RAJALAKSHMI ENGINEERING COLLEGE
An AUTONOMOUS Institution
Affiliated to ANNA UNIVERSITY, Chennai

COLOURIZEVISION : BLACK AND WHITE IMAGE TO COLOUR IMAGE USING CNN

A Project Report

Submitted by

SARAN P (221501122)
DHEERAJ KUMAR S (221501503)

AI19441 FUNDAMENTALS OF DEEP LEARNING

Department of Artificial Intelligence and Machine Learning

RAJALAKSHMI ENGINEERING COLLEGE,THANDALAM.

RAJALAKSHMI
ENGINEERING COLLEGE

BONAFIDE CERTIFICATE

NAME ………………………………………………………………………………..……..…

ACADEMIC YEAR………………..………SEMESTER…………BRANCH………………

UNIVERSITY REGISTER No.

Certified that this is the bonafide record of work done by the above students in the Mini Project titled **"COLOURIZEVISION : Colour Your Black And White Image Using CNN** the subject **AI19541 – FUNDAMENTALS OF DEEP LEARNING** during the year **2024 – 2025.**

Signature of Faculty – in – Charge

Submitted for the Practical Examination held on ____________________

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ABSTRACT

Image colorization is a complex and visually creative process that involves assigning realistic and context-aware colours to grayscale images. This project leverages Convolutional Neural Networks (CNNs) to automate this task, utilizing deep learning techniques to predict chrominance values in the LAB colour space. The approach is based on training a model with diverse datasets of colour images, enabling it to learn the relationships between luminance and chrominance channels. By focusing on the spatial and semantic context within images, the model can accurately reconstruct coloured outputs from grayscale inputs, producing visually compelling and contextually plausible results.

The proposed system demonstrates the power of CNNs in handling pixel-level prediction tasks while considering global image coherence. The model can predict colours that align with real-world visual patterns, such as assigning green hues to grass and blue to skies. This capability not only simplifies the traditionally manual and labour-intensive process of colorization but also opens new opportunities for applications in photography restoration, video enhancement, and creative design. Future enhancements aim to incorporate larger datasets and advanced architectures, such as attention mechanisms, to improve the richness and accuracy of the colorized outputs, paving the way for more sophisticated and user-adaptive systems.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Image colorization, the process of adding realistic colours to grayscale images, is a challenging task that requires understanding both low-level pixel information and high-level semantic contexts. Traditionally, this has been a manual and time-consuming process, often relying on the skill of artists and photo editors to assign accurate and context-aware colours. With advancements in deep learning, particularly Convolutional Neural Networks (CNNs), automating the image colorization process has become a viable solution. CNNs excel in extracting features from images, making them ideal for predicting complex relationships between grayscale luminance and the missing colour information.

This project employs a deep learning model based on a CNN architecture to automate the colorization of black-and-white images. The system uses OpenCV for image pre-processing and visualization, while the LAB colour space is leveraged for separating luminance ('L') and chrominance ('a' and 'b') channels. The model is trained to predict the 'a' and 'b' chrominance channels from the grayscale 'L' channel, effectively restoring the full-colour image. The approach not only captures local textures and details but also considers global semantic cues to ensure that colours are contextually plausible. By utilizing pretrained models and custom adjustments, such as loading specific kernel points for the chrominance layers, the system delivers visually compelling results with minimal computational complexity. This project demonstrates how deep learning can simplify and enhance the image colorization process, offering applications in photography restoration, video enhancement, and digital content creation.

CHAPTER 2
LITERATURE REVIEW

[1] Colorful Image Colorization by Zhang et al. (2016).

This work introduced one of the first deep learning-based frameworks for fully automated image colorization. The authors employed a CNN to predict chrominance ('a' and 'b') channels in the LAB colour space, given the luminance ('L') channel as input. The network was trained on large-scale datasets, allowing it to generalize across diverse scenes and objects. To handle the inherent ambiguity in colorization (e.g., a Gray apple could be green or red), the authors incorporated a probabilistic loss function, which enabled the model to predict plausible colours even in uncertain scenarios. This study highlighted the potential of CNNs for colorization tasks and set a benchmark for future research.

[2] Image-to-Image Translation using GANs.

Generative Adversarial Networks (GANs) have been applied to colorization tasks to improve the realism and diversity of the generated colours. Works such as "Pix2Pix" introduced adversarial training, where a generator network learns to produce colorized images, and a discriminator network evaluates their quality. Although GAN-based models produce vivid and lifelike colours, they often require extensive computational resources and careful tuning to ensure stable training.

[3] Deep Priors and Self-Supervised Learning.

Some studies have explored the use of deep priors and self-supervised learning for colorization. These methods leverage the spatial and structural consistency of images to train models without requiring extensive labeled datasets. By learning to reconstruct missing color information from grayscale inputs, these models demonstrate robustness in handling unseen image distributions.

[4] Transformers for Long-Range Dependencies.

While CNNs are effective at capturing local features, transformer-based architectures have recently been applied to image colorization for capturing global context. Transformers leverage self-attention mechanisms to model relationships between distant regions in an image, which is particularly useful for assigning consistent colors across large objects or backgrounds.

[5] Domain-Specific Colorization Approaches.

Several studies have focused on domain-specific applications of colorization, such as restoring historical photographs, colorizing medical images, or enhancing satellite imagery. These approaches often involve fine-tuning pre-trained models on domain-specific datasets to improve performance and reliability in specialized tasks.

CHAPTER 3

SYSTEM REQUIREMENTS

3.1 HARDWARE REQUIREMENTS:

- CPU: Intel Core i5/Ryzen 5 or higher
- GPU: NVIDIA GTX 1050 or higher (optional for faster training)
- RAM: Minimum 8 GB

3.2 SOFTWARE REQUIRED:

- OS: Windows/Linux/MacOS
- Framework: OpenCV, NumPy, Python 3.8 or above
- Libraries: TensorFlow/Keras, OpenCV, NumPy

CHAPTER 4

SYSTEM OVERVIEW

4.1 EXISTING SYSTEM

Existing image colorization systems rely heavily on manual techniques or semi-automated rule-based methods, which are labor-intensive, time-consuming, and lack the ability to generalize across diverse image types. Early approaches used statistical correlations or user-provided hints to propagate colors, but they often failed to understand the semantic context of the image, resulting in unnatural or inconsistent outputs. While modern deep learning methods like GANs and RNNs have improved the automation and quality of colorization, they come with challenges such as high computational demands, unstable training, and limited generalization to unseen data. These limitations underscore the need for a more robust, scalable, and fully automated approach, like the CNN-based system proposed in this project, which balances efficiency, contextual awareness, and output realism.

4.2 PROPOSED SYSTEM

The proposed system is a real-time image and video colorization platform that leverages a pre-trained deep learning model to transform black-and-white images or video streams into fully colorized outputs. Using OpenCV and a Caffe-based model, the system processes grayscale inputs by extracting the luminance (L) channel, generating the chrominance (AB) channels, and combining them to produce a colorized image. The system supports both static images, where users can upload and instantly view colorized results, and dynamic video streams, where each frame is colorized in real-time. The back-end processing leverages GPU acceleration for fast inference, while the front-end offers a simple UI for user interaction. This tool has applications in areas such as historical photo restoration, creative projects, media, and education, with the potential for future enhancements like interactive color adjustments and AR integration.

4.2.1 SYSTEM ARCHITECTURE

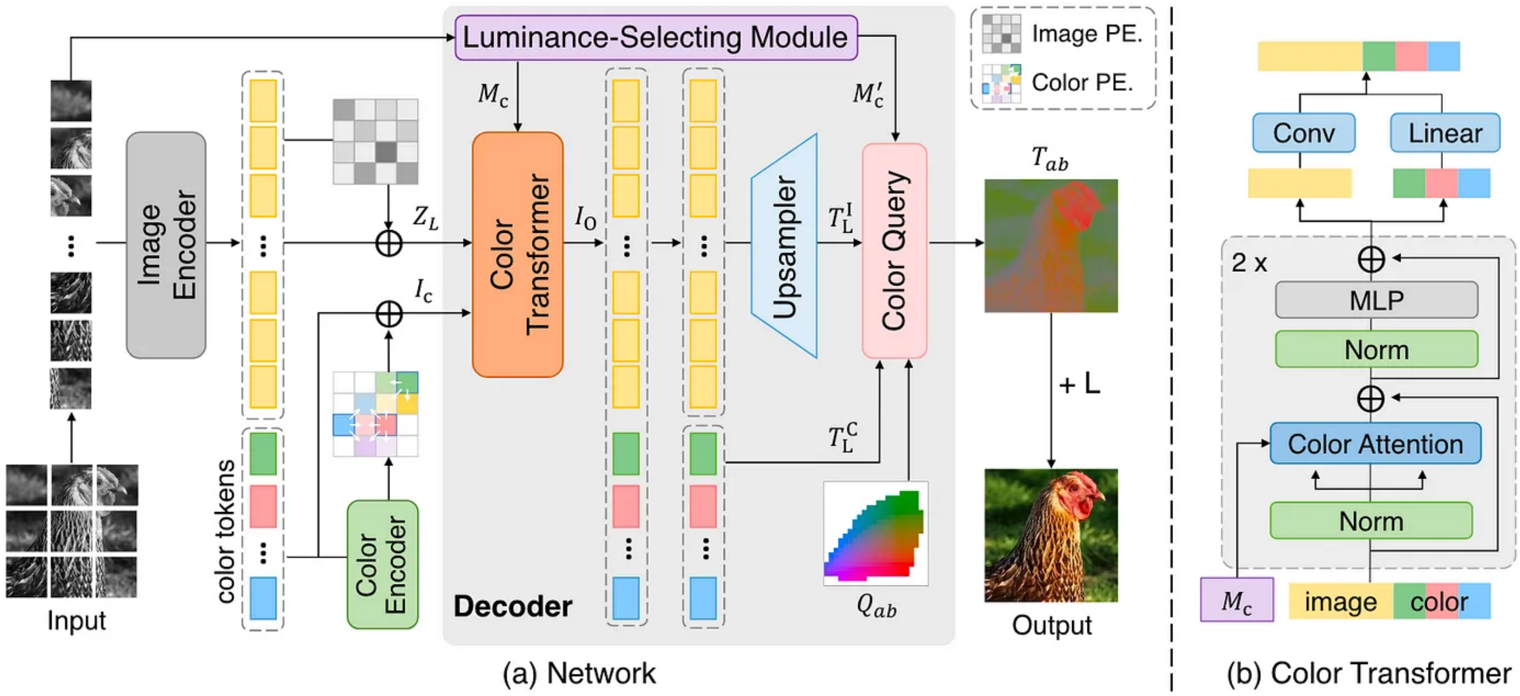

Fig 1.1 Overall diagram of b&w to colour

4.2.2 DESCRIPTION

The project involves developing a real-time image and video colorization system that utilizes deep learning to automatically convert black-and-white images and video streams into colorized versions. By using a pre-trained Caffe-based deep learning model, the system processes grayscale images by separating luminance (L) and chrominance (AB) channels, then combines them to produce colorized outputs. The system is capable of processing both static images and video streams, offering real-time colorization with GPU acceleration. The project aims to be useful for applications such as historical photo restoration, creative design, media enhancement, and educational tools, with potential future improvements including interactive color controls and augmented reality integration.

CHAPTER-5

IMPLEMENTATION

5.1 LIST OF MODULES

- Image Preprocessing Module
- Model Development and Training
- Colorization Processing Module
- GPU Acceleration and Performance Optimization Module
- Evaluation and Analysis

2. MODULE DESCRIPTION

1. Image Preprocessing Module: This module converts grayscale images to the LAB color space and extracts the luminance (L) channel for model input. It normalizes the image to fit the model's expected range. The preprocessing ensures proper formatting for accurate colorization.

2. Model Development and Training: The module focuses on creating and training the deep learning model using a large dataset of paired grayscale and colorized images. The model learns the relationship between luminance and chrominance channels to predict realistic colors. During training, hyperparameters and network architecture are optimized for better performance. Once trained, the model is fine-tuned for accuracy and generalization across diverse images.

3. Colorization Processing Module: This module takes the luminance (L) channel and applies the trained model to generate the chrominance (AB) channels. These channels are combined to form a complete LAB image, which is then converted to the BGR color space for display. The result is a colorized image that reflects the model's learned understanding of color relationships. The output can be used for both static images and video frames.

4. GPU Acceleration and Performance Optimization Module: Leveraging GPU acceleration, this module speeds up the processing of images and video, enabling real-time colorization. CUDA and other optimization techniques are used to offload heavy computations to the GPU, improving the system's overall performance. This ensures low latency and efficient handling of large datasets, especially important for video colorization. The module guarantees that both images and video can be processed smoothly.

5. Evaluation and Analysis: The module evaluates colorization quality using objective metrics such as PSNR and SSIM to ensure accuracy and realism. User feedback is gathered to assess how natural the colorized outputs appear in different contexts. Performance analysis helps identify potential areas for model improvement, such as handling edge cases or enhancing color accuracy. These insights drive further model refinement and system optimization.

5.1.1 ALGORITHMS

1. **Data Preparation**: Convert black and white images to LAB color space, using the L channel as input and a, b channels as target color information.

2. **Model Architecture**: Use a CNN trained on large datasets to predict color channels (a and b) from the grayscale L channel input.

3. **Image Preprocessing**: Normalize the input image, convert it to LAB color space, and resize it to 224x224 pixels for the model.

4. **Colorization (Forward Pass)**: Feed the L channel into the CNN to predict the a and b color channels, and resize the output to match the original image.

5. **Post-Processing**: Merge L, a, and b channels, convert to BGR, and show the colorized image.

CHAPTER-6

RESULT AND DISCUSSION

The colorization process using a pre-trained CNN model effectively transforms black-and-white images into plausible colorized versions by predicting the a and b channels, which represent color information, from the grayscale L channel in the LAB color space. The results demonstrate that the model produces visually appealing and realistic colorization, with the colors generated being consistent with the grayscale intensity values. However, the quality of the colorized image can vary, especially in more complex scenes or detailed textures, as the model's performance is influenced by the diversity and quality of its training dataset. In some cases, the model may introduce slight inaccuracies, particularly in regions where the available training data may be limited or lacking. Despite these limitations, the approach provides a robust solution for image colorization, and future improvements, such as fine-tuning the model on more diverse datasets or employing advanced techniques like attention mechanisms, could help further enhance the accuracy and realism of the colorized results.

REFERENCES

[1] **R. Zhang, P. Isola, and A. A. Efros**, "Colorful Image Colorization," *European Conference on Computer Vision (ECCV),* 2016, pp. 649-666. Available: https://arxiv.org/abs/1603.08511

[2] **I. Goodfellow et al.,** "Generative Adversarial Nets," *Advances in Neural Information Processing Systems (NeurIPS),* vol. 27, 2014, pp. 2672-2680. Available: https://papers.nips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf

[3] **Z. Cheng, Q. Yang, and B. Sheng**, "Deep Colorization," *International Conference on Computer Vision (ICCV),* 2015, pp. 415-423. Available: https://arxiv.org/abs/1605.00075

[4] **G. Larsson, M. Maire, and G. Shakhnarovich**, "Learning Representations for Automatic Colorization," *European Conference on Computer Vision (ECCV),* 2016, pp. 577-593. Available: https://arxiv.org/abs/1603.06668

[5] **M. Noroozi and P. Favaro**, "Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles," *European Conference on Computer Vision (ECCV),* 2016, pp. 69-84. Available: https://arxiv.org/abs/1603.09246

APPENDIX

SAMPLE CODE

```
import numpy as np
import cv2
from skimage.metrics import structural_similarity as ssim
import matplotlib.pyplot as plt

prototxt_path = 'models/colorization_deploy_v2.prototxt'
model_path = 'models/colorization_release_v2.caffemodel'
kernel_path = 'models/pts_in_hull.npy'
image_path = 'jii_image.jpg'  # Black-and-white image

net = cv2.dnn.readNetFromCaffe(prototxt_path, model_path)

points = np.load(kernel_path)
points = points.transpose().reshape(2, 313, 1, 1)
net.getLayer(net.getLayerId("class8_ab")).blobs = [points.astype(np.float32)]
net.getLayer(net.getLayerId("conv8_313_rh")).blobs = [np.full([1, 313], 2.606, dtype="float32")]

bw_image = cv2.imread(image_path)
if bw_image is None:
    raise ValueError(f"Unable to load image at path: {image_path}")

normalized = bw_image.astype("float32") / 255.0
lab = cv2.cvtColor(normalized, cv2.COLOR_BGR2LAB)
```

```
resized = cv2.resize(lab, (224, 224))
L = cv2.split(resized)[0]
L -= 50

net.setInput(cv2.dnn.blobFromImage(L))
ab = net.forward()[0, :, :, :].transpose((1, 2, 0))
ab = cv2.resize(ab, (bw_image.shape[1], bw_image.shape[0]))

L = cv2.split(lab)[0]
colorized = np.concatenate((L[:, :, np.newaxis], ab), axis=2)
colorized = cv2.cvtColor(colorized, cv2.COLOR_LAB2BGR)
colorized = (255.0 * colorized).astype("uint8")

plt.figure(figsize=(10, 5))
plt.subplot(1, 2, 1)
plt.title("Black and White Image")
plt.imshow(cv2.cvtColor(bw_image, cv2.COLOR_BGR2RGB))
plt.axis("off")

plt.subplot(1, 2, 2)
plt.title("Colorized Image")
plt.imshow(cv2.cvtColor(colorized, cv2.COLOR_BGR2RGB))
plt.axis("off")
plt.show()
```

```
bw_gray = cv2.cvtColor(bw_image, cv2.COLOR_BGR2GRAY)
colorized_gray = cv2.cvtColor(colorized, cv2.COLOR_BGR2GRAY)

print(f"Black-and-White Image Shape: {bw_gray.shape}")
print(f"Colorized Image Shape: {colorized_gray.shape}")

ssim_score = ssim(bw_gray, colorized_gray)

print(f"SSIM (Structural Similarity Index) Accuracy: {ssim_score:.4f}", flush=True)
```

OUTPUT SCREENSHOT

Fig 5.1 Output for black and white to colour image

Colour Your Black And White Image Using Convolution Neural Network

Saran P
dept. Artificial Intelligence and Machine Learning
Rajalakshmi Engineering College
Chennai, India
heysaran29@gmail.com

Sangeetha K
dept. Artificial Intelligence and Machine Learning
Rajalakshmi Engineering College
Chennai, India
sangeetha.k@rajalakshmi.edu.in

Dheeraj Kumar S
dept. Artificial Intelligence and Machine Learning
Rajalakshmi Engineering College
Chennai, India
dheerajtrl007@gmail.com

Abstract—This paper presents a method for black-and-white image colorization using Convolutional Neural Networks (CNNs), which excel at capturing spatial features and contextual relationships. The model is designed to operate in the LAB color space, where the luminance (L-channel) serves as input, and the chrominance (a and b channels) are predicted to re-colorizationstic and natural-looking colorized images. Unlike traditional heuristic-based techniques, this CNN-based approach effectively learns local textures and global structures, ensuring contextually accurate colorization while preserving fine details. Experimental results highlight the model's ability to handle various image types, including portraits and landscapes, producing smooth and visually appealing outputs. Future enhancements include expanding datasets and incorporating attention mechanisms to improve the model’s ability to generate more diverse, expressive, and nuanced colorizations suitable for applications like digital restoration and creative editing.

Keywords—Image Colorization, Convolutional Neural Networks (CNNs), LAB Color Space, Grayscale to Color Conversion, Chrominance Prediction, Luminance, Spatial Features, Contextual Learning, Deep Learning, Digital Restoration, Artistic Re-coloring, Visual Enhancement, Computer Vision, Neural Networks, Automatic Image Processing.

I.INTRODUCTION

Image colorization, traditionally a manual and highly skill-dependent process, has been revolutionized by advancements in deep learning, particularly through Convolutional Neural Networks (CNNs). This project automates the task of restoring realistic colors to black-and-white images by employing a CNN-based architecture that effectively learns the relationship between grayscale luminance ('L') and chrominance ('a' and 'b') channels in the LAB color space. Using OpenCV for image pre-processing and visualization, the model predicts the 'a' and 'b' channels while retaining the luminance channel, producing contextually accurate and visually compelling results. By leveraging pretrained models and fine-tuned kernel weights for enhanced precision, the system captures both local textures and global semantic cues. This efficient and scalable approach has wide-ranging applications, including photo restoration, video enhancement, and content creation, showcasing how AI can transform and simplify complex creative tasks.

II. RELATED WORK

Image colorization has evolved from manual methods, requiring user input like scribbles, to advanced deep learning approaches that automate the process.

(2015), focused on leveraging local image textures and segmentations, while Zhang et al. (2016) introduced a framework that predicts chrominance channels ('a' and 'b') in the LAB color space, incorporating global semantic context for realistic results. GAN-based methods, like Pix2Pix and CycleGAN, further advanced the field by employing a generator-discriminator architecture to create highly lifelike colorizations, ensuring both accuracy and diversity. Recent innovations include self-supervised and zero-shot learning techniques, which reduce reliance on large paired datasets by learning directly from grayscale images. Hybrid models combine traditional manual input with neural networks, enabling user-guided adjustments for finer control. Beyond static images, extensions like video colorization address challenges such as temporal consistency between frames, opening new applications in fields like photo restoration, video enhancement, and digital content creation, demonstrating the broad potential of automated colorization systems.

III. PROBLEM STATEMENT

The primary challenge in image colorization is transforming grayscale images into realistic, full-color representations while maintaining semantic accuracy and visual appeal. Traditional manual methods are time-consuming, skill-dependent, and lack scalability, while fully automated solutions must handle diverse contexts, ambiguous color assignments, and varying image textures. The problem lies in designing a system that can accurately predict the missing chrominance information for grayscale images while ensuring the results are contextually plausible, computationally efficient, and applicable across diverse scenarios. This project addresses these challenges by leveraging deep learning techniques, particularly Convolutional Neural Networks (CNNs), to automate the colorization process, aiming to deliver realistic results with minimal user intervention and broad applicability in areas like image restoration, video enhancement, and digital content creation.

IV. SYSTEM ARCHITECTURE AND DESIGN

The system architecture for image colorization begins with an input module that accepts grayscale images and converts them into the LAB color space, separating the luminance ('L') channel from the chrominance ('a' and 'b') channels. A pre-processing module normalizes and augments the data before feeding it into a CNN-based deep learning model. The model includes an encoder to extract features from the grayscale input, a decoder to predict the 'a' and 'b' channels, and optional skip connections to preserve details. Post-processing combines the predicted chrominance channels with the 'L' channel to reconstruct the full-color image, which is then converted back to RGB. The training process involves paired grayscale and color datasets, with loss functions like MSE, perceptual loss, and optionally adversarial loss for enhanced realism. Outputs are visualized or saved, with scalability to video frames for dynamic applications. This design ensures semantic accuracy, computational efficiency, and high-quality results.

V. PROPOSED METHODOLOGY

The proposed methodology for black-and-white image colorization begins by preprocessing the input grayscale image to make it suitable for deep learning. The image is normalized to scale pixel values to the range [0, 1] and converted to the LAB color space, which separates luminance (L) and chrominance (a and b) components. The luminance channel (L) is extracted and resized to match the input size expected by the pre-trained Convolutional Neural Network (CNN). This channel is passed through the CNN, which has been trained on large datasets to predict the chrominance channels (a and b). The predicted a and b channels are then resized to the original image dimensions and combined with the original L channel to reconstruct a complete LAB image. The LAB image is converted back to the BGR color space, producing a visually plausible. The LAB image is converted back to the BGR color space, producing a visually plausible colorized version of the input image. The resulting image is displayed and can be saved for various applications, such as historical photograph restoration, artistic enhancement, and media processing. This methodology ensures accurate and natural-looking colorization by leveraging a specialized deep learning model optimized for this task.

VI. IMPLEMENTATION AND RESULTS

The black-and-white image colorization is implemented using a pre-trained Convolutional Neural Network (CNN) designed for this purpose. The input grayscale image is normalized and converted to the LAB color space to extract the luminance (L) channel. This L channel is resized to 224x224 pixels and passed through the pre-trained model, which predicts the a (green-red) and b (blue-yellow) chrominance channels. The model uses learned priors and kernel points to generate realistic color distributions.

The predicted chrominance channels are resized to match the original image dimensions and combined with the original L channel to reconstruct a LAB image. Finally, this reconstructed LAB image is converted back to the BGR color space to produce the final colorized image. The implementation is executed using OpenCV for image processing and the OpenCV deep learning module (cv2.dnn) for model inference.

The colorized output demonstrates visually plausible and natural colors, effectively enhancing the grayscale input image. For example, historical photographs or grayscale inputs are restored with colors that align well with real-world expectations, such as blue skies, green vegetation, and skin tones. While the model is not perfect in all scenarios—especially when encountering ambiguous or unfamiliar objects—it delivers impressive results for general use cases, such as artistic enhancements or media processing. The results are displayed alongside the original grayscale image for comparison, and the generated images can be saved for further use. This approach proves to be efficient and versatile, showcasing the capability of deep learning in image restoration and enhancement.

VII. CONCLUSION AND FUTURE WORK

The black-and-white image colorization using a pre-trained Convolutional Neural Network (CNN) demonstrates a highly effective approach for restoring colors to grayscale images. By leveraging the LAB color space and the model's learned ability to predict chrominance channels from luminance data, the method produces visually realistic and vibrant results. The implementation is efficient, requiring minimal computational resources while delivering outputs suitable for practical applications such as restoring historical photographs, enhancing grayscale media, and creative art generation. While the results are impressive for general use cases, the approach can sometimes struggle with complex or unfamiliar scenes, where accurate color prediction may require contextual understanding beyond pixel patterns.

Future work can focus on fine-tuning the model using domain-specific datasets to improve accuracy in specialized applications, such as medical imaging or underwater photography. Incorporating semantic segmentation or user-guided inputs can enhance contextual understanding for more accurate colorization. Exploring advanced architectures like Generative Adversarial Networks (GANs) or transformers may further improve the realism and diversity of the colorized outputs. Additionally, optimizing the system for real-time performance and deployment on mobile or embedded devices can broaden its accessibility and practical usage.

REFERENCES

[1] R. Zhang, P. Isola, and A. A. Efros, "Colorful Image Colorization," *European Conference on Computer Vision (ECCV),* 2016, pp. 649-666.

[2] I. Goodfellow et al., "Generative Adversarial Nets," *Advances in Neural Information Processing Systems (NeurIPS),* vol. 27, 2014, pp. 2672-2680.

[3] Z. Cheng, Q. Yang, and B. Sheng, "Deep Colorization," *International Conference on Computer Vision (ICCV),* 2015, pp. 415-423..

[4] G. Larsson, M. Maire, and G. Shakhnarovich, "Learning Representations for Automatic *Colorization*," *European Conference on Computer Vision (ECCV),* 2016, pp. 577-593. Available: https://arxiv.org/abs/1603.06668

[5] M. Noroozi and P. Favaro, "Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles," *European Conference on Computer Vision (ECCV),* 2016, pp. 69-84.