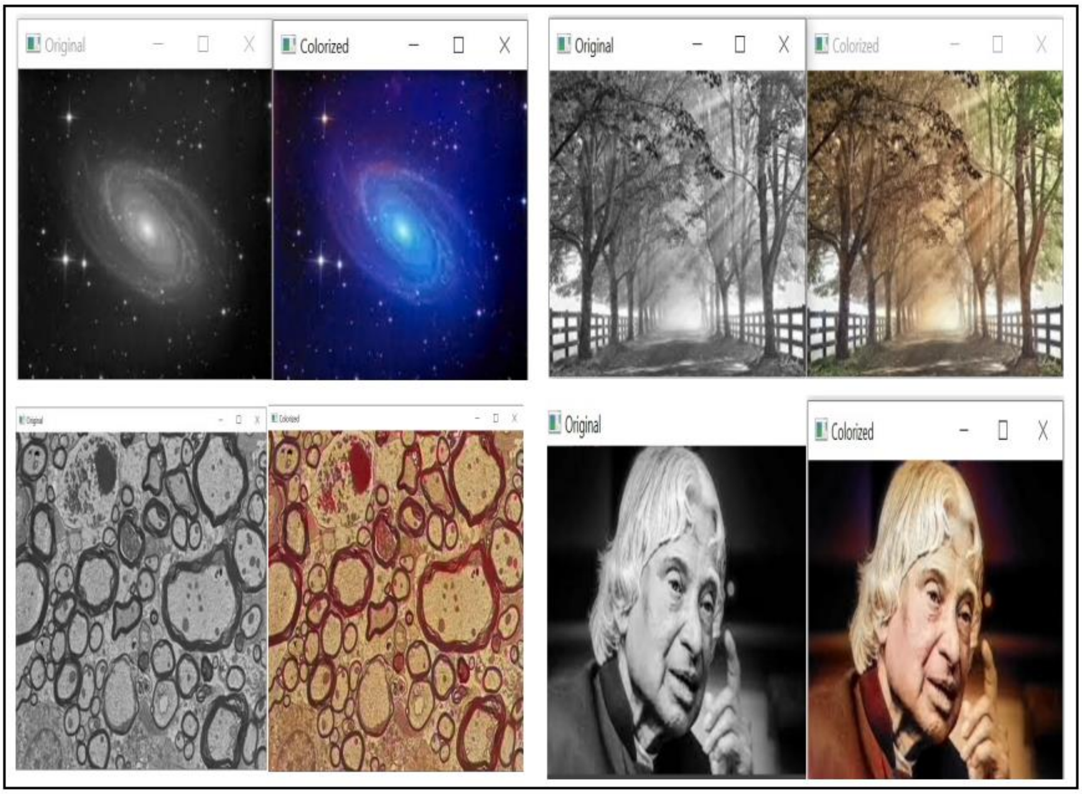
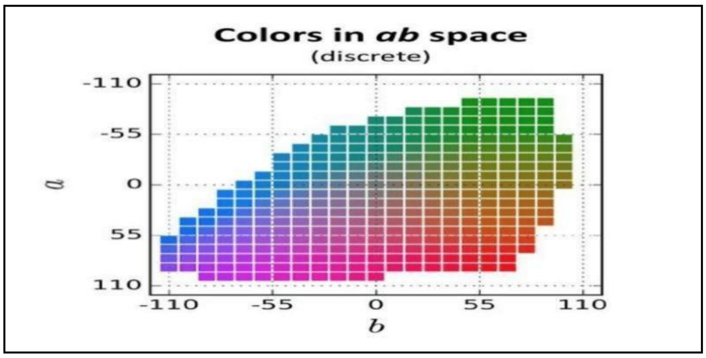
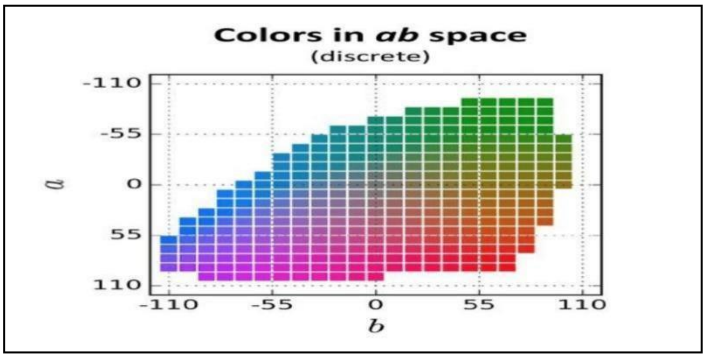


Figure 2: Results obtained on a) astronomical images b) Nature images c)Electron microscopy images d)Human images





7. CONCLUSION

It can be difficult for image processing and computer vision to colourize real images. The approach for multi-modal ill-posed uncertainty is present. If the spectator believed the coloured image to be real, the colorization was successful. This paper evaluates various topologies and levels of user guidance while taking into account the colorization time and criteria for objective image quality.

8. APPLICATIONS

Applications for the picture colorization systems include electron microscopy, CCTV footage, and astronomical photography. The various methods integrate user inputs with colour data from sizable data sets to create a model for accurate and effective colorization of greyscale photos.

9. FUTURE WORK

By carefully choosing training parameters like the number of layers, number of epochs, and learning rate to achieve a balance between training time and complexity of the neural network structure, existing deep learning colorization approaches can be improved. A greater understanding of the characteristics and context of images may result from further advancements in neural networks. The results can be enhanced by separating the extraction of local and global information. The common loss functions utilised in existing neural networks yield unsaturated hues, making selecting the right loss function in the colorization problem extremely difficult. Color, detail, perceptual information, and semantics are a few examples of variables that should be considered by more specialised loss functions.

References

[1] R. K. Gupta, A. Y.-S. Chia, D. Rajan, E. S. Ng, and Z. Huang, "Image colorization using similar images," in Proc. ACM Multimedia, 2012, pp. 369–378

[2] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 4, pp. 640–651, Apr. 2017.

[3] R. Zhang, P. Isola, and A. A. Efros, "Colorful image colorization," in Proc. ECCV, 2016, pp. 649–666.

[4] X. Liu et al., "Intrinsic colorization," ACM Trans. Graph., vol. 27, no. 5, pp. 152-1–152-9, 2008.

[5] G. Charpiat, M. Hofmann, and B. Schölkopf, "Automatic image colorization via multimodal predictions," in Proc. ECCV, 2008, pp. 126–139.

[6] A. Deshpande, J. Rock, and D. Forsyth, "Learning large-scale automatic image colorization," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 567–575.

[7] R. Dahl. (2016). Automatic Colorization. [Online]. Available: http://tinyclouds.org/colorize.

[8] G.Larsson, M. Maire, and G. Shakhnarovich, "Learning representations for automatic colorization," in Proc. ECCV, 2016, pp. 577–593.

[9] M. Richart, J. Visca, and J. Baliosian, “Image colorization with neural networks,” in Proc. Workshop Comput. Vis. (WVC), Oct. 2017, pp. 55–60.

[10] U. Qayynm, Q. Ahsan, Z. Mahmood, and M. A. Chcmdary, “Thermal colorization using deep neural network,” in Proc. 15th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST), Jan. 2018, pp. 325–329.