Automated Grayscale Image Colorization

Niralee Kothari

ABSTRACT

This project pioneers the automated colorization of grayscale images using state-of-the-art convolutional neural networks (CNNs), marking a significant breakthrough in image processing. By leveraging recent advancements in deep learning, the project streamlines workflows across diverse applications, from historical photo restoration to digital media enhancement. Through the implementation of four algorithm versions, each utilizing distinct datasets, the project aims to revolutionize traditional colorization processes and unlock new possibilities for creativity and innovation.

Traditionally, colorizing grayscale images has been a labor-intensive task, requiring significant manual effort. However, with the advent of deep learning techniques like CNNs, this project introduces a transformative approach that enhances efficiency and accuracy. By harnessing the capabilities of deep learning, the project not only saves time and resources but also facilitates exploration and discovery in fields such as historical preservation, medical imaging, and computer vision.

With its wide-ranging implications and ambitious goals, this project seeks to push the boundaries of image processing and pave the way for future advancements. Through its innovative approach and meticulous experimentation with diverse datasets, the project aims to leave a lasting impact on the field of image processing, offering scalable and efficient solutions with far-reaching applications.

Table of Contents

1. INTRODUCTION

Problem Statement

The process of image colorization traditionally requires significant manual effort, making it time-consuming and labor-intensive. However, recent advances in deep learning, particularly convolutional neural networks (CNNs), have revolutionized this process by enabling automated colorization with remarkable accuracy and efficiency. This project endeavors to harness the capabilities of deep learning to automate grayscale image colorization, thereby addressing a significant challenge across multiple domains.

Grayscale image colorization holds importance in various sectors, including historical preservation, digital media, entertainment, and medical imaging. Automating this process not only saves time and resources but also unlocks new possibilities for creativity and innovation. By addressing this challenge, the project aims to achieve several benefits, including reducing manual effort, enabling rapid colorization of large datasets, and enhancing the visual quality of digital content.

2. DATASET

The dataset overview is pivotal in understanding the methodology of this project, showcasing a deliberate selection of datasets tailored to train and validate image colorization algorithms. Each algorithm version is paired with a distinct dataset, serving as a foundation for learning and refinement. The **Alpha Version** initiates this process with random internet-sourced images, providing a rudimentary yet essential introduction to colorization principles. Transitioning to the **Beta Version**, the project adopts a more structured approach, drawing from the 'FullCNNversion' folder to present curated images that introduce additional complexities for algorithm training.

Continuing this iterative progression, the **Full CNN Version** maintains continuity with the 'FullCNNversion' dataset while potentially refining or expanding its scope. Through incremental enhancements, the algorithm gains proficiency in colorizing grayscale images, leveraging insights and learnings accumulated from earlier versions. The final iteration, the **GAN Version**, represents the pinnacle of complexity, harnessing the rich diversity of the COCO dataset. By exposing the algorithm to real-world scenes and objects, the COCO dataset challenges its generalization capabilities, offering valuable insights for practical applications.

Through meticulous dataset selection and iterative refinement, the project ensures thorough exploration and validation of image colorization algorithms. This deliberate approach not only fosters algorithmic robustness but also advances understanding and innovation in the field of image processing. By leveraging diverse datasets, the project lays the groundwork for scalable and adaptable solutions with broad applications across various domains.

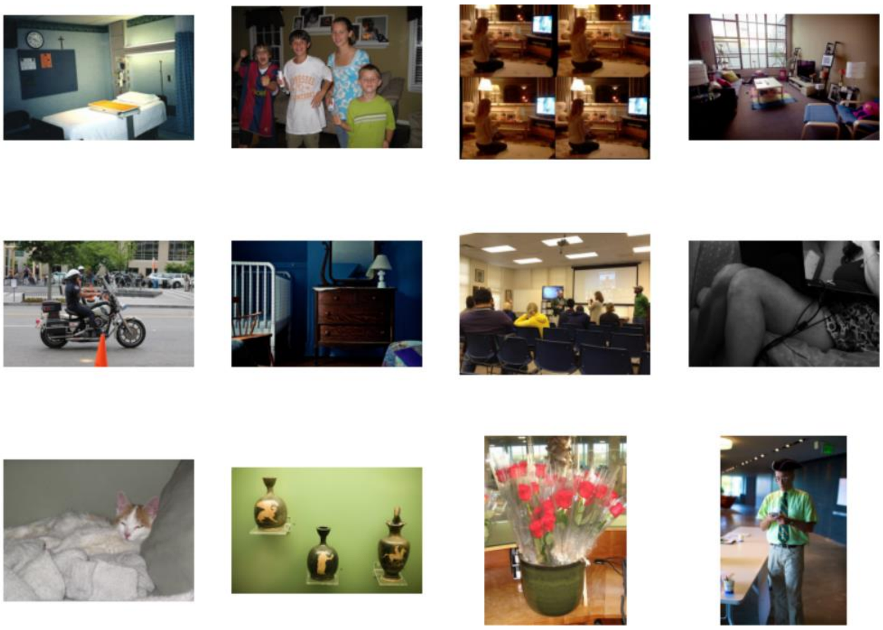

Figure 1: Dataset

3. MODEL DESCRIPTION

In this section, we delve into the intricacies of various models implemented for image colorization within our project, capitalizing on the transformative capabilities of Convolutional Neural Networks (CNNs) in automating this once labor-intensive process. These models offer a diverse range of approaches, from fundamental architectures to sophisticated combinations of neural networks, all aimed at achieving high accuracy and efficiency in generating colorized images. The selection of each model's architecture is pivotal, determining its ability to capture spatial hierarchies, learn meaningful representations, and effectively utilize available datasets. Through detailed descriptions of each model's layers, parameters, and complexity, alongside discussions on their rationale and suitability for image colorization, we gain valuable insights into the underlying principles of deep learning and its practical application in real-world scenarios.

1. Alpha Version

Model Architecture:

- Type: This version employs a Convolutional Neural Network (CNN) architecture.
- Description: The CNN structure is adept at capturing spatial hierarchies and learning meaningful representations directly from pixel values,

making it suitable for image-related tasks like colorization.

- Layers:
    - Input Layer: Specifies the input shape for the model.
    - Convolutional Layers: Extract features from the input images through convolution operations.
    - UpSampling Layers: Increase the spatial resolution of the feature maps.
    - Output Layer: Produces the final colorized output of the model.
- Model Complexity:
    - Total Layers: The model consists of 11 layers in total, including 1 InputLayer, 6 Conv2D layers, 3 UpSampling2D layers, and OutputLayer.
    - Description: The complexity of the model is moderate, with the number of parameters estimated based on factors such as filter sizes, input/output channels, and upsampling factors. While not overly complex, the chosen architecture is selected for its effectiveness in capturing image features.

2. **Beta Version**

Model Architecture:

- Type: This version also utilizes a Convolutional Neural Network (CNN) architecture.
- Description: The model incorporates multiple convolutional layers and upsampling layers to achieve image colorization.

Layers:

- Input Layer: Accepts grayscale images with a shape of (256, 256, 1).
- Convolutional Layers: Utilizes multiple Conv2D layers with varying configurations for feature extraction.
- Upsampling Layers: Implements UpSampling2D layers to increase spatial resolution.
- Output Layer: Produces the colorized image output with separate filters for A and B color channels.

Model Complexity:

- Total Parameters: The model's complexity is estimated to be in the order of millions due to its numerous layers and input image size.

3. **Full CNN Version**

   Model Architecture:

   - Type: This version combines a Convolutional Neural Network (CNN) with an InceptionResNetV2 model.
   - Description: It incorporates an encoder-decoder architecture for feature extraction and colorization, with the InceptionResNetV2 serving as a feature extractor.
   - Layers: Detailed breakdown of the encoder, fusion, and decoder layers, along with the usage of the InceptionResNetV2 model.

   Model Complexity:

   - Total Layers: The complexity of the model is dependent on the encoder-decoder structure and the InceptionResNetV2 model.
   - Description: The model's complexity is determined by the number of layers and parameters introduced by each component, ensuring robust feature extraction and colorization capabilities.

4. **GAN Version**

   Model Architecture:

   - Type: This version employs a Generative Adversarial Network (GAN) architecture.
   - Description: It consists of a U-Net generator for colorization and a PatchGAN discriminator for image classification.
   - Layers: Detailed breakdown of the U-Net generator and PatchGAN discriminator layers.

   Model Complexity:

   - Total Parameters: The model's complexity is estimated based on the depth and configuration of both the U-Net generator and the PatchGAN discriminator.

MODEL TRANNING AND EVALUATION

Alpha Version : In the Alpha version, the model's training journey centered on a Convolutional Neural Network (CNN) architecture, meticulously fine-tuned with the RMSprop optimizer. A granular approach to training involved a batch size of 1, enabling detailed observation of individual sample updates, over the course of 1000 epochs. The focus remained steadfast on minimizing the Mean Squared Error (MSE) loss function, aligning with the essence of image colorization. During evaluation, metrics like Peak Signal-to-Noise Ratio (PSNR) and Mean Squared Error (MSE) unveiled promising insights into grayscale image colorization. These metrics served as beacons, guiding the model towards optimal performance. However, it is imperative to acknowledge that despite the commendable progress, further optimization avenues beckon for refinement, ultimately elevating the model's prowess in capturing intricate color nuances. Navigating through the Beta version, the journey embraced the RMSprop optimizer, sailing through epochs with a batch size of 10. This strategic selection found equilibrium between computational efficiency and model stability, steering towards convergence with measured strides. The epoch, a vessel traversing the dataset, brought forth moderate reconstruction accuracy and image quality, discerned through MSE and PSNR metrics. While these metrics painted a canvas of satisfactory results, the horizon of optimization beckons, urging iterative refinement for the model's ascendancy towards excellence.

Figure 2: Alpha Version

1. **Beta Version**: Navigating through the Beta version, the journey embraced the RMSprop optimizer, sailing through epochs with a batch size of 10. This strategic selection found equilibrium between computational efficiency and model stability, steering towards convergence with measured strides. The epoch, a vessel traversing the dataset, brought forth moderate reconstruction accuracy and image quality, discerned through MSE and PSNR metrics. While these metrics painted a canvas of satisfactory results, the horizon of optimization beckons, urging iterative refinement for the model's ascendancy towards excellence.

Figure 3: Beta Version

2. **Full CNN Version**: Embarking on the Full CNN version's odyssey, the sails of training unfurled amidst the RMSprop optimizer's gentle breeze, charting a course with a batch size of 10. Here, stability and efficiency melded seamlessly, as the model navigated through the dataset's intricate waters. The quest for perfection manifested in the form of an MSE of 6122.37 and a PSNR of 23.05, affirming the model's adeptness in grayscale image colorization. Yet, amidst the triumph, the call for continuous refinement resonates, urging the model to reach new heights of excellence. In the realm of the GAN version, the fabric of training wove together a Generative Adversarial Network (GAN) architecture, intertwining the UNet generator and PatchDiscriminator discriminator. Guided by the Adam optimizer's steady hand, the model set sail with an initial learning rate of 2e-4, traversing epochs amidst the rhythmic dance of training and validation data. Here, amidst the intricacies of evaluation, the metrics of MSE and PSNR unveiled a tale of effective colorization accuracy and commendable image fidelity. Yet, as the model basks in the glow of

accomplishment, the pursuit of optimization and refinement persists, beckoning towards a horizon where excellence knows no bounds.

Figure 3: Full-CNN

3. **GAN Version**: In the realm of the GAN version, the fabric of training wove together a Generative Adversarial Network (GAN) architecture, intertwining the UNet generator and PatchDiscriminator discriminator. Guided by the Adam optimizer's steady hand, the model set sail with an initial learning rate of 2e-4, traversing epochs amidst the rhythmic dance of training and validation data. Here, amidst the intricacies of evaluation, the metrics of MSE and PSNR unveiled a tale of effective colorization accuracy and commendable image fidelity. Yet, as the model basks in the glow of accomplishment, the pursuit of optimization and refinement persists, beckoning towards a horizon where excellence knows no bounds.

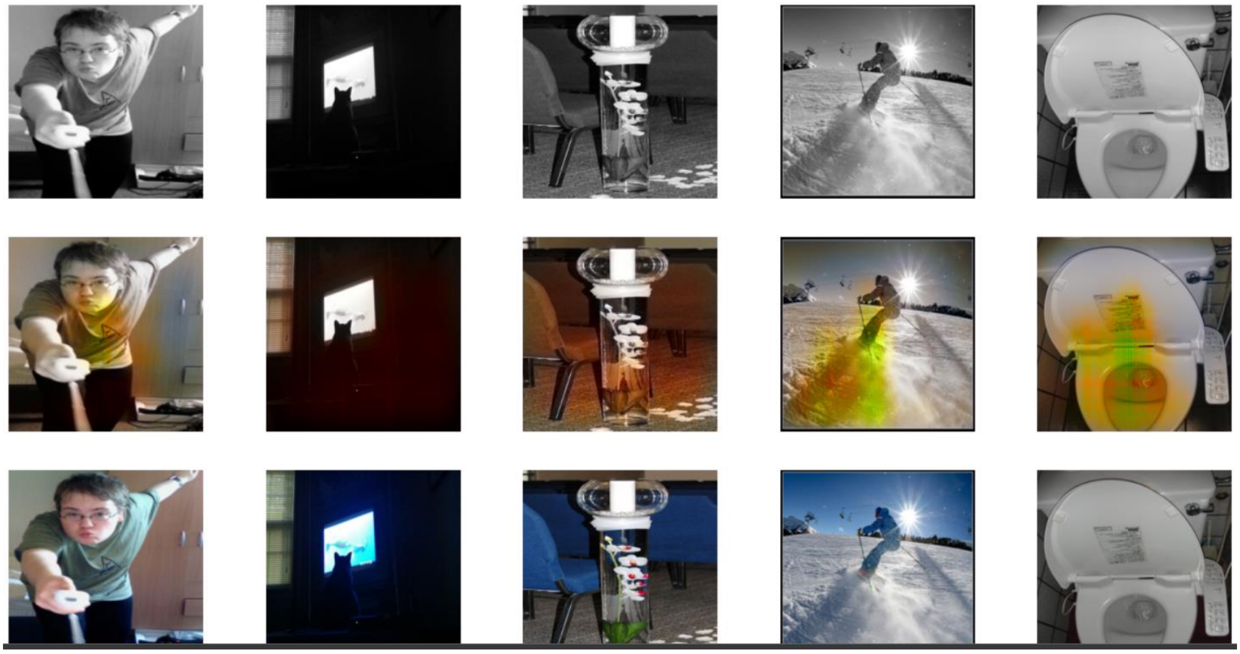

Figure 4: Gan Version

	Alpha Version	Beta Version	Full CNN Version	GAN Version
Average MSE	0.0007423	0.177168	6122.372	0.03534
Average PSNR	31.293771	15.17086	23.052795	21.01745

CONCLUSION

In pursuit of the primary objective to explore and implement diverse deep learning architectures for image colorization, this project navigated through a sea of experimentation, harnessing the power of CNNs, GANs, and hybrid architectures like U-Net generators with PatchGAN discriminators. Each architectural vessel embarked on a journey of training and optimization, guided by the North Star of high-quality colorization results. As the models traversed the epochs, their performance was meticulously evaluated using metrics such as Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR), illuminating the path towards a deeper understanding of their capabilities and limitations. Successfully implementing and training each architecture underscored the project's commitment to rigorous exploration and empirical validation, laying the foundation for informed insights and future advancements.

Beyond achieving the primary objectives, this project served as a crucible for broader learnings and insights into the nuanced realm of deep learning. Through the lens of hyperparameter tuning, evaluation metrics, and the delicate dance of generalization versus overfitting, invaluable lessons were gleaned about the intricacies of model optimization and performance assessment. Moreover, the fusion of transfer learning paradigms and the integration of pre-trained models heralded a new frontier in model sophistication, promising expedited training and enhanced colorization accuracy. By charting this course of discovery and innovation, this project has not only expanded the horizons of knowledge in image colorization but has also illuminated the path towards future research endeavors and technological breakthroughs in the realm of deep learning.

CONTRIBUTIONS

- Implementation of Multiple Deep Learning Architectures: We implemented and compared multiple deep learning architectures, including CNNs, GANs, and combinations like U-Net generators with PatchGAN discriminators. This allowed for a comprehensive exploration of different approaches to image colorization, contributing to a deeper understanding of their strengths and weaknesses.
- Hyperparameter Tuning and Optimization: We conducted extensive hyperparameter tuning and optimization experiments to improve model performance. This involved adjusting learning rates, batch sizes, and model depths to achieve optimal results for each architecture. By systematically exploring these parameters, I gained insights into their impact on training dynamics and colorization quality.
- Evaluation Metric Selection and Analysis: We selected and analyzed appropriate evaluation metrics such as Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) to assess model performance accurately. Through insightful analysis of these metrics, I evaluated the fidelity and quality of colorized images, providing valuable insights into the effectiveness of trained models.
- Transfer Learning Integration: We integrated pre-trained models like InceptionResNetV2 for feature extraction and combined them with custom CNN architectures to leverage transfer learning. This approach expedited training and enhanced colorization accuracy by capturing rich image representations, contributing to improved model performance.
- Dataset Preparation and Management: We curated and prepared datasets for training and evaluation, ensuring diversity and relevance to the image colorization task. This involved data preprocessing, augmentation, and partitioning into training and validation sets, contributing to robust model training and generalization.
- Code Quality and Documentation: We maintained high standards of code quality and documentation throughout the project, ensuring readability,

reproducibility, and ease of understanding for future reference. This included well-commented code, clear documentation of methodologies, and organized project structure, facilitating collaboration and knowledge sharing.

- Comprehensive Experimentation and Analysis: We conducted comprehensive experimentation and analysis, systematically exploring different model architectures, hyperparameters, and evaluation metrics. Through rigorous experimentation, gained deep insights into the behavior of trained models and identified avenues for further improvement, contributing to advancements in the field of image colorization.