International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 5, pp 6-11, May-2024

https://doi.org/10.47001/IRJIET/2024.805002

Colorization of Black and White Images Using a Hybrid Deep Learning Framework

[1]Baraa Qasim Ibraheem, [2]Kassem Danach, [3]Ahmad Ghandour

[1,2,3]Department of Computer and Communications, Faculty of Engineering, Islamic University of Lebanon, Wardanieh, Lebanon

***Abstract* - With the development of deep learning algorithms and their great success in the field of computer vision, the field of automatic image colorization has witnessed significant improvements in accuracy and realism. This study introduces a novel deep learning-based method for colorizing black and white photographs, utilizing the powerful feature extraction of the InceptionResNetV2 model and the generative capabilities of autoencoders. A custom data generator was developed for efficient preprocessing, data augmentation, and batch processing, enhancing memory usage and scalability. The system encodes grayscale images and extracts high-level features, which are then fused and decoded into two color channels, combined with the original luminance to recreate the image in the LAB color space. The method demonstrates strong performance with a PSNR of 22.8154 and a SSIM of 0.9097, showcasing its potential for applications like historical image restoration and media enhancement.**

Keywords: Image Colorization, Deep Learning, Autoencoders, InceptionResNetV2, LAB Color Space.

I. INTRODUCTION

Colorization, the act of adding color to black and white photos, has been a subject of interest in the field of image processing. It is particularly relevant to historical photo restoration, media production, and more recently, machine learning applications [1]. Originally, the process of colorization was carried out manually by proficient painters, necessitating a significant amount of time and exertion to attain lifelike outcomes. The implementation of digital techniques enabled quicker data processing, although frequently led to less precise and artificial colorizations due to the constraints of early technology.

The development of colorization techniques has been strongly linked to progress in computer vision and machine learning [2]. Initially, computational methods relied on scripted algorithms and human input. Users would manually choose color values for gray areas, and the software would then apply these colors based on similarities in texture or luminosity. Although these systems were more efficient than hand-coloring, they were not sophisticated and frequently yielded inconsistent outcomes. The real success of automated colorization came with the advent of deep learning, which provided the tools to address the difficulty of colorization through a data-centric approach [3]. Neural networks have proven to be well-suited for analyzing grayscale images and identifying specific colors by extracting complex patterns and subtle features from large amounts of data.

While deep learning networks have transformed the capabilities of image colorization techniques, they require significant computational resource to train [4]. This can involve high-end GPUs and large amounts of memory, making the process time-consuming and expensive, especially when dealing with the large datasets needed for large-scale generalization. Moreover, attaining a significant level of authenticity in colorized photos frequently necessitates not just accurately reproducing colors but also maintaining the texture and intricacy that communicate the original ambiance and context of the image. To tackle these issues, continuous research and innovation in the areas of deep learning and image processing are necessary. The primary objective of the study is to create a sophisticated method for enhancing black and white photographs with color. This will be achieved by utilizing deep learning networks and transfer learning [5] techniques to handle the computational demands of training on extensive datasets. Additionally, the aim is to ensure that the resulting color images maintain a high level of structural similarity and overall quality.

II. RELATED WORK

Prior research in the domain of image colorization has laid the foundation for current approaches that leverage sophisticated machine learning methods. Semi-automatic solutions frequently necessitated manual intervention, as demonstrated by Levin et al.'s [6] groundbreaking work. They introduced an enhancement-based colorization technique that mandates users to manually position colored scribbles on a grayscale image. The algorithm will subsequently distribute these colors to the other parts of the image, taking into account the similarity between pixels. With the rise of deep learning in image processing, researchers began to explore the use of neural networks for colorization. A significant study conducted by Cheng et al. [7] employed a deep belief network

ISSN (online): 2581-3048
Volume 8, Issue 5, pp 6-11, May-2024
https://doi.org/10.47001/IRJIET/2024.805002

to acquire efficient representations for colorization. This study laid the groundwork for future investigations into neural network-based approaches.

Creating numerous color palettes was demonstrated in [8] through the use of a technique that was presented on the consistent LBP texture of the inputs. When the color palettes that have been established are combined, it is possible to color an image with a particular grayscale. The model is made up of two different types of convolutional neural networks: palette-based colorization networks and palette-based texture generating networks. The first fully automatic coloring method based on reference images was presented in [9]. The network structure is subnetworks of similarity and coloration. Similarity maps between reference images and target images are found through the similarity subnetwork. The chromaticity subnetwork then aligns the pixels of the luminance channel using the similarity subnetwork and enhances the colors of the misaligned pixels using big data learning. A strategy was proposed in [10] that can produce vibrant color outcomes by extracting the corresponding characteristics. In contrast to coloring based on reference photos, the generative adversarial network (GAN) encoder is specifically designed to generate color prior to the coloring process. This enables seamless blending between various hues and produces a wider range of outcomes.

Learning-based solutions have been increasingly popular in recent years for addressing the image colorization challenge. The implemented solutions mostly vary in terms of the network architectures employed and the loss functions utilized. One of the different convolutional neural networks (CNN) -based methods is the Cycle CNN, which is introduced in [11]. This model has the capability to directly utilize real data from monochrome camera systems throughout the training process. In addition, the model has been modified to reduce spatial irregularities and promote spatial smoothness in the coloring outcomes. The approach introduced in [12] utilizes CNN to produce authentic color pictures (RGB) from thermal infrared photos. The method employs an encoder-decoder architecture that incorporates skip links. Within an encoder-decoder framework, the input undergoes a sequence of layers that progressively decrease in size until it reaches the bottleneck layer, where the process is then reversed.

The effectiveness of generative adversarial networks and convolutional neural networks on automatic image colorization tasks was contrasted and assessed in [13]. Both can automatically color grayscale photos with a satisfactory level of visual quality. While being more computationally expensive, GAN performed significantly better than the CNN-based classification method. In most applications, the output of the last layer of CNNs is commonly utilized. However, in [14], a concept known as "supercolumn" is introduced, which draws inspiration from neuroscience.

This concept aims to leverage information from all levels of the network to create a fully automated image colorization system. Obtaining a large amount of actual data is not always possible for training intricate deep learning models. Hence, the generator network incorporates the pre-trained VGG19 model, which was trained on the extensive ImageNet dataset. A way to accomplish instance-aware coloring is suggested in [15]. An instance coloring network is employed to extract object-level characteristics, and an off-the-shelf object detector is utilized to acquire images of clipped items in the network architecture. The final color prediction was achieved by applying the fusion module to both the object-level and image-level features, which were extracted from the entire images using a similar network. Learned from a massive dataset are both fusion modules and colorization networks.

To better extract features of human photos from newly created datasets, U-net was enhanced in [16]. A wide range of realistic coloring effects can be achieved in real time by forecasting the probability distribution of potential colors and human participation. The deep neural network colorization algorithm got great results, but it requires a huge dataset for training, and it also needs to be able to colorize monochrome images in real time while maintaining diversity and originality. This calls for more research to make the algorithm more effective and generalizable.

III. METHODOLOGY

The LAB color space serves as an alternative to the RGB (red, green, blue) color space commonly used in the world of digital image processing. The LAB coloring space was specifically developed to carefully mimic the way humans see colors. There are three channels in LAB color Space: L channel (brightness), A channel (spatial relationship between the red and green colors), and the B (spatial relationship between the blue and yellow colors).

Figure 1 shows the decomposition of a color image (a) into two basic components: luminance (b) and chrominance (c).LAB is very useful in image colorization because it allows the L channel (which represents the black and white image) to be preserved while enabling adjustments only on the A and B channels to provide color. This separation enables models to focus on acquiring color data (A and B) without affecting brightness.

ISSN (online): 2581-3048
Volume 8, Issue 5, pp 6-11, May-2024
https://doi.org/10.47001/IRJIET/2024.805002

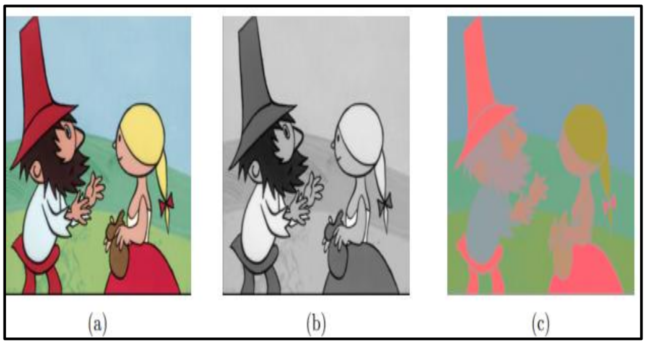

Figure 1: Decomposing a color image (a) into two basic components: luminance (b) and chrominance (c)

This section will present an approach that uses the power of GANs to create realistic colors, using the LAB color space to effectively separate luminance from color data, allowing for more focused and effective neural network training. This section will present a model that uses the power of deep learning networks to create realistic colors using the LAB color space. The model combines autoencoders [17] and uses the InceptionResNetV2 architecture for feature embedding. The approach begins with collecting images, processing them, then creating, training, and testing a coloring model.

3.1 Collect Dataset

The dataset was chosen from the “Image Coloring Dataset” available on Kaggle. 20 color images were selected for the training set and 100 grayscale images for the test set. The images were converted using the TensorFlow library into numpy arrays, a layout that allows the neural network to efficiently handle and parse visual data.Figure 2 shows examples of the various selected images.

Figure 2: Samples from the dataset

3.2 Preprocessing

Training set images are subjected to normalization by dividing the value of each pixel by 255, thus resampling the pixel values to a completely new range from 0 to 1 which helps in maintaining numerical balance and speeding up training. Data augmentation is performed dynamically during training.

3.3 Design Deep Learning Model

The architecture of the proposed system as shown in Figure 3 includes a dedicated encoder module for grayscale image processing, which works together with the pre-trained InceptionResNetV2 model to extract features. These features are then combined in a fusion layer, which combines detailed information captured by the autoencoder with high-level features extracted from InceptionResNetV2. The resulting global feature set is decoded into two color channels, which, when combined with the original luminance channel from grayscale images, reconstructs the image in the LAB color space.

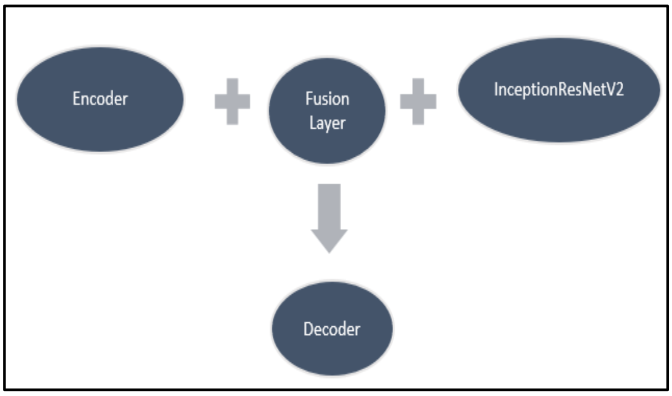

Figure 3: Components of the proposed colorization model

Figure 4 shows the detailed network structure, which takes two inputs: one for the encoder and one for the pre-trained model, InceptionResNetV2. The final output is a color image with the same spatial dimensions as the input image which, when combined with the original L channel, can be converted into a color image.

InceptionResNetV2 is a CNN architecture that combines the basic ideas of two important models, Inception and ResNet. Created by Google, InceptionResNetV2 is specifically designed to handle a range of computer vision applications such as image classification, object identification, and feature extraction. Images were resized to (299, 299) to fit InceptionResNetV2. Embeddings are obtained which are deep semantic details related to the content of images and are necessary for accurate color prediction.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 5, pp 6-11, May-2024

https://doi.org/10.47001/IRJIET/2024.805002

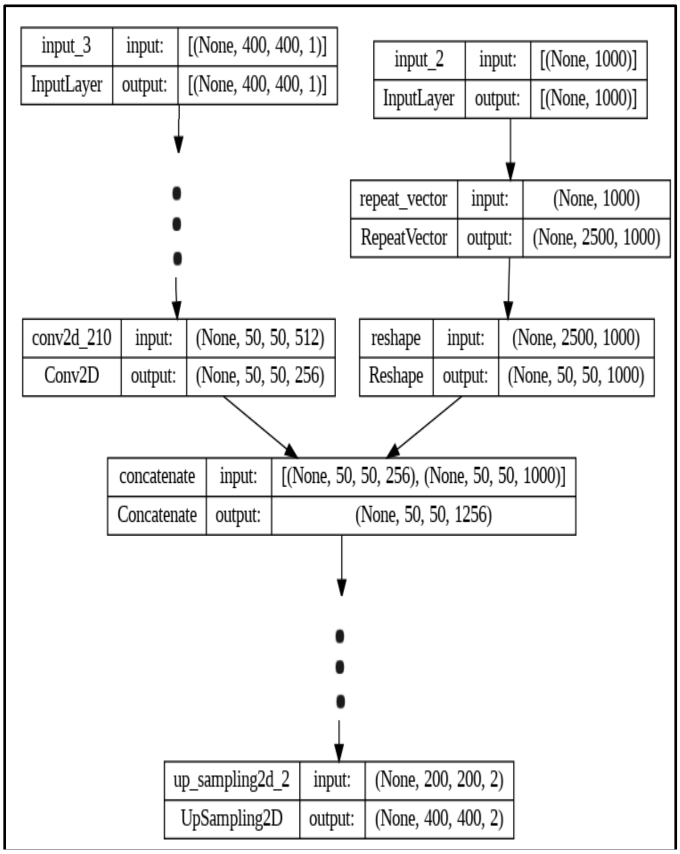

Figure 4: Architecture of image colorization network

The grayscale image in the encoder passes through a series of convolutional layers, reducing its dimension while increasing its depth. The initial layers gradually descend from 400 x 400 to 200 x 200, 100 x 100, and finally 50 x 50 pixels, with the channel depth increasing from 1 to 512.

The encoder outputs and InceptionResNetV2 feature embeddings are merged in the fusion layer to improve the feature set utilized by the decoder for color restoration. The image is progressively enlarged to its original size using a sequence of convolutional layers, while simultaneously reducing the depth from 256 to 2 (representing the A and B channels in the LAB color space).

3.4 Model Training

Proposed model was compiled by the “Adam” optimizer, with a mean square error (MSE) loss function. During training, the optimizer will change the model parameters to reduce the loss, thus enhancing the predictive function. A custom data generator was created to automatically handle the preprocessing steps and insert records into the training step in batches, which optimizes memory usage and speeds up the training process. Data generator augmented the data by making various adjustments such as cropping, zooming, rotating and flipping.

3.5 Model Testing

A total of 100 grayscale test photos were transformed into a NumPy matrix, which was subsequently normalized to the range of values between 0 and 1. Finally, the images were converted to RGB format. The InceptionResNetV2 model was employed to acquire feature embeddings. Subsequently, the data was transformed into the LAB color space, and the trained model was utilized to create predictions on the input photos.

IV. EVALUATION AND RESULTS

The model produced a stable spatial distribution of coloration as shown in figure 5, which represents one of the images colored using the model versus the original image, which indicates the good performance of the proposed model.

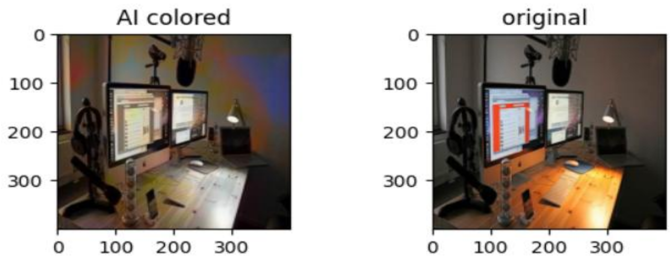

Figure 5: Sample of a color image by model against the original

After visual evaluation of the results, two common metrics were used to accurately evaluate the quality of the results:

- Peak signal-to-noise ratio (PSNR): Determines the quality of a color image by comparing it to the original image. Larger values indicate higher quality.
- Structural Similarity Index (SSIM) quantifies how similar color images are to the original image based on their structural properties, brightness, and contrast. Values in the range from -1 to 1 represent a range of similarity, where 1 indicates a state of complete similarity.

The model gave a mean PSNR of 22.815429563261745, and a mean SSIM of 0.9096590317983823. These values indicate that the proposed method works perfectly.

In [18], DenseUnet GAN architecture is presented for the purpose of image colorization which combines the features of Unet and DenseNet resulting in a complex neural network that effectively preserves details. Four state-of-the-art grayscale image colorization networks, Color CycleGAN, CycleGAN, AmetricCycleGAN, and Pix2pix GAN, are selected for comparison. In Table 1, the mean PSNR values and mean SSIM values of different coloration models are compared with our proposed model.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 5, pp 6-11, May-2024

https://doi.org/10.47001/IRJIET/2024.805002

Table 1: Comparing with other models

	Mean PSNR	Mean SSIM
CycleGAN	29.204	0.5664
Pix2pixGAN	31.2109	0.8000
Asymmetric CycleGAN	28.4802	0.4733
DenseUnet GAN	32.8144	0.8106
Our Model	22.815429563261745	0.9096590317983823

It is clear that our model has a lower mean PSNR value compared to other models, indicating good image quality and also outperforms all models in terms of mean SSIM value, indicating superior structural similarity with real images. This ensures that the proposed model effectively preserves the structural complexities and overall visual representation of color images.

V. CONCLUSION

The process of adding color to black and white photographs has been a difficult task in the field of image processing, especially due to the restrictions of existing approaches in generating realistic and precise color representations. This research presents a new technique based on deep learning that efficiently enhances grayscale photos by leveraging the capabilities of autoencoders and the InceptionResNetV2 model. This method utilizes the powerful feature extraction abilities of InceptionResNetV2 and the creative powers of autoencoders to develop a system that can generate high-quality colorized images. A specialized data generator was created to handle the tasks of pre-processing, data augmentation, and entering data into the model in batches for training. This helps improve memory usage and resource consumption and make the model generalizable for processing a huge dataset. The system design includes a grayscale image encoder and a pre-trained InceptionResNetV2 model to extract high-level features. These features are then combined in a fusion layer, creating a superset of features that is decoded into two color channels (A and B). These channels are then combined with the original L channel from the grayscale images, resulting in a reconstruction of the image in the LAB color space. The effectiveness of the proposed approach is demonstrated by its ability to generate realistic and structurally consistent colors, with the parameters obtained such as an average PSNR of 22.8154 and a SSIM of 0.9097. The results demonstrate the effectiveness of the deep learning framework in efficiently applying colors to monochrome images. This opens up possibilities for its use in historical image restoration, media enhancement, and other related applications.

REFERENCES

[1] R. Zhang, J. Y. Zhu, P. Isola, X. Geng, A. S. Lin, T. Yu, and A. A. Efros, “Real-time user-guided image colorization with learned deep priors,” *ACM Transactions on Graphics*, vol. 36, no. 4, 2017.

[2] H. Dalal, A. Dangle, M. J. Radhika, and S. Gore, “Image Colorization Progress: A Review of Deep Learning Techniques for Automation of Colorization,” *Int. J. Adv. Trends Comput. Sci. En*, vol. 10, no. 4, 2021.

[3] R. Dhir, M. Ashok, and S. Gite, “An overview of advances in image colorization using computer vision and deep learning techniques,” *Rev. Comput. Eng. Res*, vol. 7, no. 2, pp. 86-95, 2020.

[4] L. Jiao, and J. Zhao, “A Survey on the New Generation of Deep Learning in Image Processing,” *in IEEE Access*, vol. 7, pp. 172231-172263, 2019.

[5] K. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning,” *Journal of Big data*, vol. 3, pp. 1-40, 2016.

[6] A. Levin, D. Lischinski, and Y. Weiss, “Colorization using optimization,” *ACM SIGGRAPH 2004 Papers*, pp. 689-694, 2004.

[7] Z. Cheng, Q. Yang, and B. Sheng, “Deep colorization,” *Proceedings of the IEEE international conference on computer vision*, pp. 415-423, 2015.

[8] J. Li, H. Xiao, D. Tan, M. Zhang, and Y. Liu, “Image colorization based on texture by using of CNN,”*2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC)*, pp. 167-171, 2019.

[9] M. He, D. Chen, J. Liao, P. V. Sander, and L. Yuan, “Deep exemplar-based colorization,” *ACM Transactions on Graphics (TOG)*, vol. 37, no. 4, pp. 1-16, 2018.

[10] Y. Wu, X. Wang, Y. Li, H. Zhang, X. Zhao, and Y. Shan, “Towards vivid and diverse image colorization with generative color prior,” *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 14377-14386, 2021.

[11] X. Dong, W. Li, X. Wang, and Y. Wang, Y, “Cycle-CNN for colorization towards real monochrome-color camera systems,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 07, pp. 10721-10728, 2020.

[12] D. Tao, J. Shi, and F. Cheng, “Intelligent colorization for thermal infrared image based on CNN,”*2020 IEEE International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*, vol. 1, pp. 1184-1190, 2020.

[13] C. Qiu, H. Cao, Q. Ren, R. Li, and Y. Qiu, “Automatic Image Colorization with Convolutional Neural Networks and Generative Adversarial Networks,”

ISSN (online): 2581-3048
Volume 8, Issue 5, pp 6-11, May-2024
https://doi.org/10.47001/IRJIET/2024.805002

Automatic Image Colorization with Convolutional Neural Networks and Generative Adversarial Networks, 2020.

[14] L. Kiani, M. Saeed, and H. Nezamabadi-Pour, "Image colorization using a deep transfer learning,"*2020 8th Iranian Joint Congress on Fuzzy and intelligent systems (CFIS). IEEE*, pp. 027-032, 2020.

[15] J. W. Su, H. K. Chu, and J. B. Huang, "Instance-aware image colorization," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7968-7977, 2020.

[16] N. Wang, G. D. Chen, and Y. Tian, "Image colorization algorithm based on deep learning," *Symmetry*, vol. 14, no. 11, 2022.

[17] Q. V. Le, "A tutorial on deep learning part 2: Autoencoders, convolutional neural networks and recurrent neural networks," *Google Brain*, vol. 20, pp. 1-20, 2015.

[18] J. Xu, K. Lu, X. Shi, S. Qin, H. Wang, and J. Ma, "A DenseUnet generative adversarial network for near-infrared face image colorization," *Signal Processing*, vol. 183, 2021.

Citation of this Article:

Baraa Qasim Ibraheem, Kassem Danach, Ahmad Ghandour, "Colorization of Black and White Images Using a Hybrid Deep Learning Framework" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 8, Issue 5, pp 6-11, May 2024. Article DOI https://doi.org/10.47001/IRJIET/2024.805002

******