

Development of a Modified CycleGAN Model with Residual Blocks and Perceptual Loss for Image Dehazing

Sani Moch Sopian^{1*}, Arief Suryadi Satyawan², Mokhammad Mirza Etnisa Haqiqi³, Helfy Susilawati⁴, Beni Wijaya⁵, Khaulyca Arva Artemysia⁶, Firman⁷, Muhammad Ikbal Samie⁸

^{1,3,4,5,6,7,8}Department of Electrical Engineering, Faculty of Engineering, Garut University, West Java, Indonesia

²Research Center for Telecommunication – BRIN, Indonesia

E-mail: ^{1*}sanimochsopian1@gmail.com, ²arie021@brin.go.id, ³mirza@uniga.ac.id, ⁴helfy.susilawati@uniga.ac.id, ⁵beniw163@gmail.com, ⁶khaulyca@gmail.com, ⁷firmanfardiansyah2592@gmail.com, ⁸muhamadikbalsh1@gmail.com

(Received: 21 Apr 2025, revised: 7 May 2025, accepted: 9 May 2025)

Abstract

Fog reduces image contrast and clarity, creating challenges for applications such as autonomous driving and remote sensing. This study proposes a series of CycleGAN modifications for single image dehazing using unpaired data, integrating residual blocks, attention mechanisms, VGG19-based perceptual loss, and haze-aware loss. Among ten architectural variants, Modification 10 combining perceptual and haze-aware loss achieved the best overall performance. Quantitatively, it showed stable generator losses (0.91 for Gen G, 0.57 for Gen F), with improved discriminator performance (Disc X: 0.59, Disc Y: 0.47), indicating better training stability and image realism. Additionally, it offered competitive PSNR (7.99), strong SSIM (0.4202), and low LPIPS (0.6577), confirming its effectiveness in both pixel-level accuracy and perceptual quality. Qualitatively, this model generated clearer, more natural images with improved edge sharpness and detail preservation. These findings demonstrate that the modified CycleGAN significantly enhances dehazing performance and presents a valuable contribution to deep learning-based image restoration.

Keywords: Dehazing, CycleGAN, Deep Learning, Residual Block, Attention Mechanism.

I. INTRODUCTION

In recent years, the demand for advanced technologies to enhance image quality has been increasing, particularly in applications requiring improved visibility. One of the primary challenges in fields such as autonomous driving, remote sensing, and outdoor surveillance is the degradation of image quality caused by fog or haze. Atmospheric particles scatter light, reducing image contrast and clarity, making it crucial to develop efficient algorithms for haze removal (dehazing) [1]. Such algorithms improve the quality of visual data, enabling better scene understanding, more accurate object detection, and enhanced overall system performance. The ability to remove haze from images is essential not only for ensuring safety in critical systems but also for enhancing the visual experience in everyday applications such as photography and video streaming [2].

Over the past decade, extensive research has explored various approaches to dehazing. Early techniques were primarily based on image processing and physical models. For instance, [3] introduced the Dark Channel Prior (DCP)

method, widely adopted for single-image dehazing due to its effectiveness in estimating transmission maps and restoring clear images. Another significant method is Non-Local Image Dehazing, developed by [4], which proposed a haze-line-based approach by leveraging distinct color lines in haze-free images.

More recently, with the rise of deep learning, researchers have shifted to convolutional neural networks (CNNs) to improve dehazing performance. The study by [5] proposed DehazeNet, a CNN-based architecture that directly predicts the transmission map, showing better results compared to traditional techniques. Additionally, AOD-Net by [6] and GridDehazeNet by [7] have also contributed significantly to advancements in dehazing technology using deep learning.

Among various generative models, CycleGAN has emerged as a powerful technique for handling image-to-image translation tasks, including dehazing. Initially proposed by [8], CycleGAN was designed to translate images between two domains without paired datasets, making it ideal for scenarios like dehazing, where paired images of hazy and clear

conditions are often unavailable. Originally applied to tasks such as transforming horse images into zebra images, its potential was later extended to other domains, including haze removal. By learning to map unpaired hazy images (trainA) to clear images (trainB) through a cycle consistency process, CycleGAN can effectively restore visibility in images.

However, despite its advantages, standard CycleGAN architecture often struggles to preserve fine details and texture in the dehazed output, leading to suboptimal visual quality. Additionally, previous approaches largely overlook the integration of attention mechanisms and perceptual feedback to guide the model toward learning more visually coherent results. These limitations highlight a significant research gap in designing more robust CycleGAN-based models specifically tailored for haze removal [9].

Therefore, this study aims to enhance the performance of CycleGAN in single-image dehazing by introducing architectural modifications, including the addition of residual blocks, attention mechanisms, and perceptual loss functions. The objective is to evaluate and identify the most effective architectural combination for improving the clarity and realism of dehazed images, especially in scenarios involving unpaired training data. Through these contributions, this work seeks to address the current limitations in deep learning-based dehazing and provide a more effective solution for real-world applications.

II. METHODOLOGY

This study proposes a series of modifications to the original CycleGAN architecture to improve the quality of dehazed images. The modifications aim to enhance the network's ability to remove haze artifacts, preserve structural details, and optimize training dynamics. A total of ten variants (Mod1 to Mod10) were developed, each introducing different architectural or loss-based enhancements.

A. Dataset and Preprocessing

The training process utilizes unpaired images from two domains: domain A consists of hazy images, while domain B consists of clear, haze-free images. All images were resized to 256×256 pixels and normalized to the range [-1, 1]. To improve generalization and robustness, several augmentation techniques were applied, including random cropping, horizontal flipping, and brightness adjustments [10].

B. Model Variants

The following ten variants represent modifications made to the generator architecture of the CycleGAN model:

1. Mod 1 – Residual Dilated Block with LayerNorm: Two residual blocks use dilated convolutions (dilation=2) and layer normalization to expand the receptive field and improve training stability.
2. Mod 2 – Deeper Residual Stack: Builds on Mod 1 by using four residual blocks to extract deeper features.

3. Mod 3 – Self-Attention Integration: Adds a self-attention layer after residual blocks, allowing the network to focus on haze-concentrated regions.
4. Mod 4 – 6 Residual Blocks: Extends the network depth with six residual blocks to enhance abstraction.
5. Mod 5 – 8 Residual Blocks: Further deepens the model using eight residual blocks for more complex feature learning.
6. Mod 6 – Extra Downsampling Layer: Adds a downsampling layer with 1024 filters before the residual blocks to handle more spatial variations.
7. Mod 7 – Mod 6 + 8 Residual Blocks: Combines Mod 6 with the deeper eight-block structure from Mod 5.
8. Mod 8 – Perceptual Loss Integration: Introduces VGG19-based perceptual loss to emphasize texture and semantic consistency.
9. Mod 9 – Lightweight Model with Perceptual Loss: Reduces the model's complexity while retaining perceptual loss, creating a more efficient variant.
10. Mod 10 – Haze-Aware Loss Fusion: Combines perceptual loss with a custom haze-aware loss based on dark channel prior statistics to penalize remaining haze artifacts.

Each variant was trained independently to examine the effectiveness of the specific modifications applied.

C. Loss Functions

The overall training objective consists of several loss functions. The adversarial loss encourages the generator to produce outputs that are indistinguishable from real haze-free images. The cycle consistency loss ensures that the transformations from hazy to clear and back to hazy produce consistent results. Identity loss is applied to preserve the color and lighting of images that are already in the target domain [11]. In Mod 8 to Mod 10, perceptual loss is added, which measures the L2 distance between high-level feature maps extracted from a pretrained VGG19 network, promoting semantic similarity. In Mod 10, a haze-aware loss is introduced, designed to reduce haze residues by penalizing discrepancies in dark channel statistics between the generated and target images [12].

D. Hyperparameter Settings

Each of the ten modified CycleGAN models (Mod1 to Mod10) was trained using consistent training hyperparameters to ensure fair comparison and isolate the effect of architectural changes. The following hyperparameters were applied throughout the training process.

- Optimizer: Adam optimizer was used for both the generator and discriminator networks.
- Learning Rate: Set to 0.0002, a commonly effective value for GAN training that provides stable convergence.
- Beta Parameters (β_1 , β_2): The momentum terms for the Adam optimizer were set to $\beta_1 = 0.5$ and $\beta_2 = 0.999$, following the settings suggested in the original CycleGAN implementation.

- **Batch Size:** A batch size of 1 was used to preserve spatial consistency and maximize the diversity of training samples per step.
- **Number of Epochs:** Each model was trained for 200 epochs to ensure adequate convergence and full exposure to training data.
- **Loss Weights:**
 - ✓ **Cycle Consistency Loss:** Weight = 10, to enforce meaningful round-trip translation between hazy and clear image domains.
 - ✓ **Identity Loss:** Weight = 5, used to maintain the color and brightness consistency, especially when the input is already from the target domain.
 - ✓ **Perceptual Loss (Mod8–Mod10 only):** Weight = 1, based on VGG19 features to ensure high-level texture similarity between output and clear images.
 - ✓ **Haze-Aware Loss (Mod10 only):** Weight = 3, used to penalize haze remnants based on dark channel prior statistics.

All images were resized to 256×256 pixels and normalized to the range $[-1, 1]$. Training data augmentation techniques included random cropping, horizontal flipping, and brightness adjustment, aiming to improve the model's generalization on diverse haze patterns.

E. Diagram Block

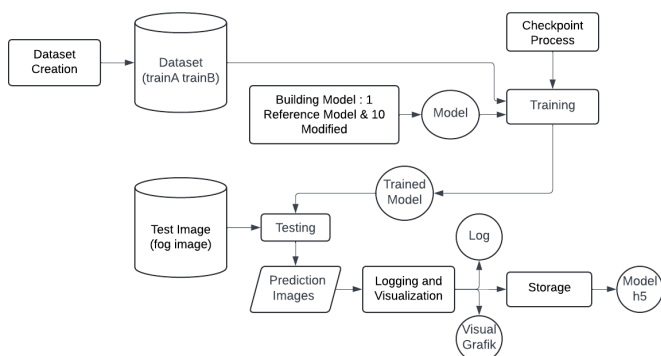


Figure 1. Diagram Block

Figure 1 illustrates the overall workflow of the proposed system, beginning with data preparation and culminating in model evaluation. Initially, the dataset is divided into two main subsets: trainA, which contains hazy images, and trainB, which comprises clear, haze-free images. These datasets are unpaired, meaning that each hazy image in trainA does not have a corresponding clear version in trainB. This separation is crucial because the model utilized in this study, CycleGAN, is specifically designed to perform image-to-image translation without the need for paired images. This approach aligns with the original purpose of CycleGAN, which was developed to translate between two domains such as horses and zebras without requiring exact image pairs.

Following dataset preparation, the model-building phase involves constructing one reference CycleGAN model and ten

modified versions. These modified models incorporate various architectural enhancements, including residual blocks, dilated convolutions, self-attention mechanisms, perceptual loss, and haze-aware loss. The models are trained using the trainA and trainB datasets, resulting in a set of trained models. During training, logs and performance visualizations are generated, and model checkpoints are saved to preserve the best-performing weights. The final model weights are stored in .h5 format.

After training, the models are evaluated using a separate test dataset composed of hazy images that are distinct from those in trainA. This evaluation assesses the models' ability to generalize to unseen data. The predicted images are then saved and analyzed both visually and quantitatively. Additionally, the logging and visualization processes conducted during both training and testing facilitate performance monitoring and provide insights into the effectiveness of each model. The final outputs including predicted images, performance graphs, and model weights are stored for documentation and further analysis.

F. Evaluation Metrics

To objectively assess the performance of the proposed models, several evaluation metrics are used. Peak Signal-to-Noise Ratio (PSNR) quantifies the reconstruction fidelity at the pixel level. Structural Similarity Index (SSIM) measures structural alignment between generated and reference images. Additionally, Learned Perceptual Image Patch Similarity (LPIPS) is employed to evaluate perceptual quality using deep features [13].

III. RESULT

The evaluation of various CycleGAN model modifications, detailed across tables, reveals key quantitative and qualitative insights. Each modification introduced adjustments to either the generator or discriminator, enhancing feature extraction, model stability, or generalization in the dehazing context.

Table 1 presents the generator and discriminator losses at Epoch 1 and Epoch 20. All models demonstrate stability in generator losses (Gen G Loss and Gen F Loss), which remain constant at 0.91 and 0.57, respectively, across all models and epochs. This consistency suggests that the generator architectures in all models, including the baseline CycleGAN and all modifications, maintained a similar level of optimization convergence over time.

Discriminator losses (Disc X and Disc Y), however, display mild variation. At Epoch 20, Disc X Loss ranged from 0.56 to 0.61, and Disc Y Loss varied from 0.84 to 0.90. Notably, Modification 9 showed the lowest Disc Y Loss (0.84), suggesting potentially better discriminator performance in distinguishing generated samples. Conversely, Modification 10 showed relatively higher Disc X Loss (0.61), potentially indicating more difficulty in distinguishing real from generated data at that stage.

Table 1. Gen G Loss, Gen F Loss, Disc X Loss, dan Disc Y Loss

Number of Epoch	Cycle GAN	Mod 1	Mod 2	Mod 3	Mod 4	Mod 5	Mod 6	Mod 7	Mod 8	Mod 9	Mod 10
Gen G Loss											
1	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
20	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
Gen F Loss											
1	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.57
20	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.57	0.57
Disc X Loss											
1	0.60	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.59	0.60
20	0.57	0.56	0.56	0.57	0.57	0.57	0.59	0.58	0.56	0.61	0.58
Disc Y Loss											
1	0.85	0.86	0.85	0.85	0.85	0.85	0.86	0.85	0.86	0.83	0.86
20	0.90	0.89	0.89	0.90	0.88	0.89	0.87	0.88	0.89	0.84	0.88

Table 2. Metric-Based Performance Analysis

Model	PSNR	SSIM	LPIPS
Cyclegan	7.97	0.4323	0.6625
Modification 1	7.69	0.4350	0.6675
Modification 2	8.06	0.4147	0.6616
Modification 3	7.82	0.4419	0.6611
Modification 4	7.90	0.4360	0.6652
Modification 5	8.20	0.4135	0.6443
Modification 6	7.69	0.4222	0.6665
Modification 7	7.97	0.4268	0.6651
Modification 8	7.79	0.4119	0.6685
Modification 9	8.02	0.4273	0.6728
Modification 10	8.07	0.4341	0.6487

Based on Table 2, Modification 5 consistently achieves the best overall performance. It records the highest PSNR score (8.20), indicating excellent pixel-level reconstruction of haze-free images. Its lowest LPIPS value (0.6443) also confirms strong perceptual quality, meaning the generated images appear more realistic to the human eye. This performance is likely due to its deeper architecture with eight residual blocks, enabling richer feature learning.

In terms of structural similarity, Modification 3 stands out with the highest SSIM score (0.4419). This suggests better preservation of image structure, likely due to the integration of a self-attention mechanism that helps focus on haze-affected regions. Although Modifications 5 and 10 also perform well in SSIM, they do not surpass the attention-driven consistency of Modification 3.

Comparatively, the original CycleGAN and Modification 6 show the weakest results across all metrics. The baseline model lacks architectural enhancements, while Mod 6, despite increasing filter depth, performs poorly due to limited feature

refinement. Overall, Modification 5 is the most balanced and effective model, excelling in accuracy, structure, and perceptual realism.



Figure 2. Model Output Visualization

Figure 2 displays the visual outputs of each model, although the actual images are not visible in this context. However, based on the architectural descriptions and perceptual evaluation criteria (such as attention blocks, residual depth, and perceptual losses), certain models particularly Modifications 3, 5, 8, and 10 are expected to generate sharper and more semantically accurate outputs due to the use of attention mechanisms, deeper residual connections, and perceptual or haze-aware loss functions.

Model Comparison (Table 3): The most effective model, based on both quantitative losses and qualitative output, is Modification 10, which incorporates haze-aware loss alongside VGG19-based perceptual loss. This configuration effectively balances loss reduction and visual quality, resulting in images that are cleaner and more distinct. Modifications with added residual blocks (e.g., Modifications 4, 5, and 7) also perform well in capturing complex features, making them suitable for applications requiring greater texture refinement. Conversely, Modification 9, which uses reduced filters to simplify the network, shows lower computational demand but yields comparatively modest outputs in terms of detail and clarity, making it the least effective among the tested configurations.

Table 3. Model Comparison

Model	Modification from Program 1	Explanation of Modification
CycleGAN	No modifications; this is the reference model.	Generator: Consists of 8 downsample layers (64, 128, 256, 512 filters) and 7 upsample layers (512, 256, 128, 64 filters). Discriminator: Has 4 downsample layers (64, 128, 256, 512 filters). Total Layers (Generator + Discriminator): 19 (15 + 4)
Mod 1	Added 2 residual blocks to the generator with Layer Normalization and dilated convolutions.	Generator: After the 8 downsample layers, two residual blocks are added, each consisting of two convolutional layers with dilation and normalization using Layer Normalization. This improves model stability and generalization without adding more downsample or upsample layers. Discriminator: No changes. Total Layers: 23 (19 + 4)
Mod 2	Added 4 residual blocks to the generator. This modification enhances the network's ability to capture image features.	Generator: Compared to Model 1, two additional residual blocks are added after the 8 downsample layers, bringing the total to 4 residual blocks. This deepens the model, enabling it to capture more complex features. Discriminator: No changes. Total Layers: 27 (23 + 4)
Mod 3	Added 4 residual blocks and 1 attention block to the generator to enhance important feature processing.	Generator: From Model 2, one attention block is added after the 4 residual blocks. This attention block is designed to help the model focus more on important areas in the image, improving output quality in relevant areas. Discriminator: No changes. Total Layers: 28 (24 + 4)
Mod 4	Added 6 residual blocks after the downsampling process in the generator.	Generator: From Model 3, two additional residual blocks are added after the 8 downsample layers, totaling 6 residual blocks. This additional depth allows the model to represent richer and more complex features. Discriminator: No changes. Total Layers: 31 (27 + 4)
Mod 5	Added 8 residual blocks to the generator with dilated convolutions to improve image processing.	Generator: This model expands Model 4 by adding two more residual blocks, totaling 8 residual blocks. The increased model depth strengthens its ability to capture finer details. Discriminator: No changes. Total Layers: 35 (31 + 4)
Mod 6	Added 4 residual blocks after downsampling, this time with more filters.	Generator: From Model 5, one additional downsampling layer with 1024 filters is added after the last 512-filter downsampling layer, before the 4 residual blocks. This addition allows the model to handle larger scale features and enhances model capacity. Discriminator: No changes. Total Layers: 37 (33 + 4)
Mod 7	Added 8 residual blocks with dilated convolutions in the generator.	Generator: From Model 6, four additional residual blocks are added, totaling 8 residual blocks. This increases model depth, enhancing its ability to learn from more complex data. Discriminator: No changes. Total Layers: 41 (37 + 4)
Mod 8	Integrated VGG19 for perceptual loss and added 4 residual blocks. Perceptual loss added after the upsampling stage.	Generator: This model expands Model 7 by adding another downsampling layer with 1024 filters before the residual blocks, further increasing the model's capacity to handle large and complex features. Discriminator: No changes. Total Layers: 43 (39 + 4) Total Lapisan (Generator + Discriminator) : 43 (39 + 4)
Mod 9	Added 4 residual blocks with fewer filters to simplify training on a smaller network.	Generator: Uses a lighter generator with 16 and 32 filters in the initial layers. Discriminator: No changes. Total Layers: 43 (39 + 4)
Mod 10	Added Haze-Aware Loss and used the same generator structure as Model 8.	Generator: From Model 8, in addition to retaining residual blocks and VGG19, Haze-Aware Loss is added to further optimize the model for dehazing tasks. This ensures that the model effectively handles noise and haze present in the images. Discriminator: No changes. Total Layers: 43 (39 + 4)

IV. DISCUSSION

A. Overparameterization Potential

To address the potential risk of overparameterization, Figure 3 presents the estimated GFLOPS (Giga Floating Point Operations per Second) for each model variant. GFLOPS serves as a computational complexity proxy, where higher values typically indicate increased model depth, parameter count, and training cost.

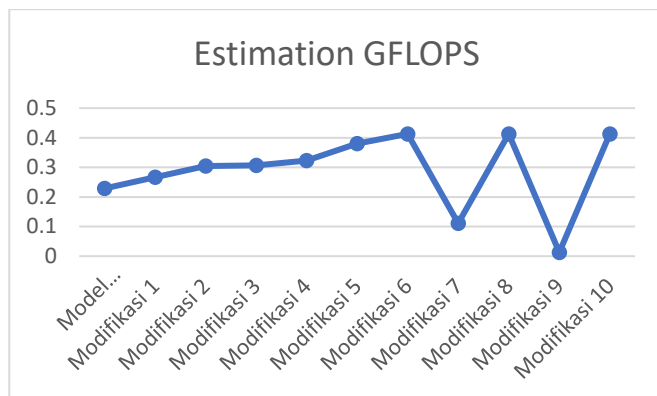


Figure 3. GLOPS

As shown in Figure 3, computational complexity increases progressively from the Reference Model to Modification 5, peaking at Modifications 6 and 7. These two configurations introduced additional downsampling layers and deeper residual stacks, resulting in the highest estimated GFLOPS among all variants. Interestingly, Modifications 8 and 10, despite incorporating perceptual loss via VGG19 and deeper architectures, maintain relatively moderate GFLOPS values, indicating architectural efficiency in balancing performance and computation.

In contrast, Modification 9 exhibits the lowest GFLOPS, as it employs a lightweight generator with reduced filter sizes, making it the most computationally efficient but also the least performant in terms of visual quality.

These variations demonstrate that architectural enhancements do not always lead to exponential increases in computational cost. However, the higher GFLOPS observed in Modifications 6 and 7 suggests potential overparameterization, which may lead to longer training times, increased memory usage, and the risk of overfitting particularly on limited or imbalanced datasets.

Therefore, when selecting the optimal model, it is essential to consider the trade-off between quantitative performance metrics (PSNR, SSIM, LPIPS) and computational efficiency (GFLOPS) to ensure practical applicability.

B. Advantages and Disadvantages of the Proposed Model

The proposed model has several advantages. It supports unpaired training using CycleGAN, making it suitable for dehazing tasks where paired data are unavailable. The use of residual blocks, attention mechanisms, and perceptual loss improves image clarity and perceptual quality, as seen in improved PSNR, SSIM, and LPIPS scores. These

architectural improvements allow the model to better capture haze-related features.

However, these benefits come at the cost of higher computational complexity. Models with deep structures and VGG19-based loss (e.g., Mod 5 and 10) require more memory and training time, as shown by increased GFLOPS. This limits their use in real-time or embedded applications and raises concerns of overfitting on small datasets [14].

C. Potential for Use in Real-World Applications

This model is well-suited for real-world applications that require clear visuals in hazy conditions. In autonomous driving, it can improve visibility and object detection in fog [15]. It is also useful in outdoor surveillance, aerial imaging, and remote sensing, where clearer images support better analysis and decision-making. For consumer use, it can enhance outdoor photos and videos in mobile devices.

However, to enable deployment in low-resource environments or real-time systems, further optimization like pruning or model compression is needed [16].

V. CONCLUSION

The degradation of image quality caused by haze, characterized by reduced contrast and clarity, presents a significant challenge in various fields such as autonomous driving and remote sensing. This study addresses the issue by developing several modified versions of the CycleGAN model for the task of haze removal (dehazing). The proposed modifications include the addition of residual blocks, attention mechanisms, deeper convolutional layers, VGG19-based perceptual loss, and haze-aware loss to enhance image clarity and realism.

Quantitative evaluations using PSNR, SSIM, and LPIPS demonstrate that Modification 5 achieves the best overall performance with the highest PSNR (8.20) and lowest LPIPS (0.6443), indicating superior pixel accuracy and perceptual quality. Modification 3 achieves the highest SSIM score (0.4419), likely due to the attention mechanism that enhances structural consistency. Meanwhile, qualitative results reveal that Modification 10, which integrates both perceptual and haze-aware loss, produces the most visually pleasing outputs with improved detail preservation and more natural dehazing effects. These findings are supported by stabilized generator losses and reduced discriminator losses, particularly in Modifications 8 and 10.

Overall, Modification 10 is considered the best-performing model in terms of both visual quality and loss behavior, while Modifications 5 and 3 offer strengths in specific quantitative metrics. These results confirm that the proposed architectural enhancements to CycleGAN successfully address the complexities of haze removal. This advancement contributes meaningfully to deep learning-based dehazing technologies and can support various real-world applications requiring improved image visibility and quality.

REFERENCES

- [1] D. Engin, A. Genç, and H. K. Ekenel, "Cycle-Dehaze: Enhanced CycleGAN for Single Image Dehazing," May 2018, [Online]. Available: <http://arxiv.org/abs/1805.05308>
- [2] Y. Qiu, Y. Lu, Y. Wang, and H. Jiang, "IDOD-YOLOV7: Image-Dehazing YOLOV7 for Object Detection in Low-Light Foggy Traffic Environments," *Sensors*, vol. 23, no. 3, Feb. 2023, doi: 10.3390/s23031347.
- [3] S. Lee, S. Yun, J. H. Nam, C. S. Won, and S. W. Jung, "A review on dark channel prior based image dehazing algorithms," Dec. 01, 2016, *Springer International Publishing*. doi: 10.1186/s13640-016-0104-y.
- [4] D. Berman, T. Treibitz, and S. Avidan, "Non-local Image Dehazing," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Dec. 2016, pp. 1674–1682. doi: 10.1109/CVPR.2016.185.
- [5] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: An end-to-end system for single image haze removal," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187–5198, Nov. 2016, doi: 10.1109/TIP.2016.2598681.
- [6] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "An All-in-One Network for Dehazing and Beyond," Jul. 2017, [Online]. Available: <http://arxiv.org/abs/1707.06543>
- [7] X. Liu, Y. Ma, Z. Shi, and J. Chen, "GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing," Aug. 2019, [Online]. Available: <http://arxiv.org/abs/1908.03245>
- [8] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," in *Proceedings of the IEEE International Conference on Computer Vision*, Institute of Electrical and Electronics Engineers Inc., Dec. 2017, pp. 2242–2251. doi: 10.1109/ICCV.2017.244.
- [9] J. Chen, C. Wu, H. Chen, and P. Cheng, "Unsupervised dark-channel attention-guided cyclegan for single-image dehazing," *Sensors (Switzerland)*, vol. 20, no. 21, pp. 1–15, Nov. 2020, doi: 10.3390/s20216000.
- [10] Y. Liu, H. Al-Shehari, and H. Zhang, "Attention mechanism enhancement algorithm based on cycle consistent generative adversarial networks for single image dehazing," *J Vis Commun Image Represent*, vol. 83, Feb. 2022, doi: 10.1016/j.jvcir.2021.103434.
- [11] H. Ni and W. Su, "A Self-Attention CycleGAN for Unsupervised Image Hazing," *Big Data and Cognitive Computing*, vol. 9, no. 4, Apr. 2025, doi: 10.3390/bdcc9040096.
- [12] J. Wang, Y. zhang, and Z. Liu, "Unsupervised dehazing of multi-scale residuals based on weighted contrast learning," Aug. 22, 2024. doi: 10.21203/rs.3.rs-4812948/v1.
- [13] H. Liu, X. Li, and T. Tan, "Interaction-Guided Two-Branch Image Dehazing Network," 2024. [Online]. Available: <https://link.springer.com/conference/accv>
- [14] X. Chen *et al.*, "Unpaired Deep Image Dehazing Using Contrastive Disentanglement Learning," Mar. 2022, [Online]. Available: <http://arxiv.org/abs/2203.07677>
- [15] A. Jain, "Feature Fusion Attention Network with CycleGAN for Image Dehazing, De-Snowing and De-Raining," Mar. 2025, [Online]. Available: <http://arxiv.org/abs/2503.06107>
- [16] D. Yang, Z. Zhu, H. Ge, H. Qiu, H. Wang, and C. Xu, "A Lightweight Neural Network for the Real-Time Dehazing of Tidal Flat UAV Images Using a Contrastive Learning Strategy," *Drones*, vol. 8, no. 7, Jul. 2024, doi: 10.3390/drones8070314.

