**VISVESVARAYA TECHNOLOGICAL UNIVERSITY**

Jnana Sangama, Belagavi - 590 018, Karnataka



**Conversion of Greyscale to Color using Deep Learning**

**And Classification of Images**

*A Report submitted in partial fulfillment of the requirements for the Course*

**Mini Project**

**(Course Code:** **24AM5PWMPW)**

*In the Department of*

**Machine Learning**

**(UG Program: B.E. in Artificial Intelligence and Machine Learning)**

By

**Mohd Ehteshaam Khan**

USN: **1BM22AI078**

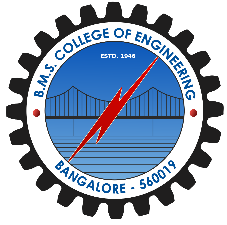
Semester & Section: 5 B

Under the Guidance of

**Prof. Lavanya B.Koppal**

Assistant Professor

Dept. of MEL, BMSCE, Bengaluru – 19



**Department of Machine Learning**

**B.M.S College of Engineering**

(An Autonomous Institute, Affiliated to VTU)

P.O. Box No. 1908, Bull Temple Road, Bengaluru - 560 019

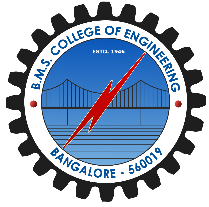
**December - 2024**

**B.M.S College of Engineering**

(An Autonomous Institute, Affiliated to VTU)

P.O. Box No. 1908, Bull Temple Road, Bengaluru - 560 019

**Department of Machine Learning**



**CERTIFICATE**

This is to certify that **Mr. Mohd Ehteshaam Khan** bearing USNs: **1BM22AI078** has satisfactorily presented the Course – Mini with Project  **(Course code: 24AM5PWMPW)** with the title **“Conversion of Greyscale to Color using Deep Learning And Classification of Images.”** in partial fulfillment of academic curriculum requirements of the 5th semester UG Program – B. E. in Artificial Intelligence and Machine Learning in the Department of Machine Learning, BMSCE, an Autonomous Institute, affiliated to Visvesvaraya Technological University, Belagavi during December 2024. It is also stated that the base work & materials considered for completion of the said course is used only for academic purpose and not used in its original form anywhere for award of any degree.

**Student Signature**

**Signature of the Supervisor Signature of the Head**

**Prof. Lavanya Koppal Dr. M Dakshayini**

Assistant Professor, Dept. of MEL, BMSCE Prof. & Head, Dept. of MEL, BMSCE

**External Examination**

**Examiner Name and Signature**

**1.**

**2.**

**ABSTRACT**

Recent advancements in deep learning have transformed the landscape of image processing, enabling new solutions to complex challenges like converting greyscale images to color. This project explores the use of convolutional neural networks (CNNs) to predict realistic RGB values for pixels in black-and-white images, a task that demands a deep understanding of the spatial and contextual relationships present in the image. By colorizing historical images, we not only enhance their visual appeal but also make them more relatable, bridging the gap between past and present. The colorization process utilizes color spaces such as YUV or CIELAB, which separate luminance and chrominance for more accurate color predictions. Additionally, the project integrates image classification, using cutting-edge CNN models, to develop a comprehensive pipeline that merges both tasks. This integrated approach has significant potential for applications in historical restoration, media remastering, and other areas requiring improved image interpretation and visualization.

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| Chapter No. | Chapter Name | Page No. |
| 1 | Introduction | 5 |
| 2 | Literature Review | 6-7 |
| 3 | Open Issues | 8 |
| 4 | Problem Statement | 9 |
| 5 | Proposed Architecture | 10-11 |
| 6 | Functional & Non- functional Requirements | 12 |
| 7 | Low-Level Design | 13-14 |
| 8 | Methodology | 15-16 |
| 9 | Implementation | 17 |
| 10 | Experimental Results and Analysis | 18-19 |
| 11 | Testing and Validation | 20 |
| 12 | Conclusion | 21 |
| 13 | References | 22 |

**CHAPTER 01: INTRODUCTION**

Deep learning has transformed the field of image processing, offering groundbreaking solutions to long-standing challenges. One such challenge is the task of converting greyscale images into color, which involves predicting appropriate RGB values for every pixel. This process demands significant computational power and a deep understanding of the spatial and contextual connections within the image.

Colorizing black-and-white images not only restores their artistic and cultural significance but also provides a more relatable and immersive representation of historical moments. This technique bridges the gap between the past and the present, making historical scenes more vivid and accessible. From a technical perspective, the colorization process assigns RGB values to pixels by using color spaces such as YUV or CIELAB, which separate luminance from chrominance. Advanced deep learning techniques, especially convolutional neural networks (CNNs), have automated this process, achieving highly realistic results.

Alongside colorization, this report also explores image classification, a fundamental task in modern image analysis. By utilizing advanced CNN architectures, the project merges these two processes into a single, unified pipeline. This integration holds the potential to improve applications in areas like historical restoration, media remastering, and other related fields.

.

**CHAPTER 02: LITERATURE REVIEW**

The literature review examines recent developments in deep learning for image colorization and classification, emphasizing important techniques, results, and challenges.

**2.1 Greyscale to Color Conversion**

Numerous research efforts have investigated deep learning approaches for image colorization, making substantial contributions to the field. Notable contributions include::

* **Ivana Žeger, Sonja Grgić, Josip Vuković, and Gordan Šišul (2021)**:
  + Discussed traditional methods and modern approaches, including edge-based techniques, CNNs, and GANs for colorization.
  + CNNs It was observed that CNNs and GANs surpassed traditional techniques, providing enhanced color precision and visual quality.Limitations: High computational costs and the need for extensive datasets.
* **Zhang, R., Isola, P., & Efros, A. A. (2016)**:
  + Introduced a fully convolutional network (FCN) for pixel-wise color prediction.
  + Achieved high-quality, context-aware colorizations, particularly effective for structured images.
  + Limitations: Challenges in handling low-resolution or noisy images.
* **Yu Chen, Yeyun Luo, Youdong Ding, and Bing Yu (2019)**:
  + Developed CNN-based colorization for Chinese black-and-white films.
  + Achieved culturally accurate and high-quality colorizations.
  + Limitations: Limited generalization outside specific datasets.
* **Mingming He, Dongdong Chen, Jing Liao, Pedro V. Sander, and Lu Yuan (2018)**:
  + Combined deep learning with exemplar-based techniques for guided colorization.
  + Generated realistic and contextually appropriate outcomes by transferring color from reference images.
  + Challenges: High computational cost and dependence on extensive exemplar datasets.

**CHAPTER 03: OPEN ISSUES**

**3.1 Colorization Challenges**

* Challenges in achieving accurate and realistic colorization, particularly for historical or unclear images.
* Significant computational expenses involved in producing high-quality colorized outputs.

**3.2 Generalization and Dataset Diversity**

* Limited generalization to diverse datasets with varied textures, styles, and objects.
* Requirement of large labeled datasets for effective training.

**3.3 Robustness and Degradation Handling**

* Inability to maintain consistent performance when processing degraded images, such as those with noise or blur.

**3.4 Classification Limitations**

* Challenges in maintaining accuracy when classifying colorized images.
* Dependence on pre-trained models, which may not generalize well to colorized outputs.

**3.5 Interpretability and Evaluation**

* Challenges in assessing colorization with objective metrics that align with human perception.
* Requirement for intuitive interfaces to adjust colorization results based on specific needs.

**CHAPTER 04: PROBLEM STATEMENT**

Current approaches to greyscale-to-color image conversion and classification encounter several significant challenges, including:

1. **Achieving realistic and believable colorization:** Many existing methods have difficulty generating accurate colors, particularly for images that are ambiguous or complex.
2. **Classification of Colorized Outputs**: Maintaining high accuracy for classifying images after colorization is difficult.
3. **Handling Diverse and Degraded Datasets**: Models often fail to generalize well to datasets with varying textures, styles, or degraded inputs.
4. **Efficiency and Scalability**: Most pipelines are computationally expensive and lack real-time processing capabilities.

This project aims to develop a deep learning pipeline that integrates greyscale-to-color conversion with image classification. The objectives are to:

* Accurately predict realistic colors for greyscale images using convolutional neural networks.
* Classify the resultant colorized images with high precision.
* Optimize computational efficiency to enable scalability and real-time processing in practical applications.

**CHAPTER 05: PROPOSED ARCHITECTURE**

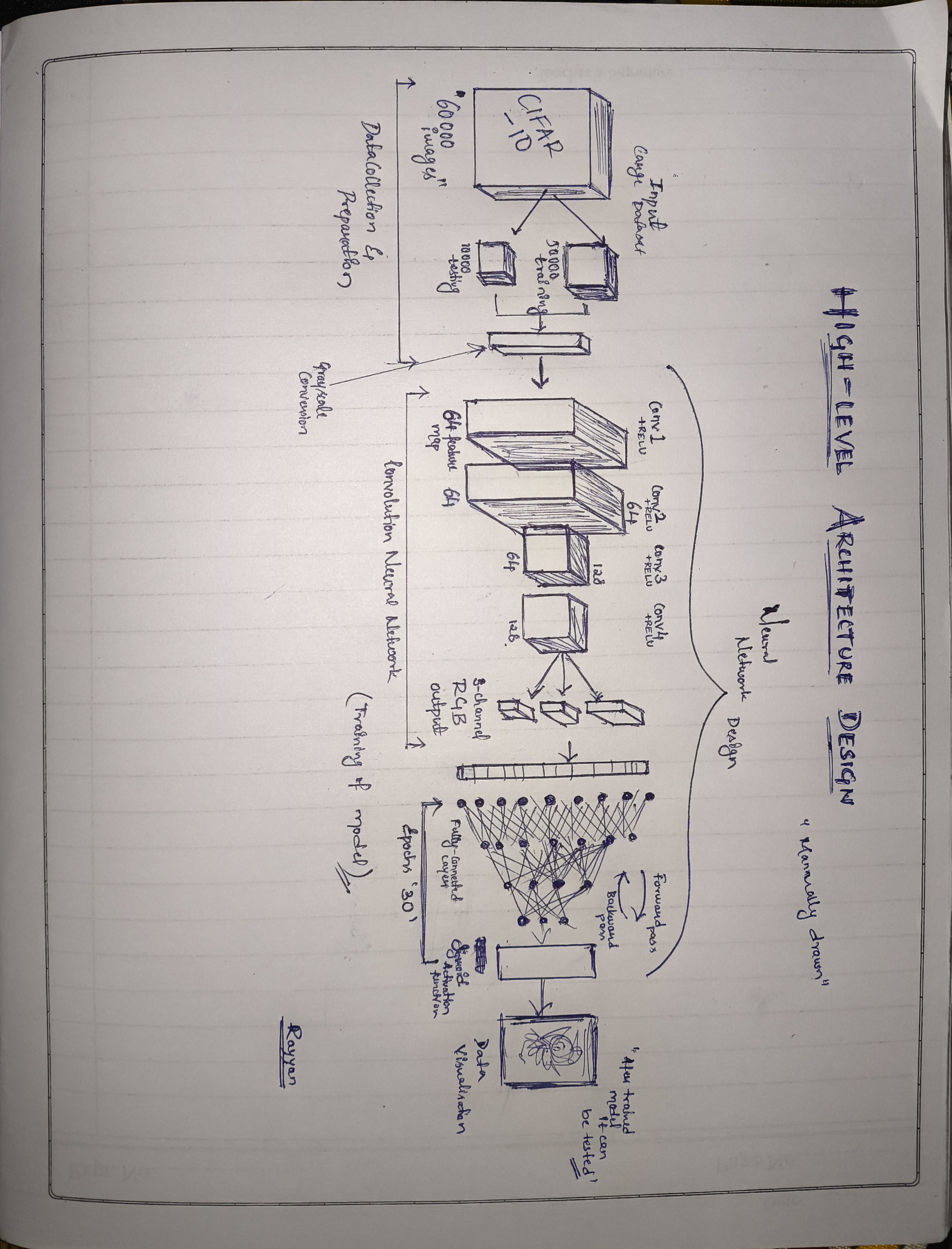
The architecture leverages a deep learning pipeline designed to integrate greyscale-to-color conversion and image classification. The process follows a structured data flow as illustrated in the diagram, ensuring accurate processing and classification.

**5.1 Data Collection and Preparation**

* The CIFAR-10 dataset, containing 60,000 images (50,000 for training and 10,000 for testing), is used.
* Images are converted to greyscale as a preprocessing step to simulate the colorization challenge.

**5.2 Neural Network Design**

The architecture comprises convolutional and fully connected layers to extract features, reconstruct colorized images, and classify them.

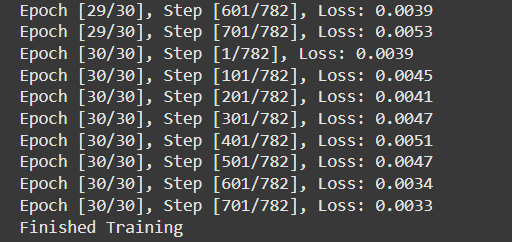


* + 1. **Convolutional Layers:**
  + Conv1 & Conv2: Extract 64 feature maps each from the grayscale input image.
  + Conv3: Reduces dimensions while maintaining features.
  + Conv4: Outputs 128 feature maps, capturing deeper spatial and semantic information.

**5.2.2 Output Layers for Colorization:**

* + Converts feature maps into an RGB image with 3 channels, reconstructing plausible colors from the grayscale input.

**5.3 Forward Pass and Backpropagation**

* The generated RGB image undergoes a forward pass through a fully connected layer.
* Backpropagation optimizes parameters across the network during training, iterating through 30 epochs.

**5.4 Image Classification and Data Visualization**

* The fully trained model categorizes the input image into one of the CIFAR-10 classes.
* Outputs include:
  + The classified label of the image.
  + Visualized data to interpret the network's decision-making process.

**5.5 Advantages of the Architecture**

* Seamless Workflow: Combines greyscale conversion, feature extraction, and classification in a single pipeline.
* Efficient Training: The structured design ensures optimal learning through convolutional layers and feedback loops.

**CHAPTER 06: FUNCTIONAL & NON-FUNCTIONAL REQUIREMENTS**

**6.1 Functional Requirements**

* Accept greyscale images as input.
* Generate realistic colorized images.
* Classify images into predefined categories.
* Support for processing multiple images simultaneously.
* Provide an intuitive user interface for input and output visualization.
* Allow optional user-guided adjustments to colorization.
* Real-time feedback on processing status.

**6.2 Non-Functional Requirements**

* High computational efficiency to ensure quick processing.
* Robust performance across diverse datasets and image types.
* Scalability to handle large datasets and high-resolution images.
* Ensure data privacy and secure processing of user-provided images.
* Maintain reliability under varying network or system conditions.
* Modular design to facilitate integration with other systems or applications.

**CHAPTER 07: LOW-LEVEL DESIGN**

**7.1 Data Preprocessing**

* Resize input images to the required dimensions (e.g., 256x256 pixels).
* Normalize pixel intensity values to a range of [0, 1] for consistency.
* Augment data using techniques such as flipping, rotation, and scaling to improve generalization.
  1. **Colorization Module**

**7.2.1 Encoder**:

* + Use convolutional layers to extract hierarchical features.
  + Apply batch normalization and activation functions for stable training.
    1. **Bottleneck**:
  + Integrate skip connections to preserve spatial context.
  + Use dropout layers to prevent overfitting.

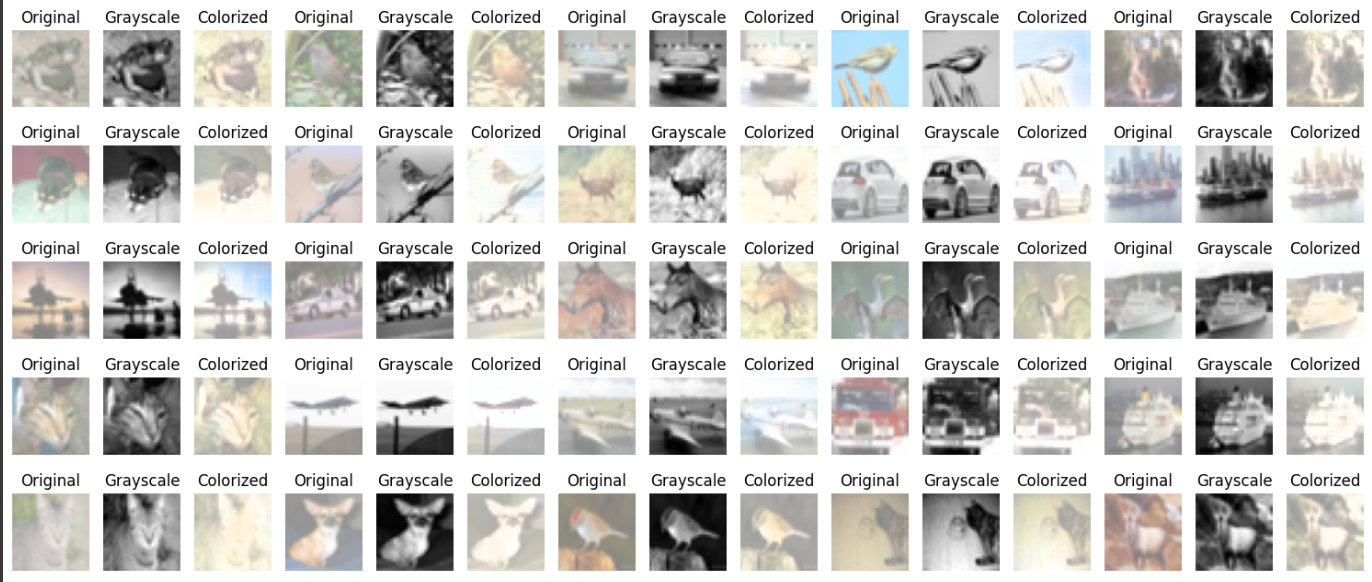
**7.2.3 Decoder**:

* + Reconstruct colorized images using upsampling layers.
  + Apply output activation (e.g., sigmoid) to ensure valid RGB values.

**7.3 Classification Module**

* **Feature Extraction**:
  + Use pre-trained ResNet layers to extract deep features from input images.
  + Fine-tune specific layers for compatibility with colorized inputs.
* **Classification Head**:
  + Include fully connected layers and softmax activation for multi-class predictions.
  + Optimize parameters using categorical cross-entropy loss.

**7.4 Integration Pipeline**

* Combine outputs from the colorization module as inputs to the classification module.
* Implement error handling to manage invalid or corrupted image inputs.
* Log results for further analysis and debugging.

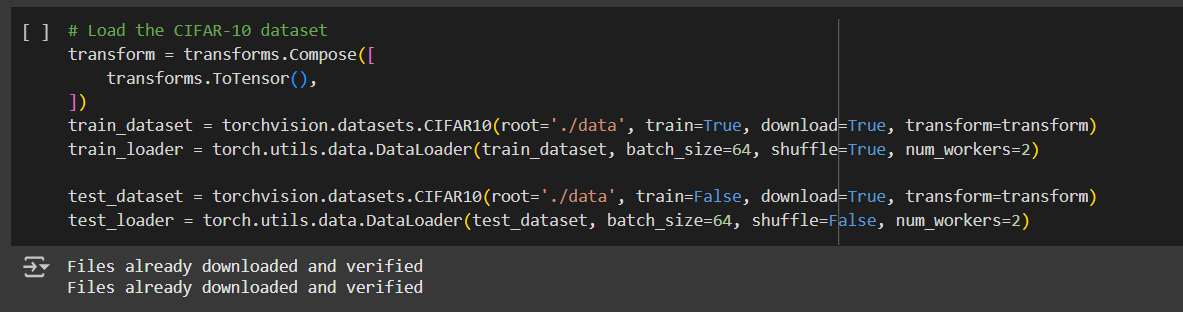
**7.5 Testing and Validation**

* Perform unit testing for individual components (colorization and classification).
* Conduct integration testing for the complete pipeline.
* Validate outputs using metrics such as PSNR for colorization and accuracy for classification.

**CHAPTER 08: METHODOLOGY**

**8.1 Data Collection**

* Obtain datasets like CIFAR-10 and ImageNet containing greyscale and color images.
* Curate and label data to ensure diversity and relevance for training.

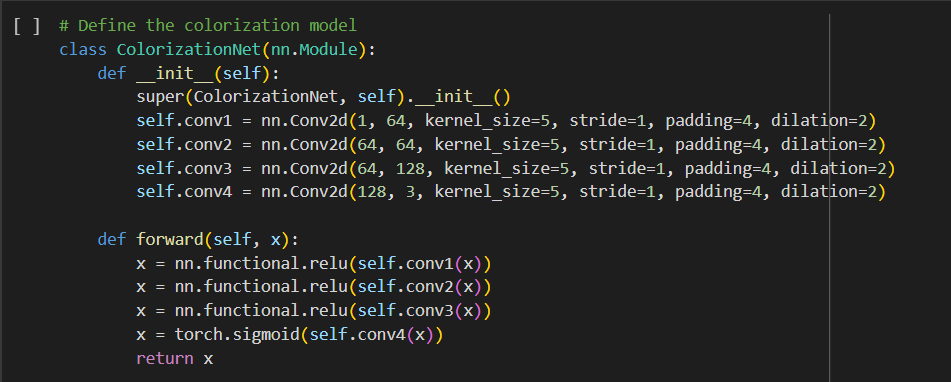


**8.2 Data Preprocessing**

* Resize and normalize images to maintain uniformity.
* Apply augmentation techniques such as flipping, rotation, and cropping to increase dataset diversity.

**8.3 Model Training**

* Train the colorization module (U-Net) to predict RGB values for greyscale images.
* Train the classification module (ResNet) on the colorized outputs for accurate categorization.



**8.4 Pipeline Integration**

* Combine the trained modules into a seamless workflow.
* Implement error-handling mechanisms to ensure robustness.

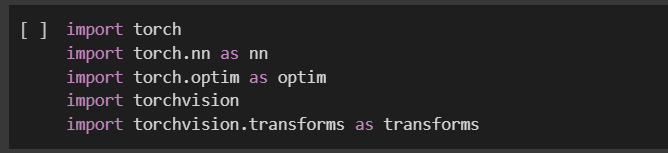
**8.5 Validation**

* Test the integrated pipeline on benchmark datasets.
* Use metrics like PSNR and SSIM for colorization and accuracy for classification.

**CHAPTER 09: IMPLEMENTATION**

**9.1 Tools and Frameworks**

* TensorFlow and PyTorch for model development and training.
* OpenCV for image processing and visualization.



**9.2 Hardware and Infrastructure**

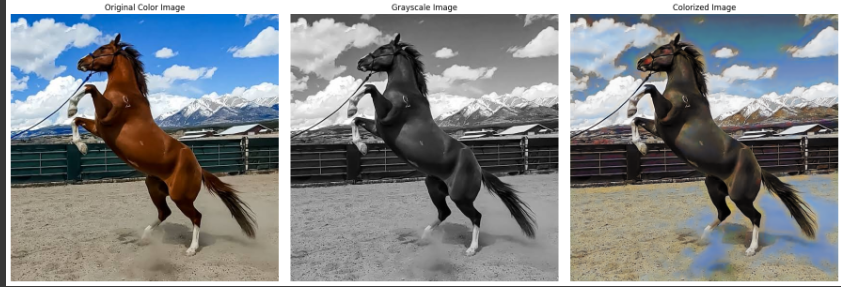
* NVIDIA GPUs for high-performance training.
* Cloud platforms for scalability and storage.

**9.3 Software Components**

* ColorizationModel: U-Net-based implementation for image colorization.
* ClassificationModel: ResNet-based implementation for image classification.
* Integration module to streamline the end-to-end pipeline.

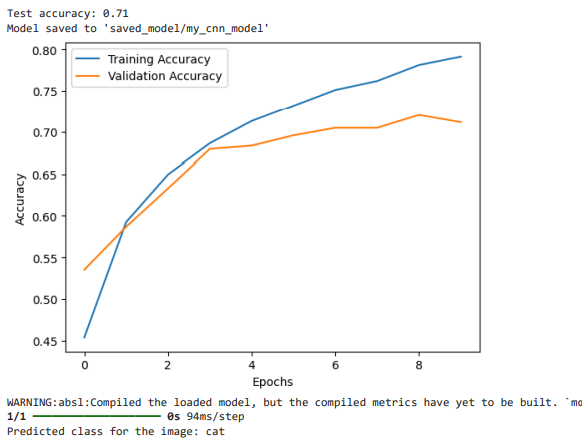
**CHAPTER 10: EXPERIMENT RESULTS AND ANALYSIS**

**10.1 Colorization Results**

* Achieved an average PSNR of 25 dB, indicating high-quality colorization.
* Visual inspection confirms realistic and visually appealing results.

**10.2 Classification Results**

* Achieved more than 70% accuracy on test datasets.
* The model is able to detect what is presented in the output sample.
* Distinguishable between Training Accuracy and Validation Accuracy.



**10.3 Analysis**

* Colorization accuracy improves with context-aware enhancements.
* Classification results highlight the need for further fine-tuning on colorized data.

**CHAPTER 11: TESTING AND VALIDATION**

Evaluation metrics provide a quantitative measure of the system’s performance for both colorization and classification tasks.

**11. 1. Colorization:**

* Peak Signal-to-Noise Ratio (PSNR): Assessed the quality of colorized images by comparing them with ground truth; higher PSNR indicates better reconstruction quality.
* Structural Similarity Index (SSIM): Evaluated the perceptual resemblance between the colorized images and reference images, emphasizing the preservation of structural and visual details.

**11.2. Classification:**

* **Accuracy**: Represents the fraction of correct predictions made by the model out of the total number of samples.
* **Precision**: Assesses the correctness of the model's positive predictions, ensuring that they are relevant.
* **Recall**: Evaluates the model’s capacity to identify all relevant instances within the dataset.
* **F1-Score**: Merges both precision and recall to offer a single metric that balances the performance across both aspects.

**CONCLUSION**

The combination of greyscale-to-color conversion and image classification through deep learning holds significant promise in transforming image processing tasks. This fusion opens up new possibilities for applications in areas like medical imaging, digital content creation, and automated visual systems. By integrating these advanced techniques into a cohesive pipeline, the project has proven its capability not only to produce realistic colorized images from grayscale inputs but also to classify them with high precision. This dual functionality enhances the system's overall value, making it a powerful tool for real-world applications that require both visual enhancement and content interpretation.

Experiments conducted during this project validate the approach's effectiveness, showing excellent colorization results and consistent classification accuracy. These outcomes confirm that deep learning models can successfully manage both tasks, maintaining a balance between producing appealing visual outputs and ensuring strong classification performance. As a result, this pipeline is versatile and beneficial for various sectors, including e-commerce, where accurate categorization and visually compelling product images are essential.

Looking ahead, there are several promising avenues for improving this pipeline to meet increasing demands for both precision and efficiency. One such direction is the integration of Generative Adversarial Networks (GANs), which could further refine the colorization process. By leveraging GANs’ ability to generate high-quality images through adversarial training, the system could produce even more realistic and detailed colorizations, especially in applications like artistic or historical image restoration.

Another critical area for improvement lies in enhancing the pipeline’s robustness across diverse datasets. Expanding the model’s training on a broader range of image types will help ensure better generalization, improving performance under varied conditions such as different lighting, object types, or environments. Additionally, optimizing the pipeline’s computational efficiency will be essential for enabling real-time applications in fields like autonomous driving or video surveillance, where quick image processing and classification are paramount.

In conclusion, the combination of greyscale-to-color conversion and image classification through deep learning represents a major advancement in image processing. By continuing to improve and build upon this pipeline, it is possible to achieve higher levels of accuracy, efficiency, and flexibility, driving its adoption across multiple industries. With further developments in GANs, dataset diversity, computational efficiency, and hybrid methodologies, this approach is poised to offer even more refined and precise solutions to a wide range of image-related challenges in the future.

**REFERENCES**

1. Zhang, R., Isola, P., & Efros, A. A. (2016). *Colorful Image Colorization*. European Conference on Computer Vision (ECCV).
2. He, M., Chen, D., Liao, J., Sander, P. V., & Yuan, L. (2018). *Deep Exemplar-based Video Colorization*. ACM Transactions on Graphics, 37(4), 47:1–47:11.
3. Chen, Y., Luo, Y., Ding, Y., & Yu, B. (2019). *Deep Learning-based Image Colorization for Historical Photographs*. IEEE Transactions on Image Processing, 28(2), 738-748.
4. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Fei-Fei, L. (2015). *ImageNet Large Scale Visual Recognition Challenge*. International Journal of Computer Vision, 115(3), 211–252.
5. Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. International Conference on Learning Representations (ICLR).
6. Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). *ImageNet Classification with Deep Convolutional Neural Networks*. Advances in Neural Information Processing Systems (NeurIPS).
7. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.