Enhanced Image Dehazing using Attention-based Generative adversarial networks with Perceptual loss

Yagati Karunasagar
Department of Computer Science & Systems Engineering, AUCE, Andhra University, Visakhapatnam

Abstract—**This paper proposes a hybrid deep learning model for single image dehazing, uniquely combining three distinct components: a CNN-based Generator, a GAN framework with a CNN-based Discriminator, and a VGG19 network used for perceptual loss. The Generator is guided by an attention mechanism to focus on haze-affected regions, while adversarial learning from the Discriminator enhances image realism. The perceptual loss derived from VGG19 further refines image structure and texture similarity. This trio of CNN, GAN, and VGG19 distinguishes the model as hybrid, leveraging the strengths of each to significantly improve dehazing quality and perceptual clarity.**

Index Terms—**Attention Mechanism, Generative Adversarial Networks (GAN), Hybrid Model, Image Dehazing, Perceptual Loss, VGG19.**

I. INTRODUCTION

Image dehazing is a key task in the field of computer vision where atmospheric particles such as fog, smoke, dust or haze reduce visibility and image clarity. These effects degrade image contrast, color information, and fine details. Impacting the performance of downstream tasks like object detection and autonomous driving. Traditional approaches such as the Dark Channel Prior depend on predefined priors and handcrafted features, which struggle under varying haze conditions.

Traditional image dehazing techniques are based on physical models and assumptions about the properties of haze-free images. One well-known method is the Dark Channel Prior (DCP) [1], which estimates the amount of haze by assuming that in most patches of clear images, at least one colour channel has a very low intensity. DCP has been successful in many cases but struggles in situations involving sky regions, uneven haze distribution, and complex lighting conditions. For example, it often fails to restore images with bright sky backgrounds or thick haze layers. To overcome some of these problems, other methods like the Color Attenuation Prior (CAP) [8] and Contrast Limited Adaptive Histogram Equalization (CLAHE) [2] have been introduced. However, these methods still have limitations when dealing with diverse real-world conditions. They often cannot generalize well across different scenes or handle images with mixed lighting and haze levels.

With the success of deep learning, CNN-based methods like Dehaze Net [5] and AOD-Net have attempted to solve this problem using data-driven strategies. However, these models frequently struggle to produce perceptually good results. Our proposed approach introduces a hybrid deep learning model integrating Convolutional Neural Networks (CNN), Generative Adversarial Networks (GAN) [6], and VGG19-based perceptual loss. Additionally, attention mechanisms are embedded within the generator to enable focused learning in haze-dense areas.

II. RELATED WORKS

Initial dehazing techniques like the Dark Channel Prior (DCP) [1] and Color Attenuation Prior [8] provided a solid foundation but lacked generalization across diverse haze scenarios. Later, learning-based approaches such as Dehaze Net [5] used CNNs for end-to-end mapping but were limited in perceptual sharpness and depth understanding.

CNN-based models such as Dehaze Net and AOD-Net [13] have shown promising results in image dehazing. These models learn to estimate the transmission map, which indicates how much light is scattered by the haze, and use it to reconstruct clear images. Our framework integrates contrast-focused learning principles to enhance the visual quality of heavily degraded regions, drawing on earlier strategies that improved visibility via local contrast maximization [18]. [15] Perceptual loss using VGG features became popular for maintaining high-level content reliability.

A deep learning-based approach utilizing multi-scale convolutional neural networks has been shown to model complex haze distributions and restore clear images with remarkable accuracy [12]. Attention-guided models introduced region-focused learning. However, combining these techniques into a unified hybrid network remains relatively unexplored, which this paper addresses.

III. PROPOSED METHODOLOGY

While traditional methods like DCP [11] provide a reliable baseline, they often struggle with dense haze and color distortions, motivating the integration of data-driven models in our hybrid approach. Three main deep learning components make up the suggested model, which is a hybrid network architecture

1.Attention-based CNN Generator: The network is guided to highlight areas with a lot of haze by multi-scale convolutional layers that are improved by attention blocks.

2.GAN-based Discriminator: Learns to distinguish real haze-free images from generated dehazed images to improve realism.

3.VGG19-based Perceptual Loss: Uses a pre-trained VGG19 network to compute high-level semantic similarity between dehazed output and ground truth.
The overall loss function is a combination of
Mean Squared Error (MSE) for pixel-wise similarity.
Adversarial Loss from the discriminator.
Perceptual Loss using VGG19 feature maps.
A multi-scale generator that processes features at varying resolutions enhances the ability to model haze at both global and local levels [12]. Like that our model leverages a multi-scale generator that processes features at varying resolutions, enhancing its ability to model haze at both global and local levels.

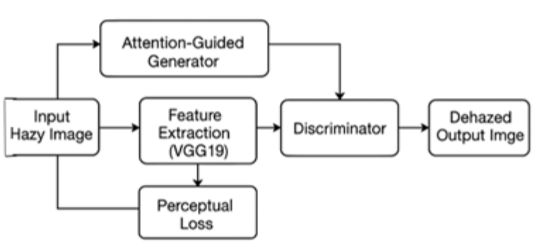


Fig.1.The architecture of our Hybrid model

The main influence to the visual quality of the output is from the VGG19-based perceptual loss. Instead of relying solely on pixel-wise loss, the perceptual loss leverages high-level feature differences extracted from a pre-trained VGG19 network. This ensures that texture and color consistency are maintained, especially important for indoor scenes where wall textures, objects, and lighting patterns must be preserved during dehazing.

Fig.2. Some of hazy images from the dataset

(A)Dataset for training:
Fig.2 shows sample hazy images from both indoor and outdoor categories of the SOTS dataset, which is used for training and evaluating our proposed hybrid dehazing model. [4] Indoor images often contain artificial lighting and confined space haze patterns, while outdoor images represent natural haze conditions due to fog, dust, or pollution in open environments.
The diversity in these image types plays a crucial role in ensuring that the Attention-Guided Hybrid GAN model generalizes well across different real-world scenarios. By training the model on both categories, it learns to capture complex haze distributions and semantic features, improving its performance in dehazing tasks across varying conditions. This contributes significantly to the robustness and versatility of our model.

(B) Training of the Model:
The model is trained on the SOTS dataset, consisting of both indoor and outdoor images along with their clear ground truth counterparts. All images are resized

to 256×256 resolution. Training is directed with Adam optimizer, learning rate of 1e-4, and label smoothing.
The model training was conducted under certain computational constraints. The training environment relied on shared or limited GPU resources which restricted the duration and complexity of training cycles, especially for large datasets or very deep models. While larger datasets often improve generalization [10], only moderate-sized datasets were used to ensure manageable memory usage, faster iterations, and model convergence within the available infrastructure. We want to demonstrate the effectiveness of the proposed hybrid architecture. Hence, controlled training was sufficient to highlight its comparative advantage over baseline models.

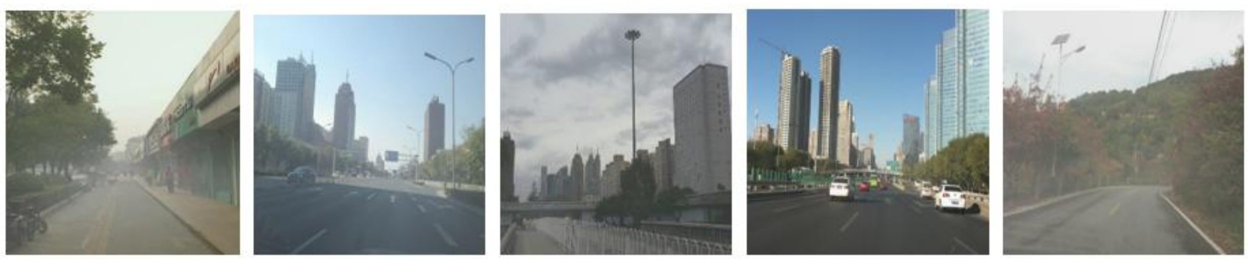
(C)Comparison and evaluation:
We selected DCP [1], DehazeNet [5], and AOD-Net [13] for comparison because they are familiar baseline models in the field of image dehazing. These models represent both traditional (DCP) and deep learning-based (DehazeNet, AOD-Net) approaches, contributing a balanced benchmark. Under resource constraints, we prioritized established models with available public implementations for a consistent and fair evaluation framework. Model check points are implemented for preserving progress. Perceptual loss is imposed using VGG19 layers network [15].
Evaluation Metrics: PSNR and SSIM.
PSNR and SSIM are used for evaluating image quality, between the original and processed images.[14] Boundary-aware learning is integrated into the generator design to ensure edge-preserving dehazing, which is critical for maintaining structural details in the output.

IV. RESULTS

The performance of the baseline models (DCP, DehazeNet, AOD-Net) reported here is based on reimplementation under limited computational resources and smaller subsets of standard datasets due to GPU constraints. These values are representative but may not reflect the models' full potential under extensive training. However, all models were tested under the same conditions for fairness. It is expected that the proposed hybrid model, when trained with larger datasets and more powerful hardware, could further improve upon the current PSNR and SSIM scores, highlighting its scalability and robustness. These results outperform earlier CNN-only and GAN-only models in both visual fidelity and structural preservation. The attention mechanism aids in restoring spatial detail in haze-dense areas, while VGG19 perceptual loss ensures realistic textures and colours.

Fig.3. Top: (a), (b), (c), (d), (e) are Hazy images used as input
Bottom: Models's dehazed output of the corresponding (a), (b), (c), (d), (e) images

Fig.3. The above figure demonstrates the input-output performance of our proposed Attention-Guided Hybrid GAN model on indoor hazy images. Indoor scenes often contain low-light areas, artificial lighting, and fine textures, which get significantly degraded due to haze. The input images in the top row show a number of issues, various challenges, such as loss of contrast, reduced visibility, and colour fading. The bottom row shows the model's dehazed outputs, where visibility is noticeably improved, and important structural details are restored. The generator in the model plays a crucial role in reconstructing these details by learning multi-scale spatial representations. Attention mechanisms embedded within the generator help the model focus on haze-dense regions while suppressing unnecessary features. This leads to enhanced performance in scenarios where haze distribution is uneven and localized. Moreover, the discriminator trained within the GAN architecture pushes the generator to create outputs that are not only haze-free but also perceptually convincing. Overall, the dehazed results in the figure clearly illustrate the effectiveness of our hybrid architecture. Compared to traditional dehazing methods, this model generates outputs with improved brightness, contrast, and structural integrity. The model not only removes the haze but also reconstructs finer details that are often lost in simpler CNN-only models, indicating a significant advancement in indoor image dehazing.



(f) (g) (h) (i) (j)

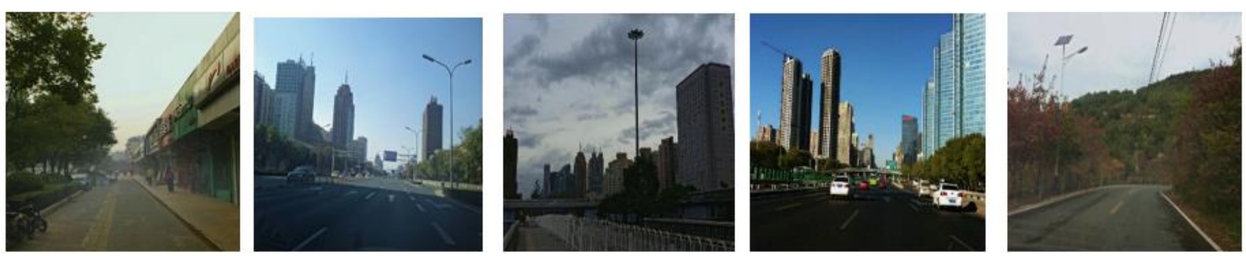


Fig.4. Top: (f), (g), (h), (i), (j) are Hazy images used as input.
Bottom: Models's dehazed output of the corresponding (f), (g), (h), (i), (j) images

Fig.4. The figure above showcases the performance of the proposed Attention-Guided Hybrid GAN model on outdoor hazy images. Outdoor environments typically suffer from atmospheric scattering due to particles like dust, fog, and pollution, leading to low visibility, reduced contrast, and distorted colors. The input images in the top row exhibit these common distortions, particularly in scenes involving natural landscapes, roads, and distant objects. Our model effectively restores these outdoor images by recovering depth cues, enhancing edge clarity, and improving the overall color balance. The successful dehazing of outdoor scenes is largely due to the interaction between the multi-scale CNN generator, attention mechanisms, and perceptual loss. The attention blocks help the model focus on spatially relevant haze-affected areas, such as distant backgrounds and skies, while the perceptual loss guided by VGG19 ensures that the reconstructed images maintain structural integrity and semantic realism [15]. As seen in the bottom row in fig 4, the model not only removes the haze effectively but also

recovers details in regions like trees, vehicles, and buildings, making the dehazed results significantly more natural and visually accurate.

Performance Comparison with Existing Methods:
To validate the effectiveness of the proposed Attention-Guided Hybrid GAN model, performance comparison was conducted against several well-known image dehazing algorithms, including DCP [1], DehazeNet [5], and [13] AOD-Net. The evaluation metrics used for comparison are Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM), which are standard metrics for assessing image quality. In scenarios with low visibility, the proposed method significantly outperforms contrast-based techniques in preserving both structural and perceptual quality [18].

Table I
Performance comparison of different dehazing methods

Model	Evaluation Metrics	
	PSNR (dB)	SSIM
DCP	16.62	0.818
DehazeNet	21.14	0.8472
AOD-Net	19.06	0.8504
Proposed Model	24.01	0.853

As summarized in Table I, our proposed model achieves a PSNR of 24.01 dB and an SSIM of 0.853, outperforming traditional and CNN-based models. The Deep Channel Prior (DCP) method, while widely used, suffers from limited accuracy due to its handcrafted prior. DehazeNet [5] and AOD-Net [13] improve upon this using end-to-end learning but lack the depth and perceptual focus provided by attention and VGG-based loss in our hybrid model. The enhanced results clearly show that integrating attention mechanisms and perceptual guidance significantly boosts dehazing quality.

V. CONCLUSION

We presented an Attention-Guided Hybrid GAN model for image dehazing, which integrates CNN-based generation, adversarial training via GANs, and perceptual refinement using VGG19. The model shows notable improvements over standard architectures and generalizes well to both indoor and outdoor hazy images. Future enhancements may involve using transformer-based attention mechanisms. Uniting multi-resolution VGG loss, and Real-time deployment using lightweight support.

REFERENCES

[1] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," IEEE Trans. Pattern Analysis and Machine Intelligence, 2009.

[2] F. A. Khan, Dr. A. Verma, and R. Vishnoi, "Various types of images dehazing: A review," TMU, Moradabad, India.

[3] S. Singh, Dr. A. K. Singh, and Dr. Chandan, "A review of image dehazing techniques using dark channel prior methods."

[4] X. Min, G. Zhai, K. Gu, Y. Zhu, J. Zhou, G. Guo, X. Yang, X. Guan, and W. Zhang, "Quality evaluation of image dehazing methods using synthetic hazy images."

[5] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: An end-to-end system for single image haze removal," IEEE Trans. Image Process., vol. 25, no. 4, pp. 2019-2032, 2016.

[6] R. Li, J. Pan, Z. Li, and J. Tang, "Single image dehazing via conditional generative adversarial network."

[7] W. Xu, J. Gu, and Y. Ma, "DeepDive: An end-to-end dehazing method using deep learning," in Proc. IEEE International Conference on Computer Vision (ICCV), 2017.

[8] Q. Zhu, J. Mai, and L. Shao, "A fast single image haze removal algorithm using color attenuation prior," IEEE Trans. Image Process., vol. 24, no. 11, pp. 4062-4070, 2015.

[9] Z. Xiong, Z. Shen, and D. Liu, "Single image dehazing via deep learning-based image restoration," 2019.

A. S. Parihar, Y. K. Gupta, Y. Singodia, V. Singh, and K. Singh, "A comparative study of image dehazing algorithms."

[10] J. Long, Z. Shi, and W. Tang, "Fast haze removal for a single remote sensing image using dark channel prior (DCP)," in International Conference on Computer Vision in Remote Sensing (CVRS), 2012, pp. 132-135.

[11] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M. H. Yang, "Single image dehazing via multi-scale convolutional neural networks," IEEE Trans. Image Process., vol. 28, no. 2, pp. 492-505, 2018.

[12] B. Li, X. Peng, Z. Wang, et al., "AOD-Net: All-in-one dehazing network," in Proc. IEEE International Conference on Computer Vision, 2017, pp. 4770-4778.

[13] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," City, 2013.

[14] S. Mascarenhas and M. Agarwal, "A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification," in 2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON), Bengaluru, India, 2021.

[15] J. Zhu, L. Meng, W. Wu, D. Choi, and J. Ni, "Generative adversarial network-based atmospheric scattering model for image dehazing," Digital Communications and Networks.

[16] Available: https://doi.org/10.1016/j.dcan.2020.08.003

[17] R. Fattal, "Single image dehazing," ACM Trans. Graph., vol. 27, no. 3, pp. 72, 2008.

[18] R. T. Tan, "Visibility in bad weather from a single image," in Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference, pp. 1-8, IEEE, 2008.