Image Colorization with deep convolutional Neural Network

Anika Nazhat
Mount Holyoke College
nazha24a@mtholyoke.edu

***Abstract*— This project explores Image Colorization using a Convolutional Neural Network (CNN)-based encoder-decoder architecture. The model takes colored galaxy images as input, where the primary task is to predict the missing color channels (a and b) in the Lab color space. The network is trained on a dataset consisting of colored galaxy images, learning to extract spatial features and associate them with accurate color predictions. The model is trained for several epochs, where the grayscale L channel of the Lab color space is used as input and the model learns to predict the chrominance channels (a and b). The trained model is evaluated on a set of test images to assess its performance and ability to generalize to new, unseen galaxy images. The aim of this project is to develop a deep learning model that can provide realistic colorization of galaxy images. Space images contain unique color patterns that correspond to various elements and cosmic phenomena. Colorizing grayscale versions of these could help train a model that can detect such phenomena through colorization.**

***Index Terms*— keyword 1; keyword 2; keyword 3**

I. Introduction

Image colorization is the process of transforming grayscale images into color based on the image's content. Colorization has been revolutionized by deep learning, particularly with Convolutional Neural Networks (CNNs). CNNs are well-suited for this task as they extract image features—such as edges and textures—through convolutional layers, allowing the model to analyze grayscale images and generate plausible color predictions by learning from large datasets.

Our approach uses a CNN-based encoder-decoder architecture to predict colors for grayscale images. Our model takes in colorized images. These images are then converted to separate the grayscale (L channel) from the color information (a and b channels). The L channel serves as the model's input. This setup allows the network to learn associations between grayscale patterns and specific color values by training on a dataset of color images, enabling it to generate plausible colorizations for new images. This approach assumes that effective colorization requires understanding both the image's structure and its content, which guides the prediction of realistic color values.

The model is trained on a set of colored images. Initially, the images are converted to the LAB color space where the L(Lightness) colorspace, which is the grayscale version of the images, is taken as the input for the Convolutional Neural networks (CNNs). The model architecture consists of an encoder-decoder structure with convolutional layers, batch normalization, and upsampling techniques. This architecture is designed to capture the spatial features in the input image and use them to reconstruct the missing color information. After training, the model is capable of generating colorized versions of grayscale images where it predicts the A and B colorspace of the images, even those it has never seen before.

The goal of this project is not only to generate aesthetically pleasing colorized images but also to demonstrate how deep learning can be used to tackle a classical problem in image processing. By the end of this project, we aim to understand the strengths and limitations of using a CNN for image colorization, and provide insights into the potential for further improving the quality of automated image colorization methods.

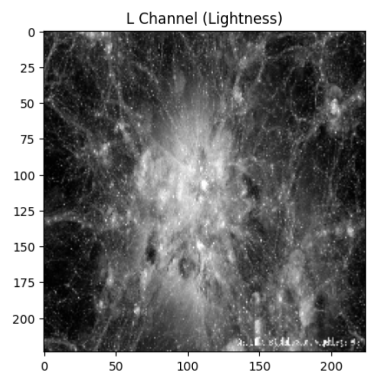

Fig. 1. grayscale (L channel)

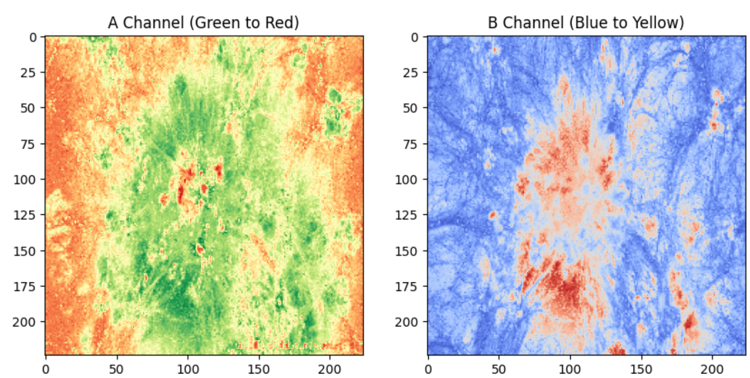

Fig. 2. color information (a and b channels)

II. Existing Methods

The task of image colorization has seen significant advancements, transitioning from interactive systems to fully automated deep learning methods. Early work focused on user-guided techniques, such as grouping pixels into coherent regions for color mapping. For example, Luan et al [1] proposed a method where users group regions with similar textures or intensities, followed by a mapping process to apply vivid colors to each region. This approach reduced user effort compared to purely manual methods but still required substantial interaction, limiting scalability to larger datasets or complex images [1].

Deep learning has greatly transformed the field, with convolutional neural networks (CNNs) emerging as the dominant approach. Hwang and Zhou presented a CNN-based system that utilized the CIELUV color space to separate the luminance (grayscale) and chrominance (color) components of an image. Their model employed a regression-based approach to predict chrominance channels, though they noted limitations in the vibrancy of the generated colors due to averaging effects inherent in their loss function [2].

Addressing the ambiguity in colorization, Deshpande et al [3] developed a method based on variational autoencoders (VAEs) to generate multiple plausible colorizations for a single grayscale input. By learning a low-dimensional embedding of color fields and employing a conditional mixture density network, their approach produced diverse and spatially coherent results, outperforming simpler CNN models in terms of visual realism [3].

The growing availability of large datasets has enabled more automated methods. Techniques discussed in the ECCV proceedings emphasize the importance of high-quality training data and robust architectures, such as encoder-decoder models, to improve generalization to new images. These models leverage features extracted from pretrained networks, often combined with transfer learning to accelerate convergence and improve performance [4] [5].

While traditional methods like those discussed by Luan et al [1]. focused on interactive segmentation and mapping, modern approaches prioritize minimizing user involvement and enhancing the realism of generated colorizations. However, challenges remain in handling high-resolution images, generalizing across diverse datasets, and managing the inherent ambiguity of colorization tasks.

III. Open challenges

Current limitations: Image colorization faces several challenges, with ambiguity being a key issue. Multiple plausible colorizations can exist for the same grayscale image, and while methods like VAEs and CMDNs aim to generate diverse outputs, they often struggle to balance realism with variation. This can result in outputs that are either overly uniform or lack spatial coherence [3] [2].

Another limitation is the difficulty of generalizing across diverse datasets. Models trained on specific datasets often fail to perform well on images with unfamiliar textures or objects. While transfer learning improves feature extraction, it falls short in adapting to unique domains like scientific or high-resolution images. Additionally, computational demands for training and inference at higher resolutions limit scalability, making efficient, accurate high-resolution colorization an ongoing challenge [2] [4].

Research Idea: The research question I am exploring is "How can deep learning, specifically convolutional neural network (CNN)-based encoder-decoder architectures, be utilized to generate realistic and scientifically meaningful colorizations of grayscale galaxy images?".

Recent studies, such as Kalvankar et al. (2021) [6], have used Generative Adversarial Networks (GANs) and pretrained ResNet-18 models to automate the colorization of astronomical images, focusing on improving resolution and visual appeal. Similarly, Rector et al. (2007) [7] discussed techniques for combining multiwavelength data to create accurate and visually compelling images [6] [7].

My research differs by aiming to not only colorize galaxy images but also teach the model to identify cosmic phenomena, like star formations, through these colorizations. To achieve this, I will focus on refining the model to learn patterns that connect specific colors with physical features in the images, making the outputs both visually meaningful and scientifically informative.

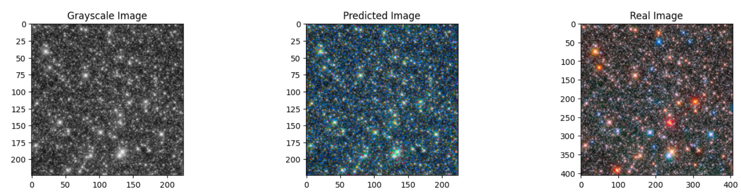

Fig. 3. Grayscale, Predicted and Real Image

IV. Concept to Code

Here is the GitHub source code folder link:
`https://github.com/MHC-FA24-CS341CV/beyond-the-pixels-emerging-computer-vision-research-topics-fa24/tree/main/code/12-im-colorization/FINAL_Image_Colorization_Using_CNN_Anika`

Description of the Dataset(s): The dataset used for this project consists of colored galaxy images, specifically curated for training and testing the colorization model.

Source: Galaxy images are typically sourced from publicly available astronomy datasets, such as those provided by NASA or the Sloan Digital Sky Survey (SDSS). For this project, the dataset was sourced from Kaggle.

Key Characteristics: Images are in RGB format and are converted into the Lab color space during preprocessing. The Lab color space separates lightness (L) from color information (a and b), enabling a focus on grayscale structure for input and chrominance for output. The dataset includes diverse galaxies with unique structures, textures, and color patterns, which provide rich training data for the model to learn realistic and scientifically meaningful colorization.

Run Instructions: To execute the Colab notebook for the project, follow these steps:

Dataset Preparation: Upload the dataset to a Google Drive folder. Ensure images are in a compatible format (jpg) and structured into training and validation directories.

Google Colab Setup: Mount Google Drive by running the provided from google.colab import drive command. Provide the appropriate path to access the dataset from Google Drive. Environment Setup:

Install the necessary libraries such as TensorFlow, NumPy, OpenCV, and any other dependencies listed in the notebook.

Running the Code: Execute the notebook cells sequentially: Data Preprocessing: Converts the dataset to Lab color space and prepares training data. Model Training: Trains the CNN model to predict a and b channels from the L channel. Evaluation: Uses the test dataset to predict colorized outputs of the grayscale images.

Visualizing Results: Generated images can be visualized directly in the notebook or saved back to Google Drive for offline access.

Connection to the Proposed Research Idea: This program supports the research idea of using deep learning to enhance image analysis for scientific purposes, specifically by enabling the colorization of grayscale galaxy images.

Scientific Relevance: Color in galaxy images corresponds to key cosmic phenomena (e.g., star formation regions, chemical composition). By training the model to infer accurate color patterns, this project attempts to train a model to not only colorize images but also detect these phenomena through this process which can be useful in the field of Astronomy.

Next Steps for Technical Implementation: To enhance the model's performance and applicability, the following steps can be taken in the future:

Dataset Expansion: Include larger and more diverse datasets, such as those with different galaxy types and spectral imaging data, to improve the model's generalizability.

Model Refinement: Experiment with advanced architectures like U-Net or GAN-based networks to enhance spatial feature extraction and produce more realistic outputs [5]. Performance Optimization:

Fine-tune hyperparameters such as learning rate, batch size, and training epochs. Implement techniques like data augmentation to make the model more robust to variations.

Integration of Domain Knowledge: Incorporate astrophysical constraints or rules into the model (e.g., expected colors for certain regions) to guide the predictions.

REFERENCES

[1] Q. Luan, F. Wen, D. Cohen-Or, L. Liang, Y.-Q. Xu, and H.-Y. Shum, "Natural image colorization," in *Proceedings of the 18th Eurographics Conference on Rendering Techniques*. Eurographics Association, 2007, pp. 309–320.

[2] J. Hwang and Y. Zhou, "Image colorization with deep convolutional neural networks," Stanford University," Technical Report, 2016.

[3] A. Deshpande, J. Lu, R. A. Yeh, M. Jin, and D. Forsyth, "Learning diverse image colorization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2877–2885.

[4] G. Charpiat, M. Hofmann, and B. Schölkopf, "Automatic image colorization via multimodal predictions," in *Computer Vision–ECCV 2008: 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part III*, vol. 10. Springer Berlin Heidelberg, 2008, pp. 126–139.

[5] R. Zhang, P. Isola, and A. A. Efros, "Colorful image colorization," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III*, vol. 14. Springer International Publishing, 2016, pp. 649–666.

[6] S. Kalvankar, B. Elger, H. Zacharia, and E. Tempel, "Astronomical image colorization and upscaling with generative adversarial networks," *arXiv preprint arXiv:2112.13865*, 2021.

[7] T. A. Rector, Z. G. Levay, L. M. Frattare, J. English, and M. Pu'uohau-Pummill, "Image-processing techniques for the creation of presentation-quality astronomical images," *The Astronomical Journal*, vol. 133, no. 2, pp. 598–611, 2007.