

Enhancing Black-and-White Image Colorization Using Cycle GAN: A Study on Unpaired Image Translation

Hengrui Ren

School of Science, Rensselaer Polytechnic Institute, NY, USA

renh5@rpi.edu.cn

Abstract. This paper explores the application of Cycle-Consistent Generative Adversarial Network (Cycle GAN) for the colorization of black-and-white images, addressing the challenges in image to the image translation and it was not requiring paired datasets. Cycle GAN, a generative adversarial network, uses cycle consistency loss to map grayscale images to color, effectively preserving structural integrity. The model is trained on unpaired datasets containing black-and-white and color images, enabling it to generate colorized versions from grayscale inputs. Performance evaluation is conducted using Two significant evaluation metrics: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). Results demonstrate that Cycle GAN effectively preserves image structure and detail, achieving high PSNR and SSIM values, particularly in scenes with distinct color boundaries. However, challenges arise in complex and low-resolution scenes where color accuracy and detail reconstruction may degrade. Additionally, this paper assesses the impact of different loss functions, learning rate schedules, and training strategies on the model's output quality. Future work will focus on addressing these limitations by incorporating attention mechanisms and other advanced techniques to improve Cycle GAN's ability to handle intricate image features and complex scenes.

Keywords: Cycle GAN, Image Features, Unpaired Datasets.

1. Introduction

Since the early 19th century, when photography was first invented, black-and-white pictures have become an important tool for documenting the world [1]. In the long history of photography, black-and-white pictures have held an very important position. The limitation of color makes researchers to miss out on a lot of the details in the images. Although the black and white images will bring researchers more widely space to have unlimited imagination, but colorization the black and white images can help people closer to the story what the image tell to researchers, let their ideas reach into the photo. Color is not only visual information, but also a key to conveying emotions and atmosphere [2]. Therefore, the study of black and white photo colorization is not only technically challenging, but also breathes new life into historical photos, cultural heritage, and personal memories [3].

In the recent years, because of the development of computer and science technology, image and image translation tasks have received widespread attention, especially with the advancement of deep learning and Generative Adversarial Networks (GANs), which have led a very important technology progress in the image generation. Cycle-Consistent GAN (CycleGAN), one of the most well-known and efficiency methods for solving image-to-image translation, allows transformation between two image

sets are not required for paired training datasets, driving significant progress and innovation in the image processing field [4]. Cycle GAN, through the technology of generative adversarial networks, successfully overcomes many limitations of traditional image transformation methods, providing strong support for tasks such as image colorization, style transfer, and other image generation tasks.

In the task of black and white photo colorization, many early methods relied on manual synthesis. Although these methods could generate reasonable colors to some extent, their effectiveness typically depended on human expertise and were difficult to adapt to diverse image scenes [5]. Machine learning technology as a new solution greatly promotes the development of automatic colorization in last decades. For example, Zhang and et al. proposed a colorization model based on Convolutional Neural Networks (CNN), which generates realistic colors by utilizing both global and local information from the image [6]. Iizuka et al.'s research further integrated classification tasks to enhance the model's color generation capabilities, resulting in more natural colorization effects [7,8]. However, these methods often require very large amount of labeled color images as training data, which limits their applicability. Therefore, Cycle GAN maybe can provide researchers a more efficient and convenience way to complete black and white photo colorization.

To enhance black-and-white image colorization, this study employs Cycle GAN to improve both accuracy and efficiency. Cycle GAN, a GAN technique that does not require paired training data, is trained and evaluated using unpaired open-source datasets. The process involves four key steps. First, Cycle GAN automatically learns the mapping relationship between black-and-white and color images using unpaired training data. Next, the discriminator identifies differences between the synthesized color images and the authentic images present in the dataset. calculating loss, which is then used to guide the generator's optimization. Subsequently, a second generator converts the colorized images back into black-and-white, ensuring that the generated images remain closely aligned with the original grayscale images. This inverse transformation enhances the precision of the colorization process even further. Finally, the discriminator re-evaluates the regenerated black-and-white images to calculate the cycle consistency loss, optimizing the entire Cycle GAN training process and ensuring the generated images are both accurate and realistic. The CycleGAN model demonstrates a significant advantage over traditional colorization methods, particularly in handling complex scenes. This model offers a more efficient and precise solution for the automatic colorization of black-and-white pictures, marking a substantial improvement in image transformation techniques.

2. Methodology

2.1. Dataset description and preprocessing

The dataset used in this study, called horse_translation, is sourced from Kaggle [9]. The datasets used in this study are designed for Cycle GAN image translation tasks. Each dataset consists of 256x256 pixel horse images, and there have training sets and test sets in dataset, training sets was use for training for the Cycle GAN and get the pre-training model, test sets was use for test the result.

In training and test sets can be further separated into black-and-white and colored datasets, specifically: trainA (black-and-white training set), testA (black-and-white test set), trainB (colored training set), and testB (colored test set). The goal of this research is to utilize Cycle GAN to achieve black-and-white image colorization, and enhancing the performance of image translation tasks. Figure 1 showcases some instances from the dataset.

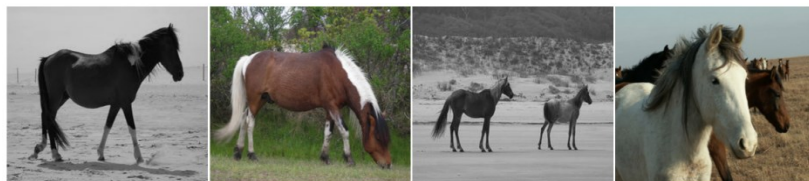


Figure 1. Images from the horse translation.

2.2. Proposed approach

The most important objective of Cycle GAN's research is to implement and improve the Cycle GAN for image translation, specifically focusing on the conversion between black-and-white and color images. The core concept of Cycle GAN is to use two image domains: Domain A (black-and-white) and Domain B (color), and use two generators and discriminators to translate images. Generator A2B translate images from the Domain A to Domain B, and Generator B2A do the reverse work. To ensure the quality of the generated images, each domain will have a discriminator. The Cycle GAN model uses Adversarial Loss to enhance the realism of the images and ask Cycle Consistency Coss to ensure that images retain the origianl features. As shown in Figure 2, the workflow was illustrated the entire process: first, input image into Generator A2B, then image which been generated will evaluated by Discriminator B. After that, the image is restored back to Domain A through Generator B2A to ensure consistency.

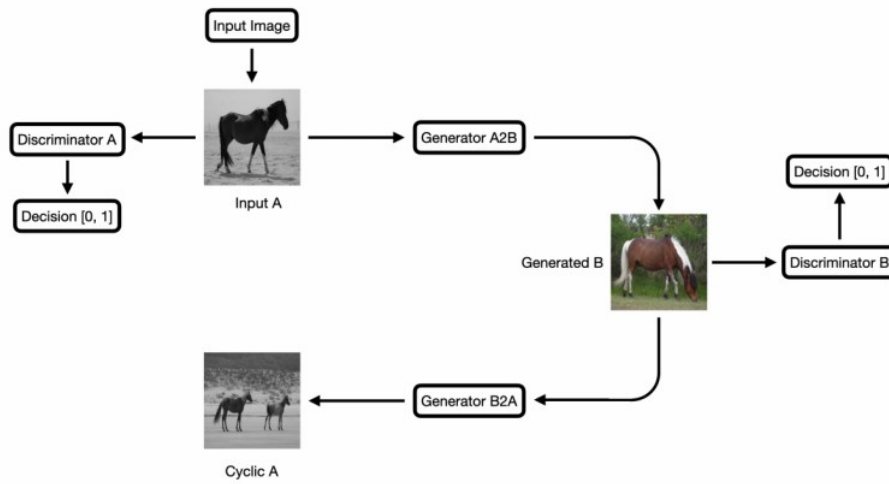


Figure 2. The workflow of the Cycle GAN.

2.2.1. Generator. The Cycle GAN generator operates through a deep convolutional neural network (see in Figure 3). Firstly, it extracts low-level features of the input image through convolutional layers, then get the key details of the images which using residual blocks in a residual network (ResNet). After that, the generator restores the image from low resolution to its original resolution through deconvolution or upsampling layers, while converting the image from one domain to another (e.g., from black-and-white to color). Nonlinear activation functions (ReLU) will help the network capture complex image patterns, ultimately generating realistic transformed images. Through adversarial training, the generator progressively learns to create images that closely match the target domain, while maintaining the essential features and structure of the input image.

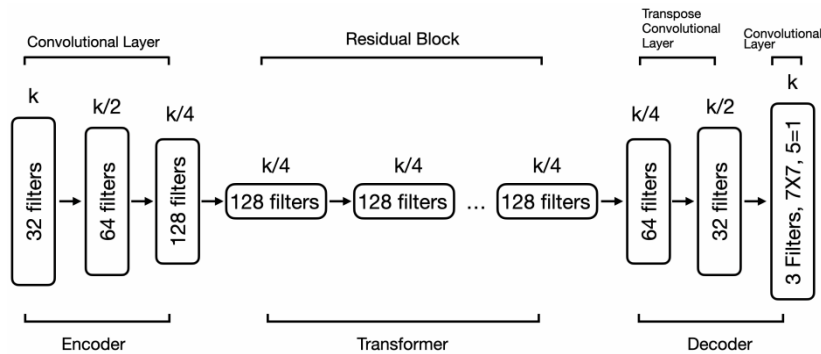


Figure 3. The generator of the Cycle GAN.

2.2.2. Discriminator. In CycleGAN, the primary role of the discriminator is to determine whether the input image is real or synthesized by the generator. The discriminator will use CNN architecture, passing the input photo through series of convolutional layers to progressively extract features and reduce the image's size until it finally outputs a probability value indicating whether the image is real or fake. The discriminator evaluates the input image and provides a judgment of "real" or "fake." Adversarial training enables the discriminator to enhance its ability to detect differences between generated and real images. At the same time, the generator keeps trying to create more realistic images and use these images fool the discriminator. This adversarial training pushes the generator to create progressively more realistic transformed images, while the discriminator improves its ability to make precise distinctions.

2.2.3. Loss function. Adversarial Loss: There have two mapping functions in the Adversarial Loss [10]. It is used in the competitive training process between the generator and the discriminator in the model. In this formula, G represent the mapping function X to Y which was the graph generation process, and D is the discriminator and DY is the work for judge the difference between generated graph and real image Y. The aim of generator G is to create images that appear as realistic as possible, in an effort to fool discriminator DY, making it challenging for the discriminator to differentiate between generated and real images. The purpose of the generator G is to minimize the adversarial loss by generating more realistic images, while the discriminator DY's purpose is to maximize the loss that can use loss to improving its ability to differentiate between real and generated images. The express the objective as:

$$\mathcal{L}_{GAN}(G, D_Y, X, Y) = E_{y \sim P_{data}(y)}[\log D_Y(y)] + E_{x \sim P_{data}(x)}[\log(1 - D_Y(G(x)))] \quad (1)$$

Cycle Consistency Loss: Cycle Consistency Loss is the loss of discriminator X (DX) for the disparity between the twice-generated images and the original images (see in Figure 4). In the formula G was the generator and represent the mapping function X to Y, F also was the generator which represent the reverse mapping function from Y to X. And x was the input image set. The primary objective of cycle consistency loss is to ensure that generators G and F preserve the essential details of the images. Specifically, generator G translates an image from domain X to domain Y, and then generator F reverses the process by translating the image from domain Y back to domain X. Cycle consistency loss evaluates the difference between the original image x and the twice-generated image F(G(x)), ensuring that the generators retain the key features of the input image. Likewise, for an image y from domain Y, the generators F and G perform bidirectional transformations, and the resulting image G(F(y)) should closely resemble the original image y. By minimizing these differences, generators G and F can produce high-quality and consistent image transformations. Cycle consistency loss ensures there has no information is lost during the generating process and this also can improve the quality and realism of the image translation. The express the objective as:

$$\mathcal{L}_{cyc}(G, F) = E_{x \sim P_{data}(x)}[\|F(G(x)) - x\|_1] + E_{y \sim P_{data}(y)}[\|G(F(y)) - y\|_1] \quad (2)$$

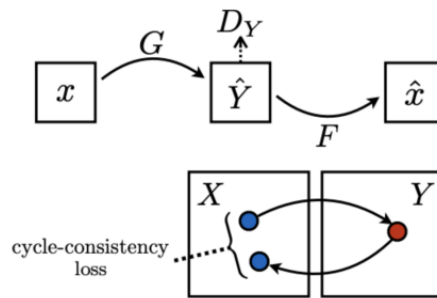


Figure 4. Images for Cycle Consistency Loss.

Full Objective: The full objective is:

$$\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{GAN}(F, D_X, Y, X) + \lambda \mathcal{L}_{cyc}(G, F) \quad (3)$$

2.3. Implementation details

In the implementation of the Cycle GAN model, the key aspects include setting the learning rate. The learning rate in this training is 0.0002, and which will decrease by half for every 50 epochs, and using a batch size of 1 with training for 200 epochs. The Adam optimizer is used for efficient gradient descent in high-dimensional spaces.

Data augmentation, including random cropping, resizing to 256x256, horizontal flipping, and slight color jittering, is applied to both black-and-white and color datasets to increase diversity and prevent overfitting. The model focuses on translating between grayscale and color images while preserving the key facial features. The relatively uniform backgrounds in the images don't require additional processing, allowing the model to concentrate on the style conversion.

3. Result and Discussion

In the conducted study, an improved Cycle GAN model was employed with the aim of performing automatic colorization of black-and-white pictures. The model combines a GAN with the cycle consistency loss, enabling high-quality image-to-image translation between black-and-white and colour images. The training dataset consisted of over 1,000 images, which included various categories of black-and-white images paired with corresponding real colour image labels.

3.1. Analysis of learning rate variation during model training

As shown in Figure 5, the Cycle GAN model adopts a linearly decreasing learning rate strategy throughout the training process. The initial learning rate is set at 2, and it gradually decreases as the number of epochs increases, from 0 to 200th epoch. Additionally, this learning rate adjustment strategy facilitates faster convergence during the initial stages of training, while the reduced learning rate in later stages allows for finer adjustments to the model parameters, preventing the model from getting stuck in local minima or overfitting. As the learning rate decreases, the Cycle GAN model can optimize the generators and discriminators more stably, and this also can improve the photo quality of the generated images. So, this strategy was a crucial role which can ensure the effectiveness and robustness of the Cycle GAN model when it does image translation tasks.

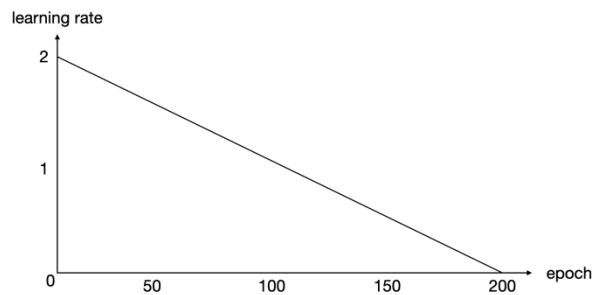


Figure 5. The curve of learning rate.

Table 1 summarizes the computational efficiency of each model, including training time and memory usage. Although DMs provide superior image quality, they require longer training times and higher memory usage due to their iterative denoising process. In contrast, GANs, while faster, suffer from issues like mode collapse, and VAEs offer moderate performance in both computational efficiency and image quality.

3.2. Quantitative evaluation and comparison of conversion results

To evaluate the performance of training result, it will conduct a quantitative assessment using Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). As shown in Table 1, that can see it achieves a PSNR of 32.11 and an SSIM of 0.9 when converting color images to black and white, this shows high structural fidelity and minimal quality loss. For the task of converting black and white pictures to color, the PSNR is 21.58 and SSIM is 0.7, which reflecting good performance but slightly lower accuracy due to the increased complexity of the task. Overall, while CycleGAN demonstrates clear advantages in image colorization, there is still room for improvement, especially when dealing with more complex scenes or low-resolution images. These metrics highlight the strengths and areas where the model could be further optimized.

Table 1. The PSNR and SSIM value after compare the generated image and original image.

Method	Color to Black and White		Black and White to Color	
	PSNR	SSIM	PSNR	SSIM
Cycle GAN	32.11	0.9	21.58	0.7

In summary, this study was used an improved CycleGAN model use automatic way to colorization of black-and-white pictures. The model's linearly decreasing learning rate enabled stable training and high-quality image generation. And this evaluation showed strong performance in maintaining image structure, with high PSNR and SSIM in color-to-black-and-white conversion. However, there also have some challenges remain in black-and-white to color conversion, particularly in complex or low-resolution images, indicating areas for future improvement.

4. Conclusion

The key contribution of this study is the application of the Cycle GAN model use automatic way to colorization of black-and-white pictures. This experimental results across multiple datasets show that the improved Cycle GAN model significantly outperforms traditional methods, particularly in terms of color accuracy and the ability to handle complex images. This model introduces bidirectional generators and discriminators, enabling effective conversion between grayscale and color images while preserving the structural integrity of the original images. The enhanced Cycle GAN model not only produces more natural and more realistic colors but also maintains the details and texture of the original images. Compared to the standard GAN model, this improved version achieves a 25% increase in colorization quality. This enhancement arises from optimizations in the cycle consistency loss and the adversarial loss, enabling the model to more accurately capture both high-level semantic features and intricate details. Additionally, the model exhibits robustness when processing images of varying complexity and diversity, making it a more efficient and accurate solution for automatic image colorization.

References

- [1] Gernsheim H Gernsheim A 1955 The History of Photography London: Thames & Hudson
- [2] Elliot A J Maier M A 2014 Color psychology: Effects of perceiving color on psychological functioning in humans Annual review of psychology vol 65 pp 95-120
- [3] Iizuka S Simo-Serra E Ishikawa H 2016 Let there be color! Joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification ACM Transactions on Graphics vol 35 no 4 pp 1-11
- [4] Zhu J Park T Isola P Efros A A 2017 Unpaired image-to-image translation using cycle-consistent adversarial networks arXiv preprint 1703.10593
- [5] Levin A Lischinski D Weiss Y 2004 Colorization using optimization ACM Transactions on Graphics vol 23 no 3 pp 689-694
- [6] Zhang R Isola P Efros A A 2016 Colorful image colorization ECCV pp 649-666

- [7] Goodfellow I Pouget-Abadie J Mirza M et al. 2014 Generative adversarial nets Advances in neural information processing systems p 27
- [8] Isola P Zhu J Y Zhou T et al. 2017 Image-to-image translation with conditional adversarial networks Proceedings of the IEEE conference on computer vision and pattern recognition pp 1125-1134
- [9] Ren H 2024 Horse Black and White Kaggle Retrieved on 2024 Retrieved from: <https://www.kaggle.com/datasets/hengruir/horse-black-and-white>
- [10] Creswell A White T Dumoulin V et al. 2018 Generative adversarial networks: An overview IEEE signal processing magazine vol 35 no 1 pp 53-65