iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: VII Month of publication: July 2026

DOI: https://doi.org/10.22214/ijraset.2026.84407

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

Gray to Glow: Colorize Grayscale Images Using Deep Learning

Mukashifa Fatima[1], Razia Begum[2]
[1]M. Tech Student of Department of Computer Science and Engineering, MJCET Hyderabad, Telangana, India
[2]Professor of the Department of Computer Science and Engineering, MJCET Hyderabad, Telangana, India

Abstract: The automatic colorization of grayscale images is a difficult computer vision problem because there are multiple possible colorizations for each grayscale image. To solve this issue, the Gray to Glow project aims to create an automatic deep learning-based system which transforms a grayscale image into a color image without the need for user scribbles or reference images, or manual intervention. The proposed system is based on the use of a CNN for color prediction in the CIE Lab color space, leaving the L channel unchanged and estimating the missing channels in a classification oriented way that minimizes the number of desaturated outputs while maximizing color vividness. It combines pretrained ECCV16 and SIGGRAPH17 colorization models, image preprocessing, output reconstruction, and it has a Flask web app to make it accessible for both command-line and browser. Experimental results presented in the thesis confirm that the system could produce visually meaningful colorized outputs, it was able to process outputs that were stable, and it was easy for users to use the system in enhancing grayscale images. The project illustrates how a deep learning model can be applied to develop a real-world and deployable system capable of image colourisation for image restoration, digital archiving and media enhancement applications.

A key point is that computer vision applications can be extremely useful in colorizing images that are previously only in grayscale.One of the interesting aspects is that an image, when it is only in grayscale, can be very helpful in computer vision for image colorization, and it can be done very efficiently with deep learning techniques.

Keywords: Image Colorization, Deep Learning, Convolutional Neural Network, Grayscale Images, Computer Vision, CIE Lab

I. INTRODUCTION

Although the color information is omitted from grayscale images, the restoration of grayscale images and enhancement of their appearance are still significant research issues in computer vision due to the fact that color images lose the appearance of visual richness, emotional depth and contextual detail when the color information is absent. The automatic image colourisation aims to reconstruct plausible colour information from the grayscale image it is given as input, with the goal of creating an image that is more realistic and informative in appearance than the original image. This problem is ill-posed, meaning that there is no unique solution to it, and many different colour assignments can result in the same luminance pattern, making it easy to get a dull or ambiguous result if one attempts to solve it by a regression-based method.

To solve this problem, the Gray to Glow project was created with a completely automatic deep learning system that is given a grayscale image and gives a photorealistic colored image as output, without any guidance from the user. A pipeline developed based on convolutional neural networks, trained to relate image structure, textures and semantic context to the most probable color distribution, and inferences made from the luminance channel to plausible chrominance values. In practical terms, this translates to assigning elements like sky, vegetation and skin visually convincing colors from a database of large number of natural images which are learned from these images.

Such systems are useful in several application areas. The thesis emphasizes applications for photo restoration, cultural heritage preservation, digital archiving, and/or visual media enhancement, where colorization can help make historical images more vivid and informative. The following analogy can be used for network intrusion detection, where the intent is to infer missing or hidden structure from complex input patterns, just like image colorization is about generating perceptually plausible visual content from incomplete image information, but unlike intrusion detection, where the inputs are traffic traces and classification of malicious actions, the inputs are images with missing parts. The comparison shows that in both domains, learned representations are used, but Gray to Glow is a generative representation of the visual enhancement system and not a security classification system .

The main contributions of this work are as follows:

1) Design and development of a deep learning based colourization system for fully automatic colourization of grayscale images using pretrained colourization networks.

2) Preprocessing, model inference, reconstruction and deployment (CLI and web). .
3) Implementation study based on literature which compares classical and modern colorization concepts and transforms them into a pipeline of a real implementation. .
4) Feasibility of functional testing, stable output supply and user-independence of grayscale enhancement.

This paper continues with the following outline. In Section II, we present a review of the related image colorization work. The proposed method and system are described in Section III. Implementation outcomes together with experimental results are discussed in section IV. Section V is the conclusion of the paper and projections for future directions.

II. RELATED WORK

The development of image colorization has progressed from user assisted and regression methods to advanced deep learning techniques that can handle automatic colorization. It is shown in the literature that direct regression in the color space did a more or less effective job of producing desaturated output for early deep learning systems, as often the correct result was the average of many possible color modes. This limitation motivated the transition toward classification-based formulations, generative models, and attention-driven architectures that better represent multimodal color uncertainty.

The study of semantic colorization of synthetic aperture radar imagery in 2026 proves that the domain of natural images has become far from the exclusive realm of colorization, and other domains have also entered the picture where semantic understanding and cross-modal alignment are important. The thesis also refers to research of comparative machine learning which demonstrated that semantic-guided and adversarial methods are better able to produce realism and preserve boundaries than simpler baseline methods. For Gray to Glow, these findings are significant in terms of the importance of scene understanding, consistency of color boundaries and perceptual quality in automatic image colourisation. .

Additionally, from the literature survey, three main families of contemporary image colorization techniques are detected: GAN, diffusion, and transformer. The GANs enhance photorealism by leveraging adversarial learning, diffusion models generate various and coherent colour distributions, and transformers capture global context and dependencies that are hard to represent by a conventional CNN. The thesis, however, reveals that the advantage of these gains is generally associated with increased computation needs, slower inference time and more complex deployment—particularly for real-time or web based applications—must be traded off.

The thesis gives a discussion of several specific models. Different decoders are employed in DDColor to enhance semantic color reasoning and preserve color boundaries; generative priors and diffusion modeling are utilized in BigColor for generating diverse output; the ColorFormer introduces memory-assisted hybrid attention in order to yield more consistent colors; autoregressive sequence modeling is employed in the Colorization Transformer to capture long-range relationships. Semantic plausibility and boundary fidelity are further enhanced by the use of semantic priors or instance segmentation for chromaGAN and instance-aware colorization, respectively, with higher architectural complexity and computational cost.

In contrast, Gray to Glow takes a more ‘how to do it' perspective. It does not design and develop a new architecture from the ground up but utilizes pre-trained models from ECCV16 and SIGGRAPH17 and puts them into a practically usable end-to-end system pipeline. This selection makes the project a deployable research implementation which compromises visual quality, speed and accessibility for grayscale image enhancement tasks.

III. METHODOLOGY

The methodology adopted for the development of the proposed grayscale image colorization system is described in this section, which encompasses the assumptions made on the dataset, the pre-processing steps, the method of integration of the models, the flow of the system, and the deployment structure.

A. Overview of the Proposed System

The architecture of Gray to Glow is general enough to be able to take both grayscale and RGB images to preprocess them into a model compatible format, apply a pretrained deep colorization model, and output reconstructed color images to the end user via command-line or browser-based interaction. The thesis organizes the system as a modular pipeline, where the image is input, the model is loaded, it is preprocessed, the color is predicted and then the output is generated and delivered using a flask web interface. Modular design to help partition concerns over implementation phases and facilitate practical usability.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

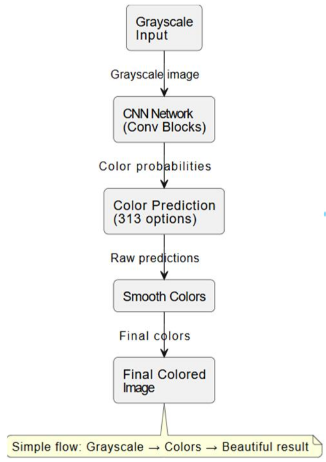

Fig. 1: System Architecture

B. Image Representation and Preprocessing

The system adopts CIE Lab color space as the basis of the colorization system. In this representation, the input image is represented by the luminance (L) channel with the missing chrominance channels (a and b) included. In the pre-processing stage, the gray information is processed as the input of the model and during the training process, the network learns to infer the chromatic information while preserving the original structure and intensity information.

The paper also mentions that pre-processing encompasses loading of images, image resizing, conversion to model-ready tensors, and normalization for inference with pretrained ECCV16 and SIGGRAPH17 models. The key part of this stage is that the colorization model is fed in a format that is compatible with the pretrained architecture, and the output reconstruction should maintain the luminosity details of the original image.

C. Core Colorization Strategy

The core novelty of the method in Gray to Glow is to view color prediction as a classification problem rather than just simple regression. This is because regression-based methods typically yield an average or washed out color since many possible colors for a pixel are combined into a mean color estimate. A classification-based formulation, on the other hand, enables the network to reason about multiple bins of color predictions and select more colorful and semantically coherent ones.

The main methodological inspiration for the project comes from ECCV16, which employs a VGG-style convolutional network with dilated convolutions and class rebalancing and annealed smoothing to predict plausible colors from a luminance input. This is the main method that has been identified in the thesis as being the most practical and balancing between fast speed, automatic operation and realistic output generation. Moreover, the project is also inspired by SIGGRAPH17, specifically by using a broader context of the scene and finer context details to achieve a more coherent color assignment.

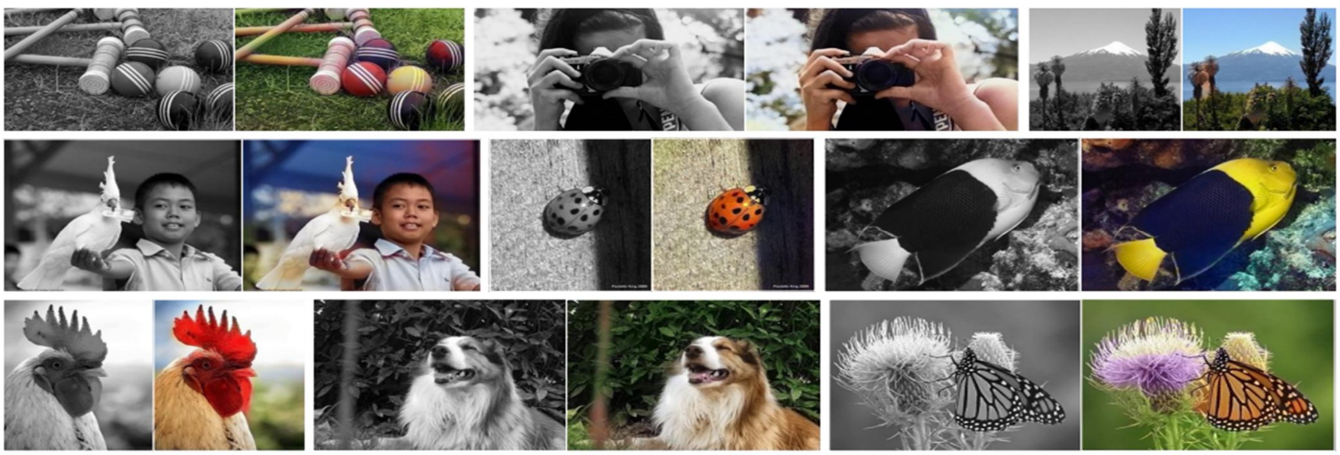

Fig. 2: Example input grayscale photos and output colorizations.

D. Implementation Overflow

The implementation is divided into eight modules:

1) Input module for selecting source images using command line interface .
2) Model loading module for loading pretrained colorization network.
3) Image preprocessing module for format conversion and tensor preparation.
4) Color Prediction Module for Outputting Chrominance Prediction.
5) Module for reconstructing final color images.
6) A web application module that is used to gain access to and interact with the browser.
7) A server-side inference web based colorization processing module.
8) A module that accepts a response and provides output of generated results.

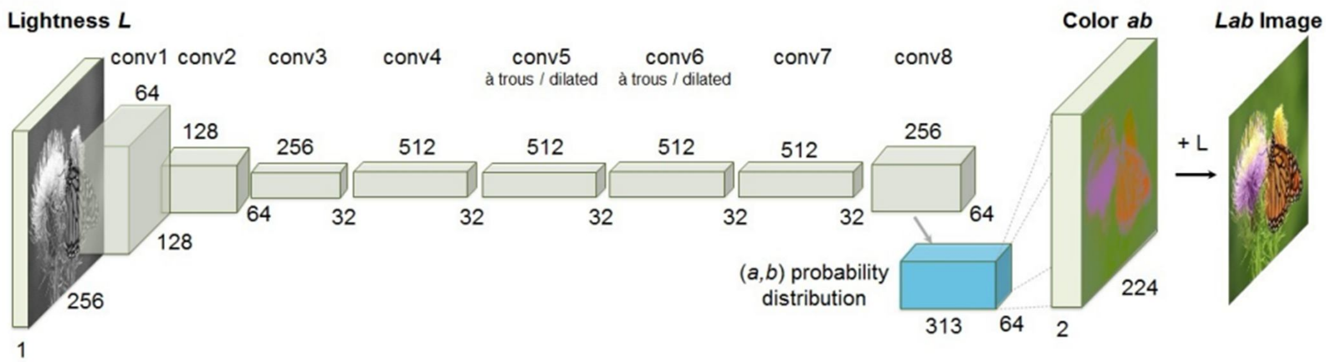

Fig. 3: Proposed CNN Architecture

The proposed implementation uses this modular decomposition which enhances the maintainability of the system and helps future processing of batches, mobile deployment and integration with newer colorization architectures.

Standardization of input is done by preprocessing, inference is done with pre-trained deep models, and post-processing is used to reconstruct color images that can be downloaded and displayed by the user. This practical implementation shows how research level colourization techniques can be incorporated into an accessible research software system that can be used for academic demonstration and applied experimentation. .

E. System Requirements and Practical Design

The usability and practicality of image upload, automatic colourisation, real-time processing, batch processing, and result export. Non-functional requirements are those that span the aspects of performance, scalability, reliability and usability that are set to assist in practical deployment. The software stack is based on python, deep learning libraries and flask or a similar web framework, and the hardware requirements are not too demanding for a typical computing environment. This is in line with the project goal of creating a research prototype that is accessible, rather than a strictly optimized lab benchmark.

IV. EXPERIMENTAL RESULTS

Overall, the paper shows the system provided realistic and stylish outputs for a wide range of grayscale inputs, particularly in cases where the semantics of the scene and the context of the objects were sufficiently explicit for the pretrained models to guess an object's color. The qualitative results revealed better color vividness and naturalness in the colorization results than the older washed-out colorization methods as reported in the literature survey. The system was also reported to be efficient and easy to use via the web-based interaction pipeline.The color distribution of natural-image datasets is very unevenly distributed, with rare but important colors being underrepresented when it comes to training. To overcome this, the ECCV framework assigns a weight to the loss based on the empirical rarity of the color, to make the model produce more vibrant and diverse color output. The vibrancy of modal decoding is then balanced with the consistency of mean decoding to get final continuous ab estimates through an annealed-mean operation from the predicted distribution. The paper states that 0.38 is an optimal compromise for visually realistic output.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

Table I shows that the ECCV16 model results in visually appealing colourized images, and has a sufficiently rapid inference time for real-time applications. Half of the human evaluators (32.3%) accepted a colorized image as real on the colorization Turing Test, suggesting that nearly one third of the colorized images fooled humans during this Turing Test.

TABLE I
COMPARISON OF MODELS

Model	**Architecture**	**Color Prediction**	**Color Vibrancy**	**Real-Time Suitability**	**Scene Understanding**
ECCV Model	Single CNN	Classification	High	Excellent	Good
SIGGRAPH Model	Dual DNN	Regression	Moderate	Good	Very Good

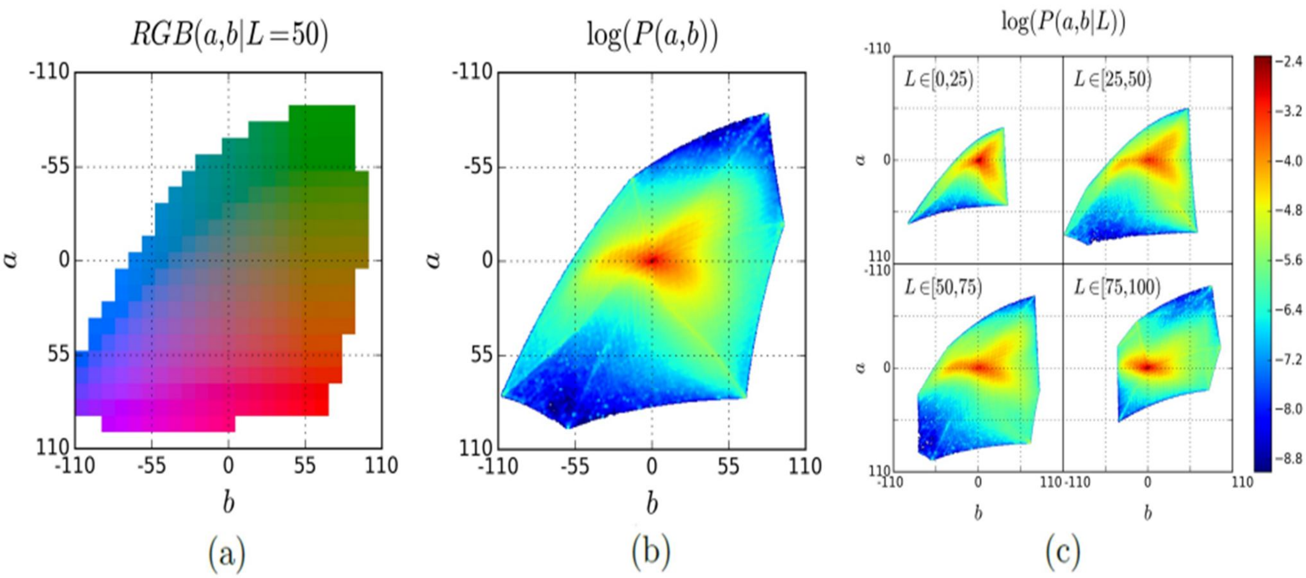

Fig. 4: (a) Quantized ab color space with a grid size of 10. A total of 313 ab pairs are in gamut. (b) Empirical probability distribution of ab values, shown in log scale. (c) Empirical probability distribution of ab values, conditioned on L, shown in log scale.

The color space visualization and color space probabilities distribution of color for image colorization are illustrated in Figure 4. Panel (a) is the quantized aaa-bbb color space at a fixed lightness (L=50L=50L=50); panel (b) is the log probability distribution of color. The probability distributions shown in Panel (c) illustrate the distribution of color frequencies over lightness (LLL) ranges, indicating the luminance-dependent differences in color frequencies.

Figure 5 presents the results for Classification accuracy Top 1 on the ILSVRC2012 data set for different convolutional layers. Our proposed method performs competitively with the state-of-the-art feature learning techniques, and is effective for image colorization. The results show that the semantic features are extracted better, especially in deeper convolutional layers which result in the more accurate and realistic color prediction.

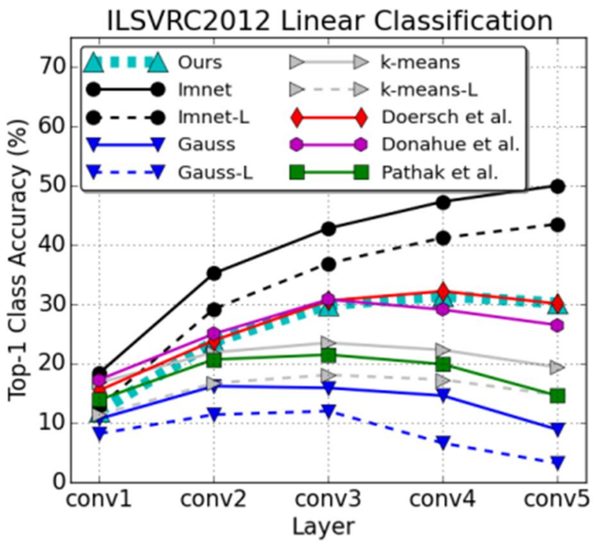

Fig. 5: ECCV Model Linear Classification

Top-1 classification accuracy on Places365 on various convolutional layers (figure 6). Results showcase the effectiveness of learned feature representations for scene recognition, with the best results obtained in deeper convolutional layers, and the possibility to train scene-aware colorization models on the Places365 database.

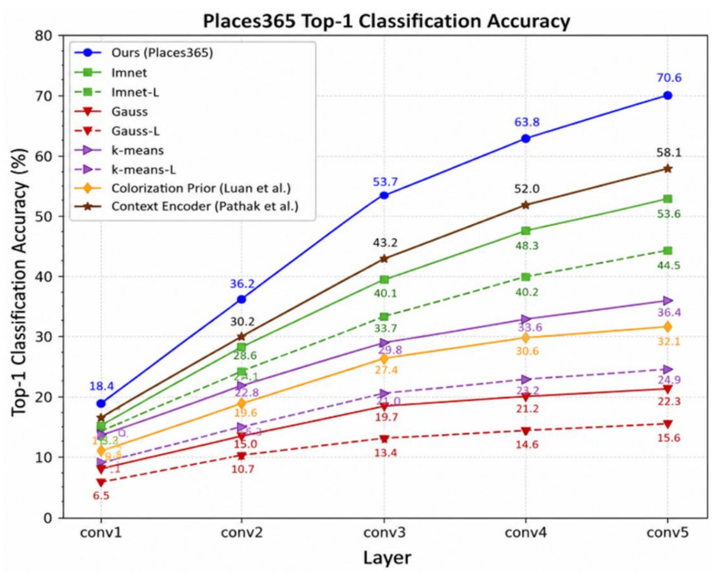

Fig. 6: SIGGRAPH Model Classification Accuracy

Comparative Analysis

- Unlike other approaches that generate dull or desaturated colors, ECCV's classification-based approach with class rebalancing allows the model to produce vibrant and realistic colors.
- SIGGRAPH uses two different feature extraction networks: one global and one local, which enables it to better account for the overall scene context and also maintain the fine image details.
- ECCV does not require multiple feed-forward CNNs to make the color prediction, hence the inference is faster.
- Generally, SIGGRAPH will give smoother transitions and better contextual consistency, but it will use more resources to do so.
- Both models generate plausible colorizations; however, ECCV is more suitable for web-based applications due to its lower computational cost and quicker processing.

The ECCV16 model is able to generate plausible color distributions and colorize the grayscale image into a realistic image as shown in Figure 7. The resulting image is rich in colors and has a good contrast, capturing the details of the butterfly and flower. The model successfully represents semantic information, giving natural colors of orange, black and yellow to the objects. In general, it shows the potential of the model in generating visually appealing and context-aware colorized images.

ECCV16

Fig. 7: ECCV Model Colorization Output

SIGGRAPH17

Fig. 8: SIGGRAPH Model Colorization Output

The SIGGRAPH17 model produces the realistic and visually coherent colorized image in Figure 8 using both global context of the scene and local image features. Color transitions are smooth and object outlines and fine structural details are retained. The butterfly and flower are correctly colored using natural colors, making the image more realistic and consistent in terms of context. In general, the SIGGRAPH17 model can colorize the image in a high quality, and its colorized result is scene-aware and has a realistic and balanced color representation. The experimental results presented above indicate that the proposed Gray-to-Glow system is able to colorize grayscale images into realistic colors with good execution time. ECCV16 offers a superior combination of speed, color gamut and usability and is thus the model of choice for the application. In addition, SIGGRAPH17 enhances the system through its additional features, which provide more consistent and sophisticated transition colors and scene-level consistency. The two models can be used together to compare different colorisation outputs and to choose the most pleasing colourisation for a particular grayscale image.

V. CONCLUSION AND FUTURE WORK

This paper revisits the Gray to Glow study for automatic image colorization of gray images with deep learning. The biggest advantage of the project is the practical end-to-end system that integrates the work of CIE Lab preprocessing, the pretrained models for the colorization with the ECCV16 and SIGGRAPH17 competition, the modular inference stages and the deployment for the user via command-line and web interfaces. The reported results demonstrate that the system is working correctly, stable, and can produce meaningful results in the process of enhancement of grayscale images in the form of colorized output results.

The literature analyzed in this, reveals that with CNN-based image colorization, the study has evolved from regression-based models to the GANs and diffusion models, which present some pros and cons regarding the realism, diversity, and computational complexity of colorization.The literature reviewed here demonstrates the shift from regression-based CNNs towards GANs, diffusion models, and transformer-based approaches, each with its own set of trade-offs in terms of realism, diversity, and computational efficiency. As such, Gray to Glow should be interpreted as a contribution to research in implementation, a search that is substantive and geared toward practice rather than a new learning architecture.

There are a number of areas where future work is possible. This system can be further improved by quantitative assessment on benchmark images, incorporation of more recent transformer- or diffusion-based colourisation models, better processing of less crisp historical photos, and larger perceptual user studies. Other topics that could be covered include mobile deployment, edge optimization, and other domain-specific colorization settings, including archives, medical imaging, or remote sensing.

REFERENCES

[1] Shivaprasad Satla, “A Deep Learning Framework for Semantic Colorization of Synthetic Aperture Radar Images” , in Springer Conference, Volume:84, 2026.
[2] Jinyi Luo, Xi Li, “Analysis and Comparison of Image Colorization Based on Machine Learning with PyTorch and ChromaGAN”, science direct conference, 2025.
[3] Muhammad Shafiq, Muhammad Asad, Asra Aslam, "A Comprehensive Review of Modern Image Colorization", in Springer Volume:84, 2025.
[4] Richard Zhang, Jun-Yan Zhu, Phillip Isola, Alexei A. Efros, "Transforming Color: A Novel Image Colorization Method", in Science Direct Volume:248, 2024.
[5] “DDColor: Towards Photo-Realistic Image Colorization via Dual Decoders”, Sungmin Kang, Myeongsu Kang, Donghyeon Cho, Sang Min Yoon, In So Kweon, 2023.
[6] Arash Vahdat, David J. Fleet, "BigColor: Colorization using a Generative Color Prior for Natural Images", in springer nature link, 2022 .
[7] Haojun Xu, Bing Li, Jianmin Jiang, "ColorFormer: Image Colorization via Color Memory Assisted Hybrid-Attention Transformer", International conference paper link, in 2022.
[8] Aditya Kumar, Richard Zhang, Jun-Yan Zhu, Colorization Transformer, in 2021.
[9] Miguel Perelló Nieto, Javier Ventura, Vitoria Pereira, "ChromaGAN: Adversarial Picture Colorization with Semantic Class Distribution", in 2020.
[10] Chaoqun Su, Weidi Xie, Tao Kong, "Instance-Aware Image Colorization", IEEE Conference, in 2020.
[11] Sukrut Oak, Young Chol, Song Da Sun, “Improved Recoloring of SAR Satellite Images, ” Stanford University, vol. 141, p. 112963, 2020, doi: 10.1/16.eswa.2019.112963.
[12] Kaushal Kishor, Chirag Sharma, Himanshu Sharma, Manmohan Fulara & Aditya Yadav, “SAR Image Colorization for Comprehensive Insight Using Deep Learning”, in 2025, springer, chapter/10.1007/978-981-96-3725-6_22
[13] Cheng, Z., Yang, Q., Sheng, B.: Deep colorization. In: Proceedings of the IEEE International Conference on Computer Vision, 2015, 415–423.
[14] Zhang, R., Isola, P., and Efros, A. A. *Colorful Image Colorization*. European Conference on Computer Vision (ECCV), 2016.
[15] Wang, X., Gupta, A.: Unsupervised learning of visual representations using videos. In: Proceedings of the IEEE International Conference on Computer Vision. 2015, 2794–2802.
[16] Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected arXiv preprint arXiv:1606.00915, 2016.
[17] Patterson, G., Hays, J.: Sun attribute database: Discovering, annotating, and recognizing scene attributes. In: Computer Vision and Pattern Recognition (CVPR), 2022 IEEE Conference on, IEEE 2751–2758.