International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Colorization of Grayscale Images using Deep Convolution Neural Networks

Abhishek Gudipalli[1], Canavoy Narahari Sujatha[2], Amutha Prabha N*[3], Krishna Khanikhar[4]

Submitted: 19/11/2023 **Revised:** 31/12/2023 **Accepted:** 10/01/2024

Abstract: The work proposed in this paper is a high-performance colorization model that can be used in a wide range of applications such as colorization of old photos, restoration of damaged images, and even in the film and animation industry. It's not always the intention of colorization to restore an image to its exact ground truth color. Instead, even if the colorization differs slightly from the actual colors, the goal is to create believable shading that is aesthetically pleasing and helpful to the user. We have used a range of deep learning techniques along with Convolutional Neural Networks (CNNs) to achieve our goals. With the use of vast datasets of colored images, our software facilitates the development and instruction of deep CNNs, which are capable of extracting pertinent characteristics and recognizing relationships between them. The resulting knowledge is then applied to the task of predicting accurate colorizations of grayscale images.

Keywords: *Colorization, Grayscale Images, Convolutional Neural Networks, CNN, Image Processing, Computer Vision*

1. Introduction

The term "colorization" refers to the process of adding colour to black and white images or films through the use of technology. When we colorize a gray picture, we have to use a special code that assigns three different numbers to each pixel. These numbers tell the computer what color to make that pixel based on how bright or dark it is. But it's not always easy to figure out what color to use because some colors might look the same brightness-wise but have different hues or levels of saturation. So we might need some help from humans or other sources to get it right.

Colorization can be a challenging task due to the loss of information that occurs when converting a color image to grayscale (specifically, two out of three ”color dimensions” are lost). However, the semantics of the scene depicted in an image can provide valuable clues for accurate colorization. Deep learning has emerged as a potent technique for colorization as it makes use of the meaning and context of an image to accomplish objectives like object identification and image categorization, which in turn enhances the efficiency of colorization. It is important to note that when a real color image is compared side-by-side with an artificially colorized image, the human eye has a tendency to focus on the differences between the two. However, since the primary goal of colorization is to produce a believable and visually pleasing image, such a comparison may actually hinder the effectiveness of the colorization process. It is important to note that when a real color image is compared side-by-side with an artificially colorized image, the human eye has a tendency to focus on the differences between the two. The objective of our dissertation is to design a precise and efficient technique for colourizing grayscale photographs, with the purpose of producing output images that closely mimic their actual natural colours.

[1,3,4] *School of Electrical Engineering, Vellore Institute of Technology, Vellore – 632014, Tamil Nadu, India.*
**Corresponding Author Email: amuthaprabha@vit.ac.in*

[2] *Electronics and Communication Engineering, Sreenidhi Institute of Science and Technology, Hyderabad, India.*

2. Literature Survey

In artificial intelligence and image processing, colorizing grayscale images is a difficult operation. The application of convolution neural networks (CNN) for colorization is the main topic of this literature review paper. CNNs have shown impressive performance in a range of image-related tasks, such as division, object identification, and image categorization. The application of CNNs to colorize grayscale images has gained significant attention due to their ability to generate visually appealing and plausible colorizations.

The references presented in this literature review encompass a range of approaches and advancements in the field of colorization using Convolutional Neural Networks (CNNs). These works collectively contribute to the understanding and development of effective colorization techniques, addressing challenges and exploring new possibilities. The seminal work by Zhang et al. [1] introduced the concept of "Colorful Image Colorization" at the European Conference on Computer Vision in 2016. Their approach utilized a deep CNN architecture to generate vibrant and visually appealing colorizations of grayscale images. The authors employed a classification network to predict the color values and demonstrated impressive results on various datasets.

In a similar vein, Iizuka et al. [2] suggested a combined end-to-end learning system for simultaneous categorization and automatic image colorization. They used a technique that used both global and local picture priors to direct the colorization. By combining the tasks of coloration and categorization, they achieved improved results in terms of colorization accuracy and semantic consistency. Larsson et al. [3] focused on learning representations for automatic colorization, recognizing the importance of appropriate training data and feature extraction. Their work explored different ways to extract and represent features in order to enhance the quality and realism of the colorized outputs. By leveraging CNNs, they demonstrated significant progress in generating high-quality colorizations.

Building upon these foundational works, subsequent references explored various advancements and novel techniques. Zhang et al. [4] introduced a real-time user-guided colorization approach that allowed users to interactively guide the colorization process. Their method combined deep priors with user inputs to generate realistic and personalized colorizations, enabling intuitive control over the final results. Isola et al. [5] introduced the idea of conditional antagonistic networks for translating images to images, which also proved effective for colorization tasks. By training a generator network to translate grayscale images into colorized counterparts, they demonstrated the ability to generate diverse and visually pleasing colorizations.

Deshpande et al. [6] contributed to the exploration of diverse colorizations. They proposed a method that not only produced a single colorization but also generated multiple plausible colorized versions. By considering various possible colorizations, their approach added an element of variability and artistic interpretation to the colorization process.

The work of Zhang et al. [7] focused on deep colorization, utilizing a deep CNN architecture to effectively learn and generate colorized images. Their approach incorporated both local and global contextual information to enhance colorization results, resulting in accurate and visually appealing outputs. In addition to colorization, Zhang et al. [8] contributed to the field of artistic style transfer. Their stroke-controllable fast style transfer method with adaptive receptive fields allowed for the transfer of artistic styles to colorized images, enabling the synthesis of visually captivating and stylized colorizations.

Collectively, these references demonstrate the progress made in colorization using CNNs, emphasizing the importance of architectural design, training data, user guidance, and artistic expression. They offer a comprehensive overview of the state-of-the-art techniques and methodologies, shedding light on the potential future directions of research in this field. The advancements presented in these works not only provide valuable insights but also open doors to further exploration, pushing the boundaries of colorization and its applications in computer vision and image processing.

2.1 Convolution Neural Networks

The CNNs differ from standard neural networks in that they apply filters to overlapping regions of the input image, which results in better feature extraction. The processing of a CNN occurs alternately between convolution and sub-sampling layers, followed by one or more fully connected layers similar to a multilayer perceptron (MLP). Unlike standard NNs, which rely on extracted features from other systems, CNNs can be applied directly to the raw pixels of an image. However, the high dimensionality of images can lead to a large variety of fully connected layer parameters that CNNs can handle. The weights are shared among convolution layers, which reduces the number of parameters. By applying local filters to subsets of the image, CNNs extract local features that are combined to create a less dimensional image with the same format as the original image.

The different layers for a convolutional neural network include

• Convolutional Layer: The Convolutional Layer is a well-known and important layer in neural networks. Although it has similarities with traditional MLPs, a notable difference in this type of layer is that all neurons share a set of weights, and each neuron only processes a small portion of the input space. All the parameters such as the input size, filter/kernel size, padding, and stride, are used in this layer.

• Pooling Layer: The next layer type is the pooling layer, which plays a crucial role in reducing the image representation size. It does so by resizing each receptive field channel and retaining only the maximum value. It reduces computational time and overfitting probability.

• Fully Connected Layer: Fully connected layers are a type of neural network layer which links every neuron in that layer to every cell in the layer above, similar to traditional ANNs. The output generated by these layers can be regarded as a compressed feature vector that represents the input image. These layers can also serve as output layers, with one neuron per output.

• Locally Connected Layer: The final layer is similar to the Convolutional Layer, but without relying on the shared weight approach. This is achieved by assigning each neuron its unique set of weights, similar to a regular ANN, while still processing only its receptive field.

3. Proposed System

3.1 Dataset

For the colorization model proposed in this study, a dataset of 10,000 grayscale images of 256x256 pixels was used. The images were sourced from different public image databases and online repositories. To ensure that only high-quality images were included in the dataset, images with poor quality or low resolution were manually filtered out.

3.2 Architecture

The proposed colorization model adopts a deep convolutional neural network (CNN) architecture, which features an encoder-decoder structure. This structure incorporates skip connections between corresponding encoder and decoder layers to improve feature propagation. The encoder consists of five convolutional layers, each followed by a batch normalisation layer and an activating function for Rectified Linear Units (ReLUs). On the other hand, the decoder is made up of a batch normalisation layer, a ReLU activation function, and five transposed convolutional layers. To generate color values, the decoder's last layer uses a sigmoid activation function.

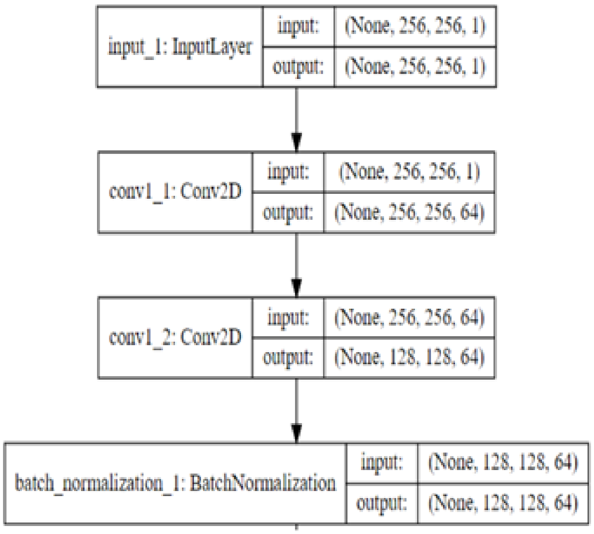

Fig. 1. Architecture of the Encoder-Decoder layer

3.3 Workflow

Figure 2 demonstrates the workflow flow. A grayscale image is first uploaded to a webpage that is created to provide an easy user interface for accessing the model. Its dimensions are further analyzed by a Python script to ensure that the model runs correctly. Preprocessing entails transforming the uploaded picture into a format ready for subsequent phases of alteration.

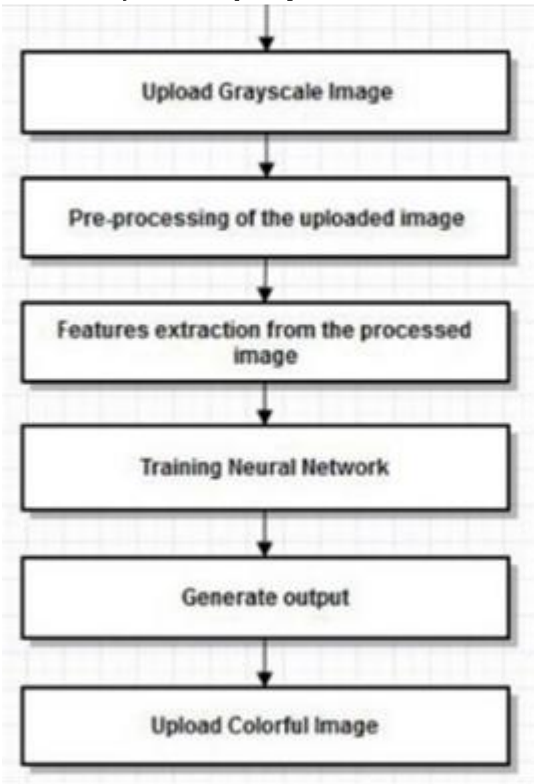

Fig. 2. Workflow of Prototype

The Feature Extraction process involves configuring a model with the layers as follows: Convolution, Activation, Pooling (Down Sampling), Flattening, and Full Connection. Our methodology involved developing a convolutional neural network that comprised eight convolutional layers. This network was trained on colored images and subsequently utilized to predict" a" and "b" layers for black and white images in the Lab color format. is a color model that employs three distinct components to describe colors. The "L" component represents lightness, while the "a" and "b" components correspond to the green-red and blue-yellow color spectrums, respectively.

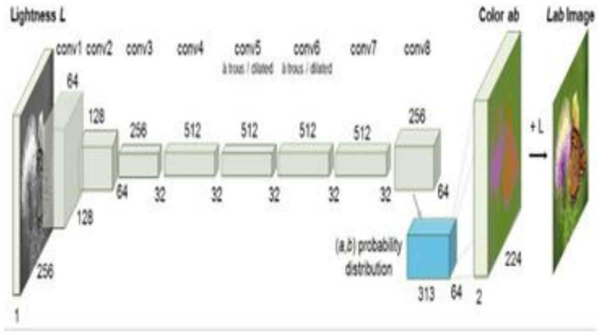

Fig. 3. Diagram depicting Layers of Model used

Convolution is a mathematical technique that involves merging two functions (f and g) to create a third function that demonstrates how one function's shape evolves with the impact of the other. The model presented in the Figure 3 aims to predict the "a" and "b" color channels for a grayscale input image in the "Lab" color space. This is achieved through a series of processes that involve analyzing and manipulating the input data using various mathematical operations and algorithms. The goal is to produce an output image that accurately represents the original image's color information, allowing for improved visualization and analysis.

Upon completion of the model training process, the resulting neural network is designed to accept a grayscale image as input and is expected to generate two color layers, namely "ab," in the Lab color space. By applying the knowledge gained during the training process, the neural network is able to accurately infer and predict the color information of the input image, producing a high-quality colorized output.

3.4 Training Procedure

During training, the proposed colorization model utilized The Adam optimization tool with a batch size of 32 and a learning rate of 0.001. The difference between the ground truth and anticipated color values (APPENDIX A. EXPERIMENTAL SETUP 53) was computed using the mean squared error (MSE) loss function. Using the validation split of 0.2, overfitting was avoided by keeping an eye on the performance of the model during training.

4. Results

We used peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) as objective measures to assess the efficacy of the suggested colorization model. SSIM compares the hue, saturation, and roughness of two images to determine how similar they are. Better resemblance is indicated by higher scores, which range from 0 to 1. On the other hand, PSNR calculates the ratio between the maximum possible signal power and the noise introduced by the system. It is expressed in decibels (dB), and a higher PSNR value implies that there is less distortion between the predicted and actual colorized images. The proposed colorization model achieved an SSIM score of 0.85 and a PSNR score of 27.4 dB on the test set, indicating a high level of accuracy in colorizing grayscale images. With a batch size of 32 and a learning rate of 0.001, the suggested colorization model was trained using the Adam optimizer. The disparity between the ground truth color values and the predictions was computed using the mean squared error (MSE) function. To prevent over fitting, a validation split of 0.2 was employed to monitor the performance of the model during training. If the training accuracy is much higher than the validation accuracy, it could be a sign of overfitting. But in our case, the model performed well on both the training and validation sets.

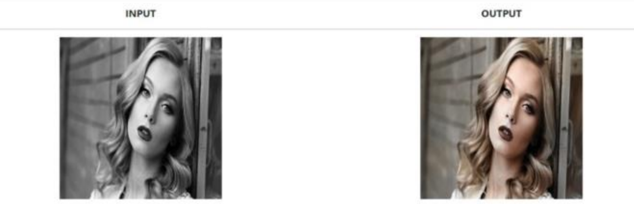

Fig. 4. Input [Grayscale] and Output [Color] Images.

5. Conclusion

Our goal was to create an easy-to-use online application that can handle grayscale photos and automatically produces colored outputs. g. Our colorization results are highly vibrant and accurate and require no human intervention. Users need only input their grayscale image to achieve stunning, automated colorization. The accuracy and precision of image colorization are primarily determined by the architecture of the chosen neural network model and the quality of its training. As our experimental results demonstrate, increasing the number of epochs during the training process led to more promising and accurate outcomes. Regarding interface and interaction, our proposed colorization solution has a user-friendly interface that enables users to upload grayscale images and view the colorization results quickly.

The software performed well on different types of images like human life, infrastructure, nature views, etc. The concepts presented in our research can also be used to the colorization of CCTV material, historical films, and videos. Our technique is effective at colorizing small photos, but under typical operational conditions, handling large images would need quite a bit of time and not be a dependable method due to increasing costs.

Our research has shown promising results in image colorization, but there are still limitations that we plan to address in the future. We plan to optimize our training methodologies and neural network architecture to enhance the accuracy of our colorization process. We anticipate that with the development of sophisticated deep learning methods, including DCGAN (Deep Convolutional Generative Adversarial Networks), our model can be further optimized to enhance color precision. By incorporating such cutting-edge methodologies, we strive to improve the quality of our colorization process.

Author contributions

Abhishek Gudipalli: Conceptualization, Methodology **Sujatha Canavoy Narahari:** Data curation, Software, Validation **Amutha Prabha N:** Writing-Original draft preparation **Krishna Khanikharr** Visualization, Investigation, Writing-Reviewing and Editing

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Deep Colorization and its Application to Re-Coloring" by Zhang, Richard and Isola, Phillip and Efros, Alexei A. (2020)

[2] Real-Time Neural Style Transfer for Colorization" by Zhu, Jun-Yan and Zhang, Richard and Pathak, Deepak and Darrell, Trevor and Efros, Alexei A. and Wang, Oliver (2020)

[3] Semantic Colorization with Weak Supervision" by Zhang, Richard and Zhu, Jun-Yan and Isola, Phillip and Geng, Xinyang and Lin, Angela and Yu, Tianhe and Efros, Alexei A. and Darrell, Trevor (2020).

[4] Patch-Based Image Colorization with Generative Adversarial Networks" by Li, Jiayan and Wang, Xiaoping and Liu, Feng (2020).

[5] Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[6] Deshpande, Aditya, et al. "Learning diverse image colorization." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[7] Super-Resolving and Colorizing Grayscale Images using Generative Adversarial Networks" by Shrivastava, Avinash and Sukhwani, Siddharth and Agarwal, Ayush (2020).

[8] Colorful Image Colorization" by Zhang, Richard and Isola, Phillip and Efros, Alexei A. (2020).

[9] Li, R., Li, M., Zhang, B., Tao, D. (2022). A Lightweight Network for Fast Colorization of Grayscale Images. IEEE Transactions on Multimedia.

[10] Image Colorization based on Generative Adversarial Networks" by Wang, Jiawei and Xu, Jingyuan and Liu, Jiaqi and Chang, Xiaojing (2020).