Grayscale Picture Colorizer using Generative Adversarial Networks(GANs)

D Chandra Mouli[1],Datla Sai Mowli[2], Giri Venkata Laasya[3], Putsala Harsha Priya[4], Yelamanchili Thrushith[5], Shaieen Begum[6],

[1]mouli227@vignaniit.edu.in, [2]dasaimouli@gmail.com, [3]laasyagirivenkata@gmail.com, [4]putsalaharshapriya@gmail.com, [5]thrushith2003@gmail.com, [6]shaieenbegum9@gmail.com

[1,2,3,4,5,6] Department of AI&DS, Vignan's Institute of Information Technology, Visakhapatnam, Andhra Pradesh, India

***Abstract*-This project Grayscale Picture Colorizer using Generative Adversarial Network, introduces a model forcolorizing black and white images into vibrant, color-enhanced pictures. The system uses deep learning techniques, using a Generative Adversarial Network for image-to-image translation. The main objective is to bring historical or monochromatic visuals to life, enhancing their aesthetic appeal and providing a new perspective on the past. This project shows importance of not only the restoration of historical content but also the enhancement of contemporary visual media through the creative application of color. Throughout the development, optimization techniques are explored to improve real-time performance, making this Image Colorizer System versatile for various applications.**

***Keywords-* Generative Adversarial Network (GAN), Image Colorization, Grayscale picture, Generator, Discriminator**

I. INTRODUCTION

The improvements in deep learning in recent years, especially in the domain of generative models have paved way to exceptional breakthroughs and achievements in various image processing tasks. Grayscale image colorization is one such task, which works to automatically colorize black and white images. Black and white images, while timeless, can't capture the full vibrancy of the world. This is why automatic colorization of grayscale images is a significant and ongoing challenge in computer vision. By adding color, researchers aim to unlock a deeper understanding and appreciation of these historical or artistic photographs. Traditional methods for coloring an image often depend on handcrafted rules or need immense human intervention, reducing their scalability and effectiveness. Deep Learning breathes new life into black and white photos. The field of grayscale image colorization has been revolutionized by the power of deep learning, especially Generative Adversarial Networks (GANs). These innovative techniques have led to substantial advancements in this area.The causes stating the necessity of this project come from the increasing necessity for automated image processing tools across various domains, including digital media, entertainment, and heritage preservation. There is a need for efficient algorithms that can improve visual content while reducing manual intervention with the rapid growth of digital content creation and consumption. This project addresses this need and supports users to produce vibrant and captivating images effortlessly by enabling automated grayscale picture colorization, thereby facilitating the generation of richer visual experiences in diverse applications and industries. We have used Image Colorization Dataset (https://www.kaggle.com/datasets/aayush9753/image-colorization-dataset/data)

II. LITERATURE SURVEY

Harshit B.,[1] created a method to add color to low-resolution images. This system utilizes a specifically designed type of Generative Adversarial Network (GAN) called a Conditional Deep Convolutional Generative Adversarial Network (DCGAN). Implemented using PyTorch, a deep learning framework, this approach tackles the colorization task by providing the grayscale image itself as a reference, eliminating the need for random noise introduction..

Richard Zhang et.al,[2] introduced a deep learning system for image colorization that incorporates user input. Unlike traditional methods with pre-defined rules, this system utilizes a Convolutional Neural Network (CNN) to directly convert a grayscale image, along with user-provided color "hints," into a colorized version. The CNN effectively combines user edits with both low-level details and high-level semantic information, which it learns from a vast dataset. This approach allows users to guide the colorization process for more customized results.

Z Cheng et.al,[3]proposed a system focused on achieving high-quality colorization entirely through automation. The authors hypothesized that a perfect patch matching technique, achievable with a large dataset of color images, could be the ideal solution for colorizing black and white images. They implemented this concept directly using deep learning methods. Additionally, they suggested incorporating a post-processing step that utilizes a technique called joint bilateral filtering.

Saeed Anwar et.al,[4] provided a thorough analysis of cutting-edge deep learning techniques used for image colorization. The authors examined the core building blocks of these methods, including their architectures, inputs, and training processes. They delved into the specific optimization algorithms, loss functions, and training data employed by each technique. Furthermore, they categorized existing colorization methods into seven distinct classes and explored the key factors influencing their performance.

Kaiming He et.al,[5] introduced a new approach called residual learning to simplify the training of much deeper neural networks than previously possible. The core idea involves reformulating the network layers to learn the differences between the input and the desired output, rather than learning the entire function from scratch. This significantly improves the training process for complex networks.

Liang-Chieh Chen et.al[6], explored using Deep Learning for image segmentation, a technique that separates objects within an image. The authors focused on a specific convolution method (without downsampling) and introduced a novel technique called atrous spatial pyramid pooling. This approach, which combines elements from deep convolutional neural networks and probabilistic graphical methods, improves the accuracy of object boundary detection in the segmentation process.

Richard Zhang et.al[7] introduced a method for automatic image colorization that generates vivid and natural-looking colors. To address the inherent ambiguity of colorizing black and white images, the authors approached the problem as a classification task. During training, they employed a technique called class-rebalancing to ensure a wider variety of colors appeared in the final results.

Alec Radford et.al,[8] proposed a new type of Convolutional Neural Network (CNN) architecture called Deep Convolutional Generative Adversarial Networks (DCGANs). These DCGANs were designed to address the limitations of using standard CNNs for unsupervised learning tasks, where labeled data is scarce. The researchers implemented specific architectural constraints within the DCGANs and demonstrated their effectiveness for unsupervised learning. This suggests that DCGANs are a promising approach for applications that lack large amounts of labeled data.

Ian J. Goodfellow et.al,[9] introduced a novel framework for training generative models. It relies on an adversarial process where two models are trained simultaneously: a generator (G) and a discriminator (D). Unlike some previous methods, this approach doesn't require complex techniques like Markov chains or unrolled approximate inference networks during either the training phase or when generating samples. This can potentially simplify the training process and make it more efficient.

Olaf Ronneberger et.al[10], introduced a new approach for image analysis tasks that leverages a powerful data augmentation strategy. This strategy allows the network to effectively utilize a limited set of labeled data. The network architecture itself is comprised of two key parts: a contracting path that extracts contextual information from the image and a mirrored expanding path that enables the network to pinpoint specific details with high accuracy.

Alex Krizhevsky et.al,[11] explored training a generative model with multiple layers. This model is designed to learn and capture features that are similar to those processed by the human visual cortex. To achieve this, the authors developed a new parallelization algorithm. This algorithm allows the training process to be distributed across multiple machines connected in a network, significantly improving efficiency and potentially enabling the training of even more complex models.

Alian Horé et.al,[12] investigated two common methods for measuring image quality: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM). The authors were able to establish a straightforward mathematical connection between these two metrics. This relationship proved effective in evaluating image quality even under various types of image degradation.

III. METHODOLOGY

This project uses a GAN architecture. The architecture consists two important components: Generator network G and Discriminator network D. Our methodology for this project comprises four main sections. Generator, feature extractor, discriminator and the working of the generator and discriminator.

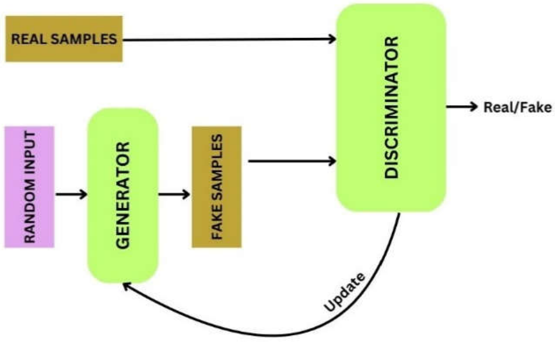


Fig. 1 Generative Adversarial Network

A) Generator

The generator network G is a Convolutional Neural Network (CNN) taking grayscale images as input and gives out the colorized versions of the same images as outputs. The generator aims to store semantic information and capture complex color distributions while learning to map from grayscale images to color images. Here, we have used ECCV16 architecture as generator.

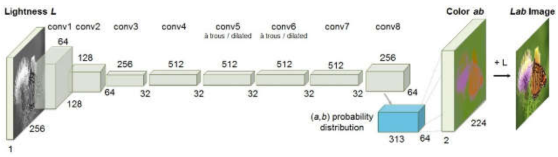


Fig. 2 ECCV16 Architecture [21]

B) Feature Extractor

We are using a pre-trained VGG19 model as the feature extractor, a novel approach to compare images and use that comparison to calculate the loss. Two images are passed in it and the activations of the 18^{th} layers are taken for both the images and then these activations are used to compute the loss using RMSE, MSE etc., between the two activations.

C) Discriminator

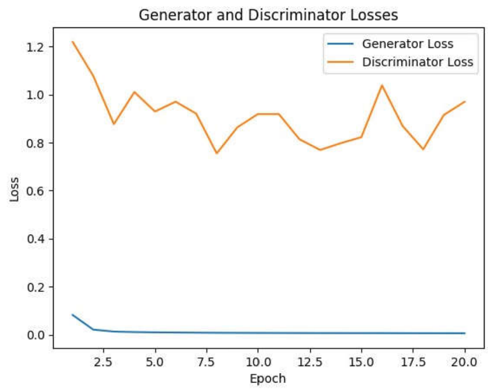
The discriminator network, denoted by D, acts like a critic in this system. It's a Convolutional Neural Network (CNN) built using convolutional and pooling layers. Its role is to distinguish between real color images and the colorized images generated by the opposing network, the generator (G).It learns features from real and generated images that are used in discrimination of the real and fake images. It outputs a probability score that shows the likelihood of each image being real by taking both the generated colorized images and real color images as inputs.

D) Loss Functions

A loss function measures the goodness of a neural network model while performing a particular task. We strive to minimize the value of the loss for the neural network to perform better. The following are the types of loss functions used in this model.

1. Adversarial Loss (GAN Loss)

The generator network is trained using adversarial loss. This loss function helps the generator create realistic colorized images that can trick the discriminator network into thinking they're real photographs.



2. Content Loss

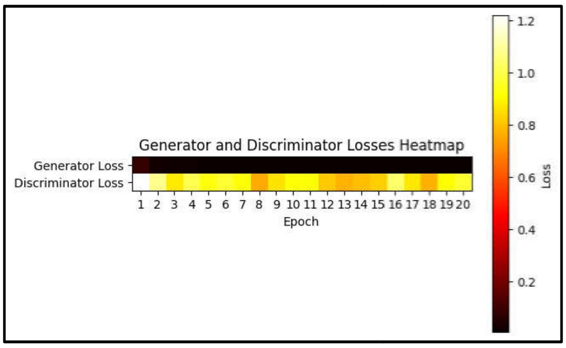
In addition to adversarial loss, the system employs a concept called content loss. This loss function evaluates how closely the features extracted from the generated colorized images resemble those extracted from real color photographs (ground truth). It achieves this by comparing the feature maps obtained from a pre-trained network called VGG19, using a method known as L1 loss.

3. Total Generator Loss

The generator's performance is measured by a combined loss function. This function incorporates both content loss and a weighted adversarial loss.

4. Discriminator loss

The discriminator's success in distinguishing real from fake images is evaluated by a loss function called discriminator loss. This loss is calculated differently for real photographs and the colorized images generated by the model (often referred to as "fake" images) to guide the discriminator's learning process.



E) Approach

A Battle for Better Colorization

Imagine a system where two neural networks compete to improve each other's skills. This is essentially what happens during Generative Adversarial Network (GAN) training for image colorization.

The Artist (Generator): A network named the generator (G) acts like an artist, constantly trying to create the most realistic colorized images possible. It studies real color photographs and aims to produce images that are indistinguishable from the originals.

The Critic (Discriminator): Another network, the discriminator (D), plays the role of a tough critic. Its job is to analyze both real and generator-created colorized images and accurately determine which is which.

The Cycle of Improvement

Both networks are trained simultaneously:

The Generator's Turn: The generator creates a batch of colorized images.

The Discriminator's Test: The discriminator examines these images alongside real color photos and tries to classify them correctly (real or fake).

Learning from Mistakes: Based on the discriminator's feedback, the generator refines its approach to create more convincing colorizations in the next round.

The Discriminator Adapts: The discriminator also learns from its errors. As the generator improves, the discriminator needs to sharpen its critical skills to stay ahead.

Loss Functions Guide the Way

Throughout this training process, special functions called loss functions evaluate how well each network performs. These losses guide the adjustments made to the networks' internal parameters, like weights and biases.

This adversarial training loop continues until both networks reach a point where the generator can create colorized images that consistently fool the discriminator. This is when the generator is considered to be a skilled colorization artist!

The working model of the project is shown in Fig. 2

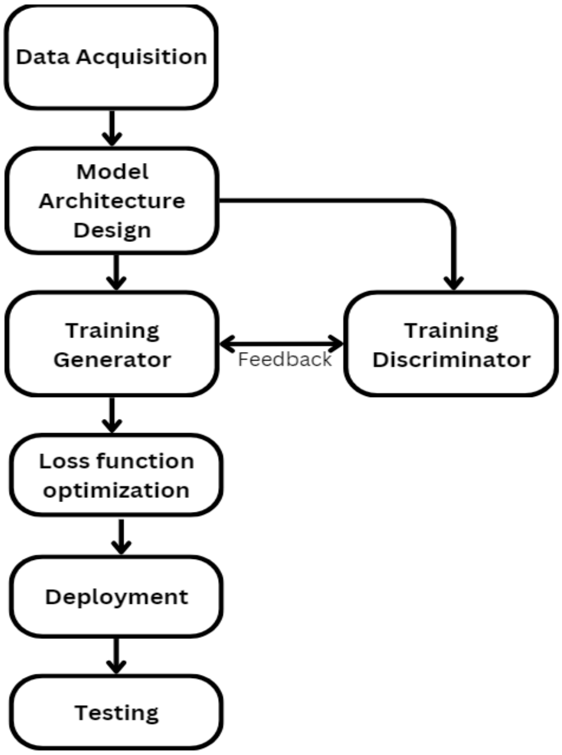


Fig. 2 Working method

IV. RESULTS

We understood that Image colorization can be used in various aspects of real-world. The results of our testing and implementation are as follows.

A) Qualitative Assessment

1. Visual Inspection

The colorized images achieved a remarkable feat - they were practically indistinguishable from real photographs when compared to the original color versions. This showed great potential of the system.

2. User Feedback

The system has been put to use in real time and the users were satisfied with the colored images generated. The images had an aesthetic appeal to them with enhanced colors.

B) Real-World Applications

1. Historical Image Restoration

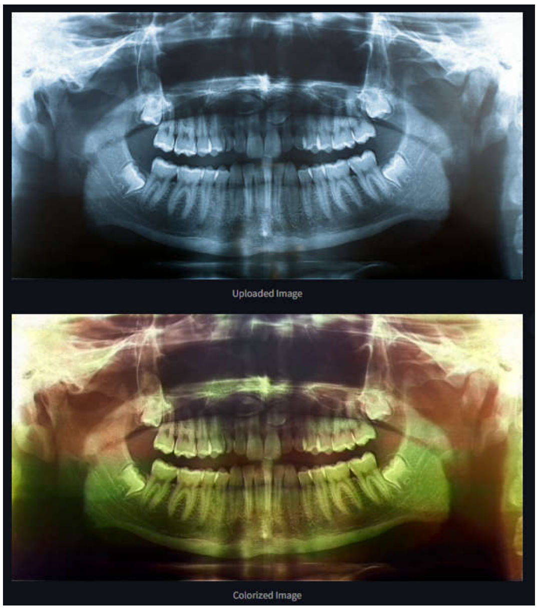
The system has been used to colorize historical images to enhance their visual appeal.

2. Digital Media Production

The system was used to colorise images to add artistic effects and arouse specific mood.

3. Medical Imaging

The system produced good results for medical imaging like X-rays.

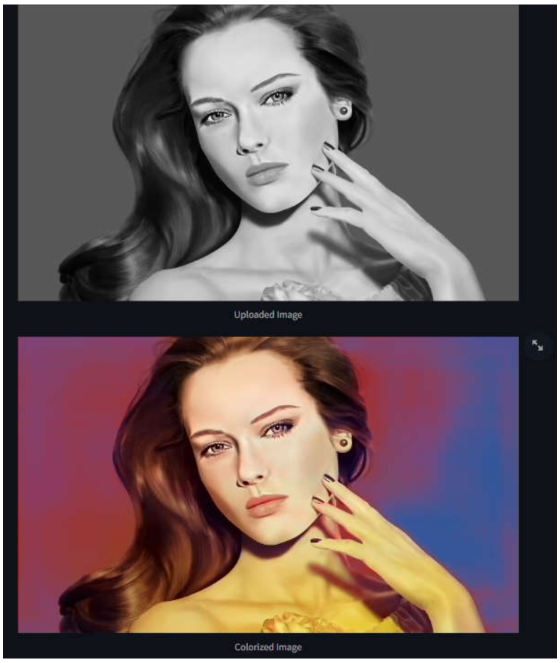


4. Forensic Analysis

CCTV footage was also colorised using the system. Colorised images can be used to identify suspects and gather crucial evidence.

5. Digital Art and Design

Grayscale portraits were also colored that added vibrancy to the sketch.

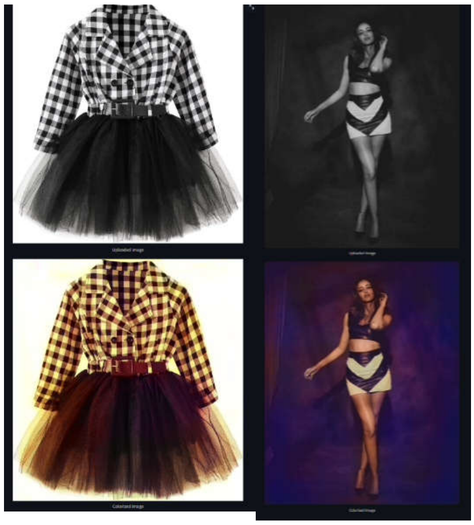


6. Culture

Image colorization of historical images with cultural history preserves the importance of culture and gives a vague idea of the history and its culture.

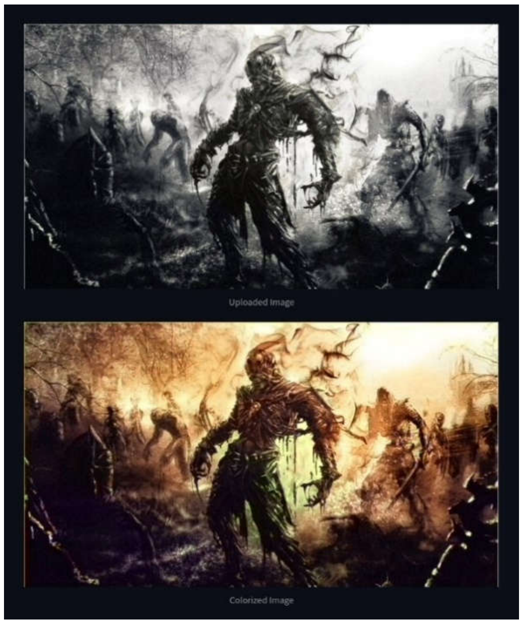
7. Fashion and Textile

Adding colors to fashion and textiles can be helpful to preview patterns, colors and designs.



8. Gaming

Adding colors to gaming scenarios would evoke and create artistic effects in which users would be indulged.



C) Generalization Performance

This system was used on unseen images, that were not used before for training of the model. It produced excellent and intriguing results. All the above depicted results were performed on unseen images. As we can see, the results were satisfying.

V. CONCLUSION & FUTURE WORK

In conclusion, this project Grayscale Picture Colorizer using Generative Adversarial Networks has depicted excellent improvements in automating the processof colorising grayscale images. We have achieved remarkable results in generating realistic and visually appealing colored images through the training and development of a GAN-based model resulting in the enhancement of overall quality and also the aesthetic appeal of grayscale images. The methodology of the project, combining adversarial training, perceptual loss and the design of network architecture has demonstrated good results in multiple quantitative and qualitative evaluation metrics. This model holds great potential for applications in photo restoration, digital art and visual content generation. It offers efficient solution for enhancing grayscale images with vibrant colors.

For further improvement of this model, we would want to develop a multi-modal and interactive colorization approach that would produce diverse colorizations for the given grayscale image and videos which would allow users to explore different colorization preferences and possibilities by providing guidance or constraints for personalized results.

REFERENCES

[1] Harshit Bansal. Image colorization. https://github.com / harshitbansal05 / Image - Colorization, 2018.

[2] Richard Zhang, Jun-Yan Zhu, Phillip Isola, Xinyang Geng, Angela S Lin, Tianhe Yu, and Alexei A Efros. Real-time user-guided image colorization with learned deep priors. *ACM Transactions on Graphics (TOG)*, 9(4), 2017.

[3] Zezhou Cheng, Qingxiong Yang, and Bin Sheng. Deep colorization. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 415–423, 2015.

[4] Saeed Anwar, Muhammad Tahir, Chongyi Li, Ajmal Mian, Fahad Shahbaz Khan, and Abdul Wahab Muzaffar. *Image colorization: A survey and dataset. arXiv preprintarXiv:2008.10774*, 2020.

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[6] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L. Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, 2017.

[7] Richard Zhang, Phillip Isola, and Alexei A. Efros. Colorful image colorization. In Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, editors, *Computer Vision – ECCV 2016*, pages 649–666, Cham, 2016. Springer International Publishing.

[8] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks, 2016.

[9] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, DavidWarde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *arXiv preprint arXiv:1406.2661*, 2014.

[10] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation,2015.

[11] Alex Krizhevsky. Learning multiple layers of features from tiny images. 2009.

[12] Richard Zhang, Phillip Isola, Alexei A. Efros, Colorful Image Colorization. In ECCV, 2016.