# Automatic Colorization of Grayscale Images Using Deep Learning

Aidan Sommer, Henock Dinberu, Thavishka Gamage, Dhriti Gada, Prateek Kharangate
*University of Cincinnati*
Email: sommerab@mail.uc.edu, dinberhw@mail.uc.edu, gamagetd@mail.uc.edu, gadada@mail.uc.edu, kharanpv@mail.uc.edu

## I. Background

Grayscale images have been the primary medium for photography for over a century. While historically significant, they lack the rich color information that is critical for human perception and vivid scene interpretation. They contain only luminance data. The process of manually adding color to these images is difficult and timely, that requires specialized artistic skill.

This project addresses the automatic colorization of grayscale images using deep learning. The motivation is to create a tool that can automate this difficult and time-consuming task, making it possible to restore historical archives or enhance media efficiently.

The core challenge is that automatic colorization is a notably ill-posed problem: a single luminance value can map to many potential chrominance values. For example, a gray circle in an image could plausibly be a red, green, or yellow. Therefore, a successful model cannot simply "solve" for the original color; it must infer a plausible and semantically consistent colorization. It must learn the statistical priors of the natural world—that sky is typically blue, grass is green, and skin exists within a specific color palette. This requires a model that learns high-level semantic features, not just low-level pixel relationships.

## II. Related Studies

This project builds upon foundational work in deep learning for image-to-image translation. While early methods relied on user-guided scribbles [5], modern approaches are fully automatic. Three state-of-the-art models are particularly relevant.

### A. Colorful Image Colorization (Zhang et al.)

- **Model Name and Authors:** "Colorful Image Colorization" by Richard Zhang, Phillip Isola, and Alexei A. Efros [1].
- **Methodology:** This work reformulates colorization from a regression problem to a classification problem. Instead of trying to predict the exact color value for a pixel (which often leads to dull, averaged results like grayish-brown), the model predicts a probability distribution over a quantized set of possible colors. It uses a VGG-style Convolutional Neural Network (CNN) trained on over a million images from the ImageNet dataset [4].
- **Performance:** Because traditional metrics like Mean Squared Error (MSE) do not capture the perceptual quality or vibrancy of an image, the authors used a "colorization Turing test." They conducted a user study where participants had to choose between the model-generated colorization and the original ground-truth color. The model successfully "fooled" human participants on 32% of trials, demonstrating a high degree of perceptual realism.

### B. Image-to-Image Translation (Pix2Pix)

- **Model Name and Authors:** "Image-to-Image Translation with Conditional Adversarial Networks" (Pix2Pix) by Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros [2].
- **Methodology:** This paper presented a general framework for image-to-image translation using a **Conditional Generative Adversarial Network (cGAN)**. The model consists of two parts: a "Generator" that creates the colorized image and a "Discriminator" that tries to distinguish the fake (generated) image from a real (ground-truth) color image. The Generator is trained to "fool" the Discriminator. A key part of its methodology is using a **U-Net** [3] based architecture for the generator, which allows low-level pixel information to be shared directly with high-level semantic layers.
- **Performance:** The model's strength is its ability to produce sharp and highly realistic images, as the adversarial loss function penalizes blurry or "unrealistic" outputs. Its performance was shown to be superior to simple L1 regression (which produces blurry results) on a wide variety of tasks, including colorization.

### C. U-Net (Ronneberger et al.)

- **Model Name and Authors:** "U-Net: Convolutional Networks for Biomedical Image Segmentation" by Olaf Ronneberger, Philipp Fischer, and Thomas Brox [3].
- **Methodology:** While not originally for colorization, the U-Net architecture is a key component in many modern solutions (including Pix2Pix [2]). Its novelty is its "skip connections," which form a 'U' shape. These connections link the down-sampling path (encoder) with the up-sampling path (decoder). This allows the network to combine deep, semantic feature information (e.g., "this

object is a car") with shallow, high-resolution feature information (e.g., "the edge of the car is right here"). This is ideal for colorization, which requires both understanding the object and precisely applying color at its boundaries.
- **Performance:** The model set a new state-of-the-art for biomedical image segmentation. In the context of colorization, its contribution is enabling the high-resolution, sharp results that were missing from earlier encoder-decoder models.

Our project will build upon the methodologies of these key papers, comparing a classification-based approach [1] with a cGAN-based approach [2].

## REFERENCES


[1] R. Zhang, P. Isola, and A. A. Efros, "Colorful image colorization," in *European conference on computer vision (ECCV)*, 2016, pp. 649–666.

[2] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 2017, pp. 1125–1134.

[3] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention (MICCAI)*, 2015, pp. 234–241.

[4] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 2009, pp. 248–255.

[5] A. Levin, D. Lischinski, and Y. Weiss, "Colorization using optimization," in *ACM transactions on graphics (TOG)*, 2004, vol. 23, no. 3, pp. 689–694.