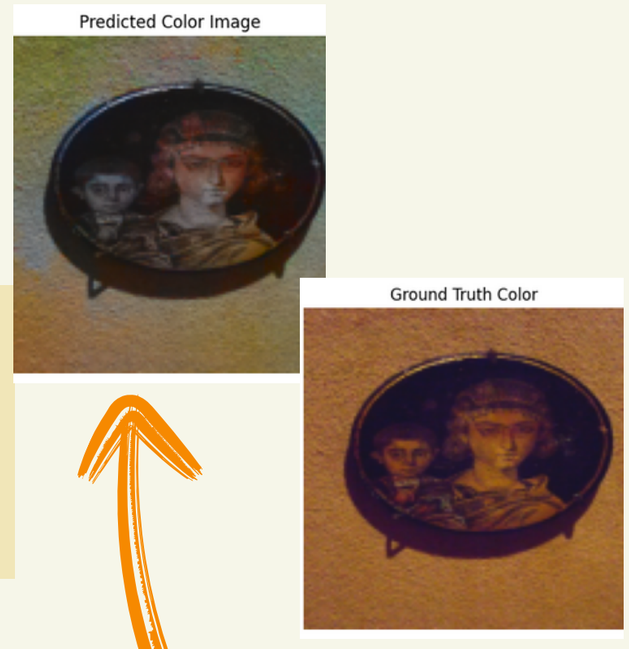
DEEP LEARNING FOR IMAGE COLORIZATION



This project addresses the task of automatic image colorization, which involves adding color to grayscale images without any manual input. We develop a deep learning model that combines the strengths of UNet and Generative Adversarial Networks (GANs). By treating colorization as a supervised image-to-image translation problem in the Lab color space, our model learns to predict the missing a and b chrominance channels from the input luminance (L) channel.

Authors
Andrijana Kešelj
Anđela Maksimović

Affiliations
Data Science and Artificial Intelligence
Faculty of Electrical Engineering
University of Sarajevo

INTRODUCTION

Image colorization is a significant task in computer vision with applications in photo restoration.
The problem is inherently ill-posed, as a single grayscale image can correspond to many valid color interpretations. Traditional techniques often relied on manual annotations or reference-based methods. With the rise of deep learning, especially Generative Adversarial Networks (GANs), it's now possible to learn color mappings directly from data. GANs are effective at producing realistic textures, while UNet architectures help preserve spatial and structural detail.

OBJECTIVE

Develop a deep learning model that can automatically and realistically colorize grayscale images using a GAN-UNet architecture.

METHODOLOGY

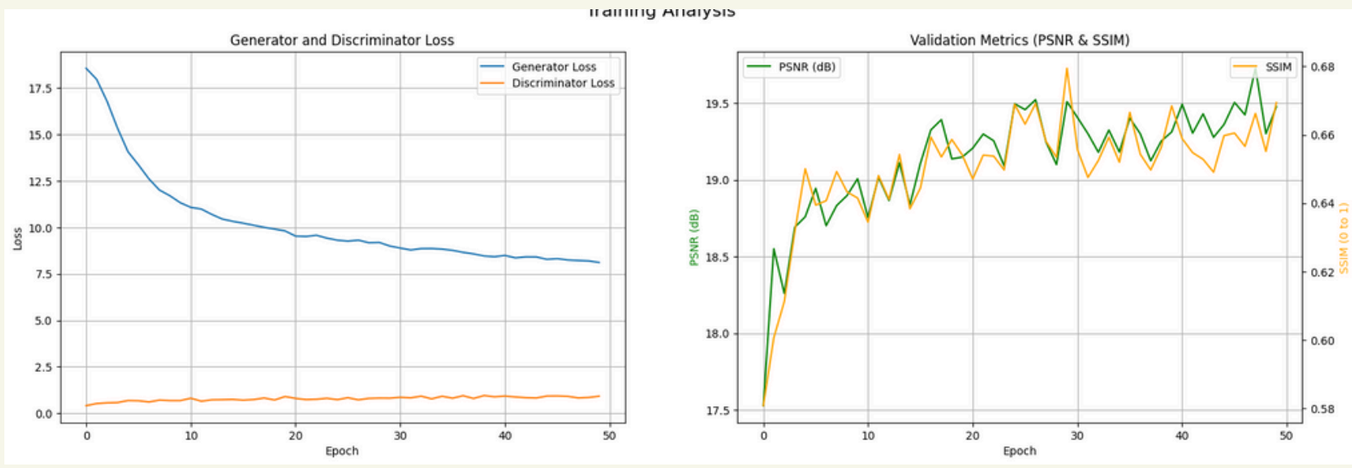
We use a UNet-based generator combined with a PatchGAN discriminator to perform image colorization. The model takes the grayscale L channel as input and predicts the a and b color channels in Lab color space.

RESULTS

The trained Generative Adversarial Network produces visually compelling colorizations, accurately rendering semantic features like skin tones. However, it struggles with ambiguous textures, substituting the ground truth's warm, golden hues with muted, greenish tones in the background. This performance is quantified by a final validation SSIM of 0.67, indicating strong structural preservation, and a PSNR of 19.48 dB, which reflects the color divergence. A Fréchet Inception Distance (FID) of 72.47 further confirms a perceptual similarity to real images, albeit with clear room for improvement.

ANALYSIS

Analysis of the training plots reveals a balanced adversarial dynamic, evidenced by a steadily decreasing generator loss against a stable discriminator loss. The consistent upward trend in validation PSNR and SSIM throughout training confirms that the model progressively improved in both pixel accuracy and structural fidelity. This dynamic suggests the generator successfully learned to produce increasingly realistic images without overpowering the discriminator. The final FID score corroborates the visual assessment that while the outputs are plausible, they have not yet fully converged with the distribution of real images.



CONCLUSION

This work successfully validates the efficacy of a conditional GAN for generating vibrant, perceptually realistic colorizations, affirming its superiority over simpler regression-based methods. While the results are compelling, the primary limitation was hardware-imposed constraints on model complexity and training resolution, which impacted the fidelity of fine-grained textures. Future work will directly address these limitations by scaling the model to higher resolutions, integrating attention mechanisms for enhanced contextual reasoning, and extending the training duration. These enhancements, systematically guided by real-time FID tracking, aim to bridge the gap between the current plausible outputs and true photorealistic quality.

Related literature

Pix2Pix

Colorful image colorization

Medium article

Stanford research

Contact:

akeselj1@etf.unsa.ba
amasksimovi1@etf.unsa.ba

Our GitHub repository:

https://github.com/aandrijana/Image-Colorization-Project/blob/main/README.md