

Acknowledgments

We extend our heartfelt gratitude to the **Samsung Innovation Campus** and the **Department of Electronics and Computer Engineering, Pulchowk Campus, Institute of Engineering, Tribhuvan University** for providing us with the opportunity to undertake this project. Our sincere thanks to the program coordinator and the Project Management Team for their continuous support and guidance throughout this journey.

This collaborative project has been an invaluable opportunity to apply our academic knowledge and gain hands-on experience in solving real-world problems. We are deeply appreciative of the support from both the department and the Samsung Innovation Campus program that enabled us to undertake this meaningful AI project.

Special recognition goes to our mentor **Bishal Pandey** and supervisor **Santosh Giri**, who generously shared their insights and expertise, and provided valuable feedback throughout the development of our project. We also extend our thanks to our friends and colleagues who contributed to this project in various ways, both directly and indirectly.

Our deepest appreciation is reserved for our family members, whose constant inspiration, encouragement, and unwavering support have been crucial to our success.

We welcome any suggestions or constructive criticism to further improve our work.

Authors:

Sandip Katel

Sharad Pokharel

Yujal Shrestha

Abstract

This project presents a Sketch to Artwork Converter—an automated deep learning system that transforms hand-drawn black-and-white sketches into colored, stylized artworks. Built upon a Pix2Pix-inspired conditional Generative Adversarial Network (cGAN) architecture with U-Net generator and PatchGAN discriminator, the system enables users to create sketches directly on a built-in canvas or upload existing drawings, which are then converted into realistic images across two domains: object-level translation (shoes and bags) and scene-level translation (nature landscapes). The generator employs an encoder-decoder structure enhanced with DeepBlocks and skip connections, trained using a combination of adversarial loss, L1 loss, and perceptual loss (LPIPS) on approximately 14,000 paired samples from Edge2Shoes and Edge2Bags datasets, plus 1,800 custom landscape pairs. Evaluation metrics reveal moderate performance with SSIM scores of 0.535 (objects) and 0.279 (scenes), PSNR values of 10.9 dB and 12.16 dB, FID scores of 120 and 88.96, and LPIPS scores of 0.44 and 0.593 respectively. While the system successfully demonstrates the feasibility of automated sketch-to-artwork conversion and provides valuable proof-of-concept for democratizing digital art creation, its restriction to limited object categories and performance gaps indicated by evaluation metrics highlight the need for more advanced architectures and expanded datasets to achieve production-quality output.