# Synthetic rain image generation via CycleGAN

**Kaiqi Yang[1,3,5], Yanfeng Lu[2,4]**

[1]USC Viterbi School of Engineering, University of Southern California
[2]Institute of Automation, Chinese Academy of Science, Beijing, China

[3]bryantykq7@gmail.com
[4]Yanfeng.lv@ia.ac.cn
[5]Corresponding author: Kaiqi Yang

**Abstract.** In the domain of autonomous driving and computer vision, the formidable challenge of adverse weather conditions, particularly rainy weather, profoundly impacts image quality and visibility. Rain streaks pose a significant impediment to accurate object detection, especially concerning pedestrians and vehicles on the road. While existing solutions have focused on training rain removal models on paired images, acquiring such data with congruent backgrounds has been a substantial hurdle. Three distinct images, each featuring a diverse background, were deliberately chosen for the rain removal experiment. The deliberate selection of these varied backgrounds allowed us to rigorously assess the efficacy of our rain removal method across diverse environmental contexts. The compelling results obtained from this experiment affirm the method's ability to effectively mitigate rain line across a spectrum of backgrounds, thus establishing its robustness and versatility in real-world scenarios. The resulting images are then employed to train our rain removal model, marking a significant advancement in our endeavor. Our innovative approach stands poised to revolutionize rain removal techniques, aligning synthetic and authentic rain images more closely in real-world scenarios.

**Keywords:** synthetic rain image, cycleGAN, image restoration, single image rain removal

## 1. Introduction

Visual perception is a cornerstone of self-driving systems, enabling vehicles to interpret and navigate the dynamic environment around them. Through the efforts of many researchers, the technology of self-driving developed rapidly. These autonomous vehicles have demonstrated remarkable capabilities in recognizing road signs, detecting pedestrians, and avoiding obstacles. [1] proposes a region proposal network to generate potential object regions, enabling accurate and real-time object detection by integrating region-based convolutional neural networks, crucial for enhancing object detection performance in self-driving scenarios. [2] introduces a novel approach to object detection by detecting objects as pairs of key points, achieving high accuracy, and reducing false positives, providing potential for improved object detection robustness in self-driving scenarios. [3] introduces YOLO, a real-time object detection algorithm, which significantly improves speed and accuracy by directly predicting bounding boxes and class probabilities in a single pass through the network, with implications for real-time perception in self-driving cars. [4] introduces a single-shot object detection approach with multiple scale predictions, optimizing speed and accuracy trade-offs, which has implications for rapid and reliable object detection in the complex environments of self-driving cars. [5] introduces an efficient method for object detection using point clouds, optimizing the encoding process and demonstrating advancements in object recognition from 3D data.

However, despite these advancements, the performance of self-driving systems can be significantly compromised under adverse weather conditions, particularly in the presence of rain. This drives the development of the single image rain removal. With the development of deep learning, the development of single rain removal has seen significant progress over the years, with various researchers and teams contributing to the advancement of rain removal techniques. [6] introduced a deep learning-based approach using a detail network to separate rain streaks from background details for rain removal. [7] proposed a layer-prior-based method to decompose an input image into rain streaks and background layers for more effective rain removal. [8] introduced a multi-stream dense network for rain removal, incorporating density-aware modules to handle different rain densities. [9] provided a comprehensive benchmark analysis of existing rain removal methods and evaluated their performance on a standardized dataset, aiding in the comparison and development of deraining algorithms. [10] developed contextualized deep networks for simultaneous rain detection and removal, leveraging the contextual information for improved performance.

The vast majority rain removal models typically necessitate paired images, one with rain and the other without, sharing the same background proves to be an arduous undertaking.

**Figure 1.** Comparison of real rain image and synthetic rain image. (a) first is a real rain image, (b) second is a synthetic rain image

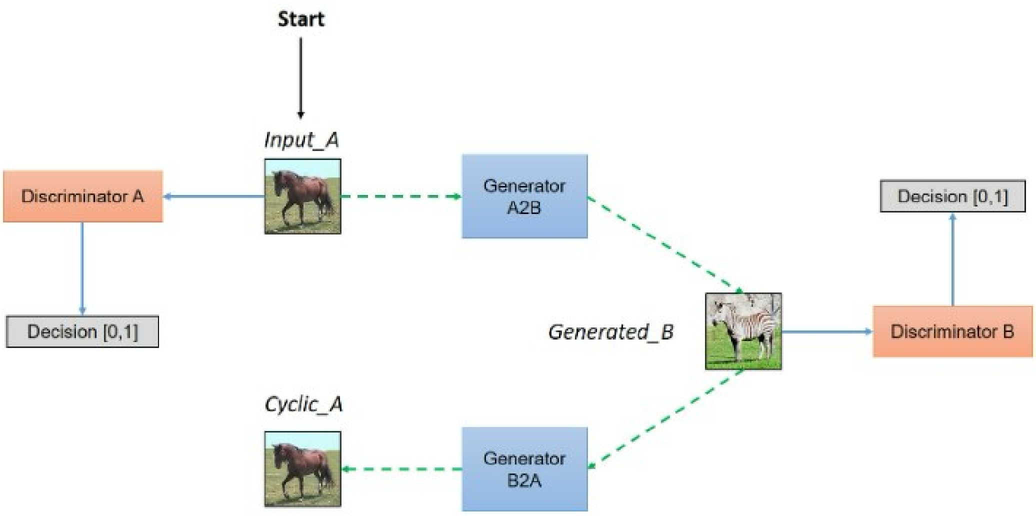


**Figure 2.** Network architecture of CycleGAN

DOI: 10.54254/2755-2721/54/20241667

Consequently, the utilization of synthetic rain images has become a prevalent choice. Although several Derain methods have showcased promising outcomes in removing rain from synthetic rainy images datasets such as Rain100T, their efficacy considerably wanes when confronted with real rain images. This discrepancy can be attributed to the fact that synthetic rain images are often crafted using tools like Photoshop or Python script, thereby failing to encapsulate the intricate attributes of real rain images. The challenges of rain removal from authentic rain images are complicated by variations in raindrop size, shapes, densities, motion patterns, as well as the intricate interplay of background and lighting condition.

Our proposed methodology presents an innovative synthesis approach for generating synthetic rain images, designed to overcome inherent limitations associated with existing synthetic rain datasets. By adopting this novel synthesis strategy, our research bridges the gap between synthetic and real rain imas, enhancing the efficacy of rain removal techniques in real-world scenarios, and increasing the object detection ability in rainy weather. Additionally, we contribute a novel dataset comprising synthetic rain images paired with corresponding ground truth images, fostering improved training and evaluation of rain removal algorithms within a controlled environment.

**Figure 3.** Image without rain (left) and image with rain (right)

## 2. Related work

### *2.1. CycleGAN*

The CycleGAN (Cycle-Consistent Generative Adversarial Network) framework has revolutionized unsupervised image-to-image translation by introducing cyclic consistency loss, enabling the conversion between two domains without the need for paired training data. This breakthrough has found particularly powerful application in tasks like adding synthetic rain to images, where conventional methods often rely on manual interventions like Photoshop or Python scripts. What sets our approach apart is its ability to leverage pairs of rain-laden and rain-free images, without the constraint of identical backgrounds. This flexibility allows for the seamless integration of realistic synthetic rain, offering a more convincing and visually appealing result compared to traditional techniques.

### *2.2. Restoration Transformer*

The Restoration Transformer represents a groundbreaking advancement in the domain of image restoration. By integrating the strengths of transformer architectures with convolutional neural networks (CNNs), this novel model excels at tasks like denoising, deblurring, and super-resolution. Its ability to capture global dependencies alongside fine-grained local details leads to remarkable gains in restoration quality. In terms of removing rain, Leveraging the attention mechanism and a multi-head self-attention mechanism, it excels in capturing both local and global contextual information crucial for effective derain.

DOI: 10.54254/2755-2721/54/20241667

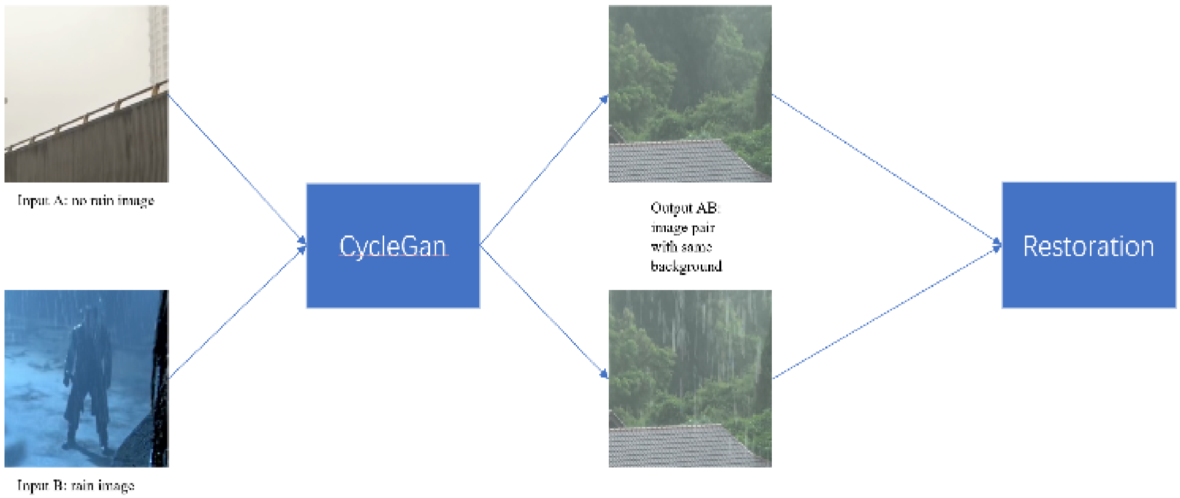


**Figure 4.** Flow chart of operation

## 3. Opeartion method

### *3.1. Datasets collection*

Addressing the challenge of obtaining paired datasets featuring rain and rain-free images sharing the same background, our research embraces the innovative paradigm of unpaired image-to-image translation facilitated by Cycle-Consistent Adversarial Networks (CycleGAN). This approach offers a pivotal solution to the inherent complexities of collecting matched datasets. The conventional prerequisite for identical backgrounds is obviated, streamlining dataset acquisition considerably. In this methodology, Figure 3 and Figure 4 represent exemplars of our dataset—both of which encapsulate the potential for unpaired training. While their backgrounds diverge, these images serve as coherent instances for the synthesis of rainy and rain-free variations. Our framework leverages the capabilities of CycleGAN to bridge the gap between these distinct images through an adversarial learning process that enables the generation of synthetic rain-infused images.

Upon the successful addition of rain to the images, the resultant dataset comprises paired samples—synthetic rain images and corresponding ground truth images. This dataset assumes a pivotal role in the subsequent phase of our research, facilitating the training of our proprietary rain removal model. The provision of coherent pairs, albeit originating from unpaired sources, enables the development of a model capable of discerning the intricacies of rain-laden scenes and, in turn, effectively mitigating the undesired effects of rain in various computer vision applications.

### *3.2. Train with Restoration Transformer nerual network*

We chose to train our single rain removal model with Restoration Transformer. Restoration Transformer makes an improvement based on Transformer architecture neural network. Transformer neural network has been introduced into image restoration to replace convolution neural network. But it usually has a high computational complexity because of global attention. But Restoration Transformer provides e propose an efficient Transformer for image restoration that is capable of modeling global connectivity and is still applicable to large images. introduces the MDTA (multi-Dconv head 'transposed' attention) block as a replacement for traditional self-attention mechanisms. Unlike standard approaches, MDTA focuses on feature relationships rather than pixel interactions.

Besides Restoration Transformer neural network. There are also many other rain removal networks feasible for this task, e.g. CNN (Convolution neural network). With the more real synthetic rain images datasets, they should have a better effectiveness on single image rain removal.

DOI: 10.54254/2755-2721/54/20241667

## 4. Experiment results

To show the advantage of using our more real synthetic rain images to train network, we will use same rain removal network to train with different datasets and compare the results of them on rain removal task of real images.

### *4.1. Implementation details*

First, we prepare a real rain image dataset [11] and a rain free image dataset. Subsequently, CycleGAN architecture is employed to train our add rain model. The trained add rain model is then applied to incorporate synthetic rain into the rain free model. We utilize the Restoration Transformer to train our rain removal model. This process is consistently implemented across various datasets (Rain13K, cityscape, our datasets) to ensure uniformity. Lastly, the trained model is evaluated using real rain images, offering an assessment of its performance under authentic conditions.

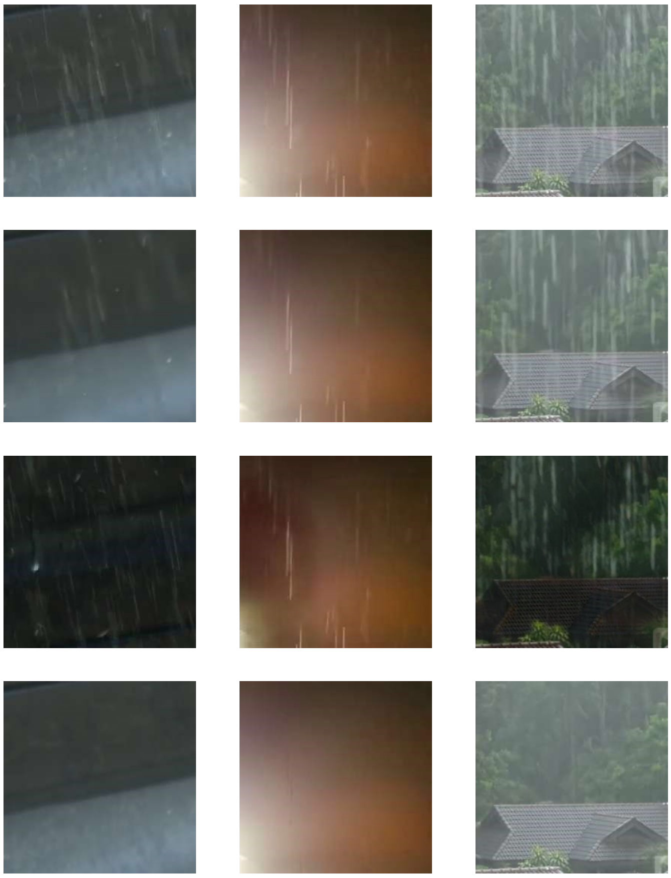

**Figure 5.** Comparison of rain removal results of different datasets.

**Table 1**. PSNR and SSIM of train with different datasets

| | Rain13K | cityscape | Our datasets | Rain 100L | Rain 100H |
|---|---|---|---|---|---|
| PSNR | 35.81 | 28.09 | 39.76 | 36.32 | 37.79 |
| SSIM | 0.9767 | 0.8147 | 0.9785 | 0.9644 | 0.9734 |

### *4.2. Evaluation*

To thoroughly assess the effectiveness of our rain removal methodology, we underwent a detailed comparative analysis involving three different datasets. In Figure 5, each vertical column presents the results of removing rain from the same image after training with different datasets. As shown in Figure

5, our datasets can remove most of the streaks of the rain image. The first row are the original images. The second row are images derain after train with Rain13K. The third row are images derain after train with cityscapes. The fourth row are images derain after train with our datasets. Three distinct images, each featuring a diverse background, were deliberately chosen for the rain removal experiment. Evident from the image, our method demonstrates exceptional efficacy in retaining minimal rain streaks and near-perfectly restoring the scene to its rain-free state. The deliberate selection of these varied backgrounds allowed us to rigorously assess the efficacy of our rain removal method across diverse environmental contexts. The compelling results obtained from this experiment affirm the method's ability to effectively mitigate rain streaks across a spectrum of backgrounds, thus establishing its robustness and versatility in real-world scenarios.

In addition, we employ Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) metrics for a more in-depth analysis. The computer average PSNR and SSIM values are presented in Table 1. Table 1 clearly indicates that our datasets exhibit higher PSNR and SSIM. Its PSNE and SSIM reached 39.76 and 0.9785 respectively. These empirical findings further validate the effectiveness of our method.

## 5. Conclusions

In this study, we address the significant challenge of adverse weather conditions, particularly rain, which severely impacts image quality and hinders accurate object detection in the domains of autonomous driving and computer vision. While previous solutions focused on training rain removal models with paired images, acquiring such data with matching backgrounds posed a substantial hurdle. To overcome this limitation, we propose a pioneering approach utilizing CycleGAN to generate synthetic rain images. Unlike conventional methods, our technique does not rely on matched backgrounds, allowing for a higher degree of authenticity in simulating rain-induced conditions. The resulting images are then used to train our rain removal model, marking a substantial advancement. Our approach holds the potential to revolutionize rain removal techniques, narrowing the gap between synthetic and authentic rain images in real-world scenarios, thereby enhancing the capabilities of autonomous systems in adverse weather conditions. The experimental results affirm the efficacy of our approach, showcasing its ability to effectively remove rain streaks and consistently outperforming other datasets in terms of image quality, as confirmed by Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) metrics. These findings collectively highlight the promise of our methodology in bolstering the performance of autonomous systems under challenging weather conditions.

## References

[1] Ren, Shaoqing, et al. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, 1 June 2017, pp. 1137–1149, ieeexplore.ieee.org/document/7485869, https://doi.org/10.1109/tpami.2016.2577031.

[2] Law, Hei, and Jia Deng. “CornerNet: Detecting Objects as Paired Keypoints.” International Journal of Computer Vision, vol. 128, no. 3, 8 Aug. 2019, pp. 642–656, https://doi.org/10.1007/s11263-019-01204-1.

[3] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 779–788). https://doi.org/10.1109/CVPR.2016.91

[4] Liu, Wei, et al. "Ssd: Single shot multibox detector." Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016.

[5] Lang, Alex H., et al. "Pointpillars: Fast encoders for object detection from point clouds." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019.

DOI: 10.54254/2755-2721/54/20241667

[6] X. Fu, J.-B. Huang, D. Zeng, Y. Huang, X. Ding, and J. Paisley, “Removing Rain from Single Images via a Deep Detail Network,” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jul. 2017, doi: 10.1109/cvpr.2017.186.
[7] W. Yang, R. T. Tan, J. Feng, J. Liu, Z. Guo, and S. Yan, “Deep Joint Rain Detection and Removal from a Single Image,” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jul. 2017, doi: 10.1109/cvpr.2017.183.
[8] H. Zhang and V. M. Patel, “Density-Aware Single Image De-raining Using a Multi-stream Dense Network,” 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Jun. 2018, doi: 10.1109/cvpr.2018.00079.
[9] S. Li et al., “Single Image Deraining: A Comprehensive Benchmark Analysis,” 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2019, doi: 10.1109/cvpr.2019.00396.
[10] W. Yang, R. T. Tan, J. Feng, Z. Guo, S. Yan, and J. Liu, “Joint Rain Detection and Removal from a Single Image with Contextualized Deep Networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 6, pp. 1377–1393, Jun. 2020, doi: 10.1109/tpami.2019.2895793.
[11] Law, Hei, and Jia Deng. “CornerNet: Detecting Objects as Paired Keypoints.” International Journal of Computer Vision, vol. 128, no. 3, 8 Aug. 2019, pp. 642–656, https://doi.org/10.1007/s11263-019-01204-1.