

Advances in Deep Learning for Automatic Image Colorization: A Survey

Ms. Kimaya Churi

Msc It, Sonopant Dandekar Arts, V.S. Apte Commerce and M.H. Mehta Science College, Palghar

Abstract—Automatic image colorization is a vital challenge in the field of computer vision that involves generating plausible colour images from grayscale input. This task has extensive applications in areas such as historical image restoration, digital media production, and content creation. Traditional approaches to this task involved manual work or relied on rule-based algorithms. However, with the emergence of deep learning techniques, particularly Conditional Generative Adversarial Networks (cGANs), the quality and realism of colorized images have dramatically improved. This survey explores the current state-of-the-art approaches in automatic image colorization, compares various techniques, and evaluates their performance using standard metrics such as PSNR, SSIM, and FID. Our findings reveal that cGAN-based models outperform traditional and CNN-based methods in generating perceptually realistic results.

1. INTRODUCTION

Image colorization has historically been a manual task requiring artistic input. Recent trends in artificial intelligence and deep learning have led to the automation of this process. The primary motivation behind automated colorization includes restoration of old photographs, enhancement of medical images, and creating artistic effects for modern media. Traditional colorization relied heavily on human effort or simple heuristics that lacked adaptability. The advent of deep learning, particularly convolutional neural networks (CNNs) and generative adversarial networks (GANs), revolutionized the field by learning from large datasets to generalize across image types.

2. PROBLEM STATEMENT

Many deep learning-based colorization models still suffer from semantic inaccuracies and lack contextual awareness, leading to unnatural colour distribution. This study investigates how effectively conditional

GANs (cGANs) address these limitations compared to both traditional methods and CNN-based architectures.

3. OBJECTIVES

- To study the evolution of image colorization techniques.
- To compare the efficacy of traditional, CNN-based, and cGAN-based methods.
- To analyse quantitative performance using PSNR, SSIM, and FID metrics.
- To visually compare results through qualitative evaluation and user studies.

4. METHODOLOGY

This research evaluates three primary models:

- Traditional colorization methods using heuristic rules.
- CNN-based models, like the one proposed by Zhang et al. (2016), using deep learning without adversarial training.
- cGANs (Conditional GANs), which condition the output on the input grayscale image and are trained adversarially.

Each model was tested on a dataset of standard grayscale-to-color image mappings. Evaluation metrics used:

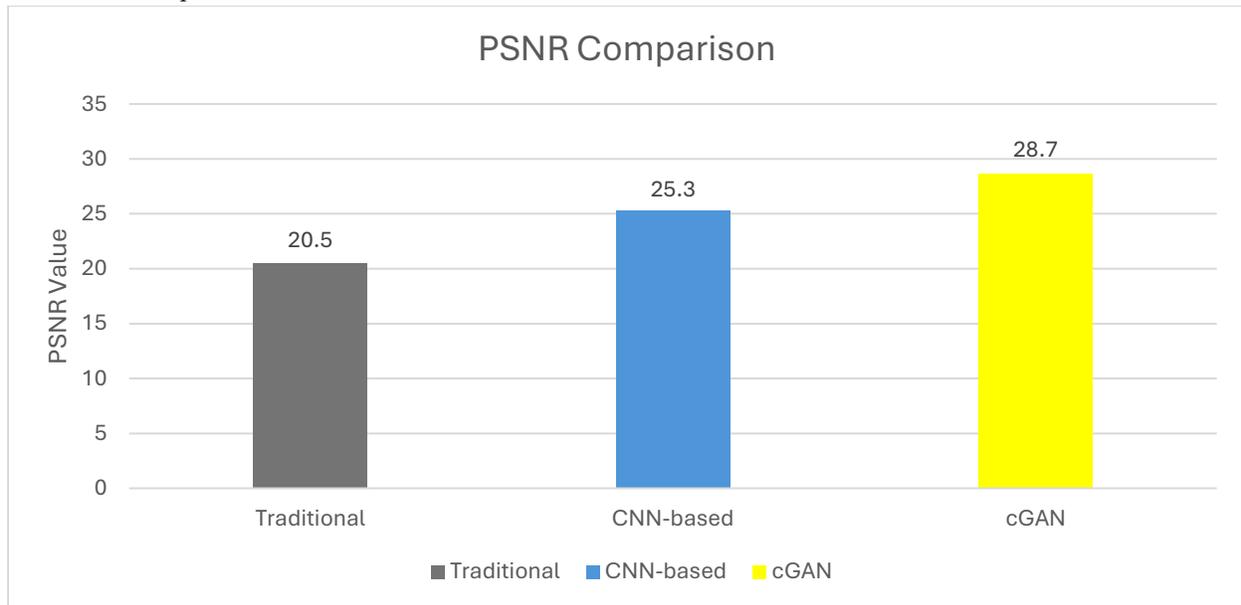
- PSNR (Peak Signal-to-Noise Ratio): Measures image reconstruction quality.
- SSIM (Structural Similarity Index Measure): Measures similarity to human visual perception.
- FID (Fréchet Inception Distance): Assesses visual quality using distributions in a feature space.

5. RESULTS

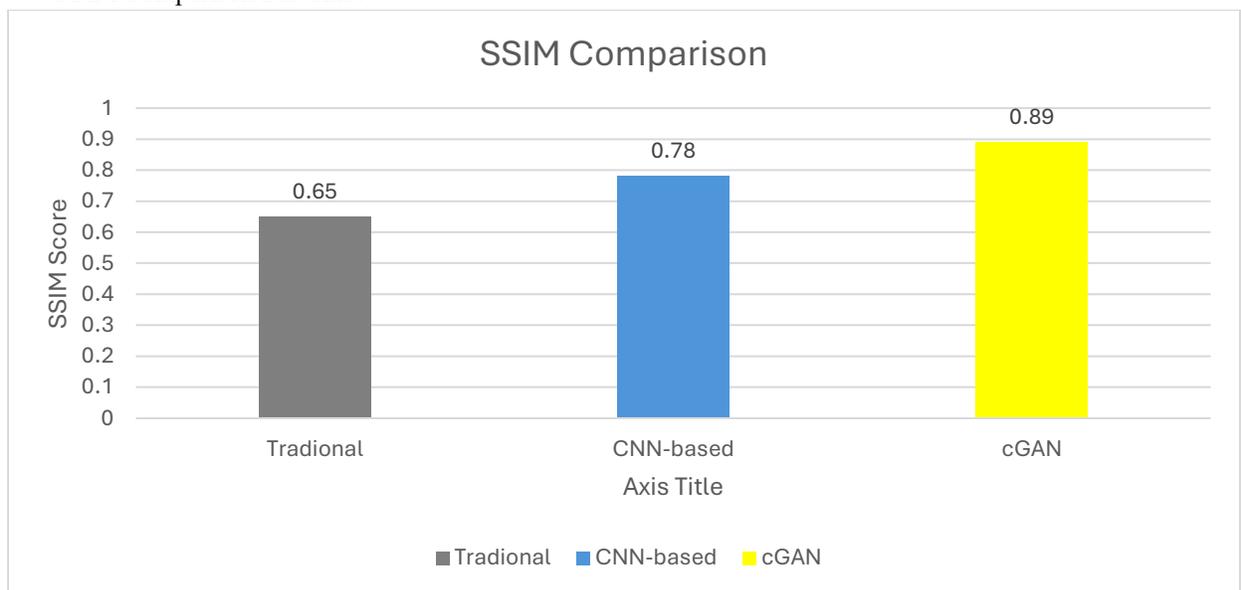
Quantitative Metrics Table:

Method	PSNR	SSIM	FID
Traditional	20.5	0.65	60.0
CNN-based	25.3	0.78	45.2
cGAN	28.7	0.89	30.5

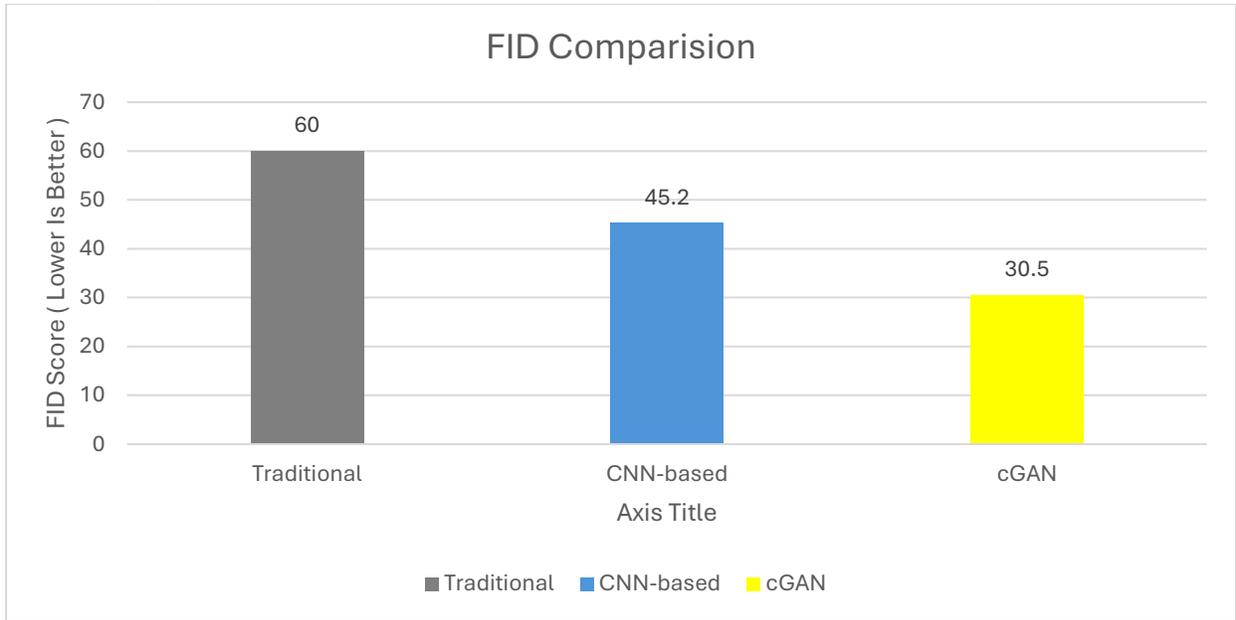
- PSNR Comparison Bar Chart



- SSIM Comparison Bar Chart



• FID Comparison Bar Chart



6. DISCUSSION

The data shows that cGANs outperform both traditional and CNN-based methods across all three metrics. While traditional methods yield lower PSNR and SSIM scores due to their reliance on heuristics, cGANs leverage adversarial training to produce results that are both structurally and visually superior. The FID score also highlights that cGANs generate more realistic and high-quality images. Additionally, cGANs are more flexible in adapting to different image types due to their deeper understanding of semantic context.

7. CONCLUSION

This study highlights the progress made in image colorization using deep learning, particularly the efficacy of Conditional GANs. With superior metrics and more realistic outputs, cGANs represent the current state-of-the-art in this domain. Nonetheless, challenges remain in training data diversity and computational demands. As technology advances, the potential applications of these models in video colorization, film restoration, and AI-assisted art tools will continue to grow.

8. FUTURE WORK

- Enhance semantic awareness using auxiliary networks.
- Integrate real-time processing for video and animation pipelines.
- Explore multimodal inputs like text guidance and reference colour images.
- Investigate the use of transformer architectures for global context understanding.

REFERENCES

- [1] Zhang, R., Isola, P., Efros, A. A. (2016). Colorful Image Colorization. ECCV.
- [2] Nazeri, K., Ng, E., Joseph, T. (2018). Image Colorization using GANs.
- [3] Larsson, G., Maire, M., Shakhnarovich, G. (2016). Learning Representations for Automatic Colorization.