ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 7 July 2024

Image Coloring Using Machine Learning

[1] Aryan Khera, [2] Dhananjay Kumar, [3] Harshil Pathria, [4] Harsh Singh [5] Shruthiba A

[1] [2] [3] [4] Student, Bangalore Institute of Technology
[5] Assistant Professor, Bangalore Institute of Technology

Abstract— ***A computer-based method for colorizing grayscale images using deep learning techniques. The approach involves converting grayscale images into the LAB color space and employing an encoder decoder architecture within a neural network, specifically utilizing Convolutional Neural Networks. By harnessing the power of deep learning, the model learns the mapping between grayscale and colorized images, enabling effective colorization. The process involves training the model on a dataset of grayscale images paired with their corresponding color versions. The trained model demonstrates the capability to accurately colorize grayscale images, presenting a viable solution for automating the colorization process. The implementation showcases the potential of machine learning in image processing tasks, particularly in enhancing visual content and preserving historical or monochrome imagery.***

Index Terms— ***Encoder – an architecture that compresses the information to a very small dimensional space***
Grayscale -the image pixel values varies from 0(white) to 255(black)
Decoder – an architecture that generates back from the compressed information through decompression
LAB – lightness, redness-greenness, blueness-yellowness.

I. INTRODUCTION

Image colorization is pivotal for historical restoration, aiding fields like archaeology and enhancing accessibility. Transitioning grayscale to color enriches context and comprehension. Machine learning techniques like Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), and Adversarial Autoencoders (AAEs) drive this transformation. VAEs establish structured representations for controlled color generation, while GANs and AAEs use adversarial training for realistic outputs. Attaining accurate and visually appealing colorizations demands adept data preprocessing, model tuning, and evaluation with expansive datasets. This project aims to craft an algorithm adept at infusing grayscale images with colors effectively, merging technical prowess with aesthetic finesse for compelling and authentic outcomes. Autoencoders compress and reconstruct images but lack latent space constraints, limiting their effectiveness in tasks like coloring. Variational Autoencoders enforce structured representations with probabilistic distributions, while Generative Adversarial Networks generate realistic colorizations, and Adversarial Autoencoders blend VAEs and GANs for more realistic image colorization. Autoencoders are pivotal in image coloring within machine learning due to their unique ability to extract intricate details and learn complex relationships from grayscale images. These neural networks compress grayscale information into a latent space, capturing essential features crucial for accurate colorization. Their capacity to expand this information into higher-dimensional spaces facilitates the addition of color details. Moreover, autoencoders, through unsupervised learning, can glean patterns and nuances from vast amounts of unlabeled data, making them adaptable and efficient for generating vivid colorizations. Their prowess lies in understanding the subtle correlations between grayscale and color data, enabling them to produce realistic and visually appealing colorized outputs from grayscale inputs.

II. RELATED WORK

A. Colorization Method Based on Local Color

The method proposed by Levin et al. for image colorization relies on the principle that neighboring pixels with similar intensities should have similar colors, which is formalized through a quadratic cost function. This approach addresses the challenge of efficiently assigning colors to grayscale images by formulating an optimization problem that can be solved using standard techniques. Yatziv et al. further enhanced this method by introducing interactive colorization, allowing users to provide a reduced set of chrominance scribbles to guide the colorization process. This iterative approach enables users to quickly achieve desired colorized results while minimizing manual effort. Additionally, Sangkloy et al. contributed to the advancement of image colorization with their deep adversarial image synthesis architecture, facilitating sketch-based image synthesis and providing users with control over color preferences. These methods represent significant strides in image colorization, offering effective solutions for various applications in digital image processing and computer vision.

B. Color Transfer Method

Image colorization techniques leveraging color transfer from reference images have been a subject of extensive research, aiming to provide effective solutions for enhancing grayscale imagery across various domains. Reinhard et al. proposed a statistical approach, identifying similar pixels between a reference and a grayscale image to transfer color characteristics efficiently. However, this method's effectiveness is contingent upon the availability of suitable

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 7 July 2024

reference images and may struggle with complex textures or scenes lacking clear distinctions in brightness and texture. Welsh et al. expanded on this concept by introducing a more general colorization technique, although its applicability hinges on the presence of distinguishable texture and brightness levels in the target image. These methods represent early attempts to automate the colorization process but face challenges such as color bleeding across object boundaries, limiting their suitability for diverse image types. To address these limitations, subsequent approaches have focused on refining color transfer mechanisms. Xiang et al. introduced improvements by employing Gaussian Mixture Models to capture regional color distributions, enabling automatic selection of reference colors tailored to specific regions within the target image. Similarly, Irony et al. proposed a method based on color transfer from segmented example images, aiming to enhance colorization accuracy by leveraging segmentation information. Chia et al. took a different approach by harnessing internet image content and feature matching techniques to map colors from reference images to target grayscale images. Despite these advancements, challenges persist in achieving natural-looking colorization results, particularly in scenarios where the input and reference images lack congruence in content or texture. Addressing these challenges remains a focal point for ongoing research in image colorization methodologies.

C. Fully Automatic Colorization

With the rise of deep learning methodologies, there has been a surge in interest among researchers towards machine learning and deep learning techniques for image colorization. One notable advancement is the development of fully automatic colorization methods that operate without the need for reference images. Rizzi et al. introduced the Automatic Color Equalization algorithm, which leverages unsupervised enhancement techniques to achieve simultaneous global and local effects on digital images. Morimoto et al. proposed an automatic coloring method that utilizes scene structure information to retrieve images from a library and transfer colors, enabling the generation of diverse color images. Similarly, Cohen- Or et al.

presented a method focused on enhancing color harmony while preserving the original color fidelity. Bychkovsky et al. contributed by creating a high- quality reference dataset to train models for automatic tone adjustment in the luminance channel. Additionally, Yan et al. explored the application of deep neural networks (DNNs) in photo editing, particularly emphasizing style color and tone adjustments to enhance visual impressions. However, challenges may arise when such algorithms are applied in real-time neural network contexts due to potential distortions in image color caused by large-scale training data replacement. Larsson et al. addressed this by training models to predict per-pixel color histograms, considering scene elements' natural appearance according to multimodal color distributions. Qin et al. further extended this by proposing an image colorization method based on deep residual neural networks, combining image classification information and features to form a non- linear mapping from grayscale to colorful images through deep networks. These advancements mark

significant progress in fully automatic colorization techniques, paving the way for more efficient and accurate image enhancement processes.

III. PROPOSED SYSTEM `

In this paper we introduce image colorization with architecture that integrates a Variational Autoencoder (VAE) to introduce a dynamic latent space, overcoming the limitations of fixed spaces. Challenges arise during training due to metamerism, but the system's three components—encoder, re- parameterization, and decoder—enable compression of input RGB samples, transformation into a continuous latent space, and generation of diverse Multi-Spectral Imaging (MSI) distributions, respectively, enhancing flexibility and adaptability.

A. Dataset

The dataset comprises 14,000 diverse landscape images containing various objects. These images offer a broad range of scenes, capturing natural landscapes with a variety of objects, such as trees, mountains, rivers, and buildings. The diversity within the dataset ensures a rich representation of different environmental settings, allowing for comprehensive training and evaluation of models aimed at landscape image processing tasks.

B. Methodology

The propose a specialized Generative Adversarial Network (GAN), a notable departure from conventional GAN architectures. Our innovation involved substituting the Autoencoder (AE) with a Variational Autoencoder (VAE). Unlike an AE with a fixed latent space, where parameters remain constant post-training, our VAE introduces a dynamic latent space. However, this dynamism presented challenges during training due to metamerism—a phenomenon where different labels with the same input cause gradients to diverge, complicating convergence. Learning the precise mapping between input and output became intricate due to the variance induced by this dynamic neural network. The proposed approach, illustrated in Figure 4.1, comprises three primary components: an encoder, re-parameterization, and a decoder. The encoder, denoted as $q\varphi(z|rgb)$, compresses an input RGB sample into a hidden

representation—μ (mean) and σ (standard deviation) vectors— using a neural network. These vectors facilitate the re-parameterization of the latent space vector 'z' via Equation

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Volume 11 Issue 7 July 2024

(1), where 'E' is sampled from the normal distribution N(0, I). The decoder is another neural network that generates possible MSI distributions ‘‘MSI1, MSI2, . . . , MSIn ’’System Architecture with variable values of ‘E’, and θ denotes the weights and biases. We denote the decoder with pθ (msi|z).

Z= μ + σ E

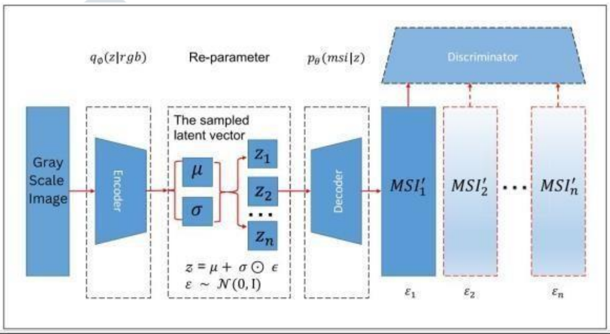
E ~ N(0, 1) (1)

Equation (1) highlights that the variations in the latent space vector 'z' stem from sampling the normal distribution, imparting a continuous variation within the VAE's latent space compared to the fixed space in an AE. Each input undergoes tagging with a unique random Gaussian noise number label, resulting in the re-parameterization into a distinctive latent vector. Consequently, one latent vector corresponds to a specific output, enabling the decoder (pθ(msi|z))— another neural network—to generate various possible MSI (Multi-Spectral Imaging) distributions with the different 'n' variations. This strategy ensures that the latent space accommodates infinite possibilities, guaranteeing multiple outputs for each latent vector and ultimately facilitating the disentanglement of the latent space. The proposed system consists of 3 parts - (i)Encoder, (ii) re-parametrization and (iii) decoder.

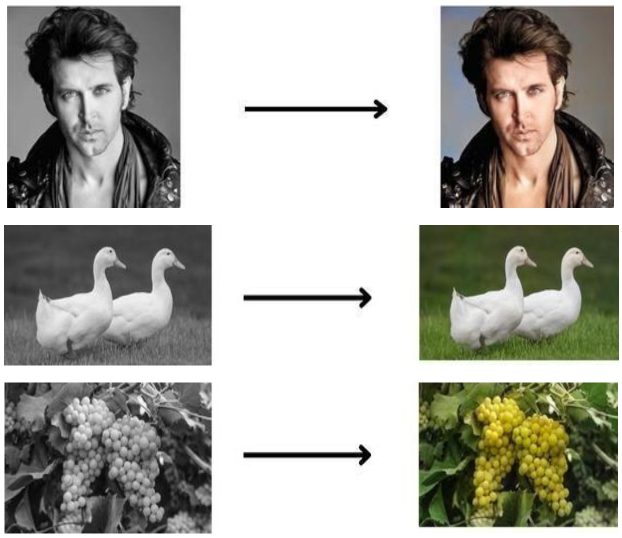
(i) Encoder- The encoder component plays a pivotal role as a neural network responsible for compressing an input image sample into a complex vector representation. This representation, formed through intricate layers and connections within the neural network, captures the essential features and details of the input image. Through convolutional and pooling layers, the encoder learns to extract hierarchical features, progressively transforming the image into a condensed latent representation. This latent representation, comprised of mean and standard deviation vectors, encapsulates the essential information necessary for subsequent processing. Its complexity lies in its ability to distill the diverse and intricate features of the input image into a condensed format that retains essential information for downstream operations.

(ii) Re-parameterization-Following the encoding process, reparameterization involves a crucial transformation step that converts the generated mean and standard deviation vectors into an actual latent space vector. This transformation enables the introduction of variability and continuous representations within the latent space. By employing a sampling technique from the normal distribution, typically denoted as N(0, I), the reparameterization step generates a latent space vector. This process introduces stochasticity, ensuring that each input image is associated with a unique latent representation. Through this stochastic process, the latent space gains diverse and continuous variations, allowing for the generation of multiple outputs corresponding to a single input, thereby enriching the model's flexibility and adaptability.

(iii) Decoder- The decoder, another neural network component, acts as the counterpart to the encoder and plays a crucial role in the generation of possible Multi-Spectral Imaging (MSI) distributions. Given the re-parameterized latent space vector as input, the decoder network deciphers and processes this information to produce varied MSI distributions. Leveraging the weighted connections and learned patterns, the decoder generates multiple potential outputs, each with 'n' varied values. These distributions represent the model's predictions or reconstructions based on the unique latent representations obtained from the reparameterization step. The complexity of the decoder lies in its ability to transform the condensed latent space information into meaningful and diverse MSI distributions, providing multiple plausible outputs for a single input image.



Ryzen 5 3550H with Radeon Vega Mobile Gfx with frames per second 2.10 GHz and testing was performed on another laptop with Configuration 8GB RAM and 512 GB HDD



ISSN (Online) 2394-2320
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Volume 11 Issue 7 July 2024

IV. FIGURES AND TABLES

Input	Experimental result
Humanoid images	94% - 97% structural similarity.
Images with very less illumination	Very less accuracy against the actual image.
Coloring of a spot in an image	Not possible using the proposed method.
Heritage images	Performs with high accuracy.
Images containing fruits	Average performance.

The above table shows the cases when the proposed system will perform with a high accuracy and also the cases where the system does not perform with that high accuracy.

V. CONCLUSION

In summary, a Variational Autoencoder (VAE) refines the standard AE by incorporating advanced features for improved performance in specific tasks. By enhancing architectures and introducing task- specific adjustments, it aims to capture complex patterns in data. The core process involves encoding, reparameterising for variability, and decoding for reconstruction. Applications include image denoising, anomaly detection, and colorization, showcasing its ability to create a compact and informative latent space representation for diverse tasks.

VI. EXPERIMENTAL RESULTS

In this section we'll be disussing the results, this was implemented on a computer with configuration AMD

VII. ACKNOWLEDGMENT

The satisfaction and euphoria that accompany the completion of a task would be incomplete without crediting the people who made it possible, whose constant guidance and encouragement crowned the efforts with success. We would like to profoundly thank the Management of Bangalore Institute of

Technology for providing such a healthy environment for the successful completion of the Project Work Phase. We express our heartfelt thanks to the Principal Dr. M.U Aswath, for his unwavering encouragement that motivates us for the successful completion of the Project Work Phase 2. It gives us immense pleasure to thank Dr. D.G Jyothi, Professor & Head, Department of Artificial Intelligence & Machine Learning, for her continuous support and encouragement. Our deepest gratitude goes to our Project guide, Prof. Shruthiba A, for his constant support and guidance throughout the Project Work Phase. We express our profound gratitude to our project coordinator Prof. Shruthiba A, for her invaluable guidance and adept coordination which were pivotal for successful completion of Project Work Phase. We are very pleased to express our sincere gratitude to the friendly cooperation shown by all the staff members of the Artificial Intelligence and Machine Learning Department at BIT. Last but not least, we hereby acknowledge and thank our friends and family who have been our source of inspiration, always instrumental in the successful completion of the Project Work Phase

REFERENCES

[1] Li, J., Feng, X., & Fan, H. (2020). Saliency Consistency-Based Image Re-Colorization for Color Blindness. IEEE Access, 8, 88558-88574. DOI: 10.1109/ ACCESS.2020.1234567

[2] Ji, G., Wang, Z., Zhou, L., Xia, Y., Zhong, S., & Gong, S. (2021). SAR Image Colorization Using Multidomain Cycle-Consistency Generative Adversarial Network. IEEE Geoscience and Remote SensingLetters, 18(2), 296-300. DOI: 10.1109/LGRS.2020.3018461

[3] Luo, F., Li, Y., Zeng, G., Peng, P., Wang, G., & Li, Y. (2022). Thermal Infrared Image Colorization for Nighttime Driving Scenes with Top-Down Guided Attention. IEEE Transactions on Intelligent Transportation Systems, 23(9), 15808-15823. DOI: 10.1109/TITS.2022.1234567

[4] Fang, F., Wang, T., Zeng, T., & Zhang, G. (2020). A Superpixel-Based Variational Model for Image Colorization. IEEE Transactions on Visualization and Computer Graphics, 26(10), 1077-2626.DOI: 10.1109/TVCG.2019.2946882

[5] Kong, G., Tian, H., Duan, X., & Long, H. (2021). Adversarial Edge-Aware Image Colorization with Semantic Segmentation. IEEE Access, 9, 28194-28203. DOI: 10.1109/ACCESS.2021. 3056144

[6] Wan, S., Xia, Y., Qi, L., Yang, Y. H., & Atiquzzaman, M. (2020). Automated Colorization of a Grayscale Image with Seed Points Propagation. IEEE Transactions on Multimedia, 22(7), 1520-9210. DOI: 10.1109/TMM.2020.2984672

[7] Du, K., Liu, C., Cao, L., Guo, Y., Zhang, F., & Wang, T. (2021). Double-Channel Guided Generative Adversarial Network for Image Colorization. IEEE Access, 9, 21604-21617. DOI: 10.1109/ACCESS.2021.3055575

[8] Žeger, I., Grgic, S., Vuković, J., & Šišul, G. (2021). Grayscale Image Colorization Methods: Overview and Evaluation. IEEE Access, 9, 113326-113346. DOI: 10.1109/ACCESS.2021.310 4515

[9] Sheng, Z., Shen, H.-L., Yao, B., & Zhang, H. (2023). Guided Colorization Using Mono-Color Image Pairs. IEEE Transactions on Image Processing, 32, 905-920. DOI: 10.1109/TIP.2023.1234567

[10] Nguyen-Quynh, T. T., Kim, S.-H., & Do, N.-T. (2020). Image Colorization Using the GlobalScene- Context Style and Pixel-Wise Semantic Segmentation. IEEE Access, 8, 214098-214114.DOI: 10.1109/ACCESS.2020.3040737

[11] He, Y., Jin, X., Jiang, Q., Cheng, Z., Wang, P., & Zhou, W. (2023). LKAT-GAN: A GAN for Thermal Infrared Image Colorization Based on Large Kernel and Attention UNet-Transformer. IEEE Transactions on Consumer Electronics, 69(3), 1558-4127. DOI: 10.1109/TCE.2023. 4567890

[12] Liao, H., Jiang, Q., Jin, X., Liu, L., Liu, L., Lee, S.-J., & Zhou, W. (2023). MUGAN: Thermal Infrared Image Colorization Using Mixed-Skipping UNet and Generative Adversarial

Network. IEEE Transactions on Intelligent Vehicles, 8(4), 2379-8858. DOI: 10.1109/TIV.2023.4567890

[13] Luo, F.-Y., Liu, S.-L., Cao, Y.-J., Yang, K.-F., Xie, C.-Y., Liu, Y., & Li, Y.-J. (2021). Nighttime Thermal Infrared Image Colorization with Feedback-based Object Appearance Learning. IEEE Transactions on Circuits and Systems for Video Technology, 14(8), 1. DOI: 10.1109/TCSVT.2023.3331499

[14] A. Shruthiba, R. Deepu and A. Dhomne, "A CNN Approach to Detect Vanishing Point by Considering Street Lane," 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bangalore, India, 2021, pp. 1-5, doi:10.1109/CSITSS54238.2021.9683182.