Research Progress in Image Dehazing Methods

Kejia Huang [1, *], Yuxue Liu [2]

[1] Portland Institute, Nanjing University of Posts and Telecommunications, Nanjing, China

[2] Department of Automation, North China Electric Power University, Baoding, China

* Corresponding Author Email: p22000503@njupt.edu.cn

Abstract. Haze in the environment, caused by airborne particles, reduces image clarity and poses challenges for subsequent in-depth analysis and processing of images. Therefore, it is imperative to evaluate the advantages and disadvantages of various image dehazing methods, identify the challenges they face, and explore prospects. This paper focuses on image dehazing techniques, discussing the latest research advancements with an emphasis on comparing deep learning-based and traditional methods. It delves into the principles and application domains of different dehazing approaches, including image enhancement methods, image restoration methods, and deep learning-related techniques. Research indicates that various image dehazing methods face challenges such as detail processing, colour distortion, and computational complexity. Future developments should integrate the strengths of traditional and deep learning approaches, and improve network architectures, evaluation systems, and generalization capabilities to further advance image dehazing technology in practical applications.

Keywords: Image Dehazing; Image Enhancement; Image Restoration; Deep Learning.

1. Introduction

In the field of digital image processing, image dehazing is a critical task aimed at addressing the degradation of image quality caused by atmospheric scattering, thereby enhancing the visual quality and analytical efficiency of images. With the continuous evolution of digital image processing technologies, image dehazing techniques have garnered widespread attention.

Traditional image dehazing techniques primarily rely on analysing the physical properties and statistical patterns of images, offering certain advantages in computational efficiency. For instance, DDNet (Density and Depth-Aware Network) employs a dual-path structure capable of enhancing both local and global features of an image, making it suitable for real-time low-light image enhancement. However, it may result in the loss of details and colour deviation during dehazing [1]. Similarly, NAFSSR (Stereo Image Super-Resolution Using NAFNet), based on the degradation model of hazy images, attempts to restore haze-free original images while enhancing resolution, making it suitable for stereo image super-resolution processing. Nevertheless, traditional image dehazing techniques require manually designed features and struggle to handle the diversity and complexity of images effectively [2].

In recent years, the rise of deep learning technology has brought transformative breakthroughs to image dehazing. Deep learning methods leverage multi-layer convolutional neural networks to automatically extract multi-level features from images, significantly improving dehazing performance [3]. However, despite their outstanding results, deep learning methods often require large amounts of training data and substantial computational resources, posing high demands on hardware platforms [4].

This study conducts a comprehensive analysis of existing image dehazing techniques, comparing the advantages and disadvantages of deep learning and traditional methods. It explores the application background of deep learning in image processing. The research method involves a thorough examination and comparison of the principles and performance of various dehazing techniques. The purpose of this study is to provide references and guidance for the future development of image dehazing techniques, assisting researchers and practitioners in selecting appropriate dehazing methods. This will promote the continuous advancement of image dehazing technology to meet the growing demands for image processing, making significant contributions in fields such as computer vision, remote sensing, and intelligent transportation, with both theoretical and practical significance.

2. Major Techniques for Image Dehazing

2.1. Image Enhancement

Image enhancement techniques are generally based on physical models and prior

Knowledge. These techniques rely on the physical model of hazy images, which is typically expressed as:

$$I(x) = J(x)\, t(x) + A\, (1 - t(x)) \tag{1}$$

Where I(x) is the hazy image, J(x) is the dazed image, A is the atmospheric light, and t(x) is the transmission map [5]. Prior knowledge is often utilized to assist the dehazing process, providing a foundation for subsequent operations.

Additionally, such methods require estimation and adjustment of the transmission map. Using the dark channel prior, the transmission map can be estimated under the assumption of known atmospheric light by applying a minimum operation within local regions. Morphological operations can then refine the transmission map, enhancing detail [6].

After obtaining an initial dehazed result, further image enhancement operations are needed. For instance, gamma adjustment can be employed: the colour model of the dehazed image is converted from Red-Green-Blue (RGB) to Hue-Saturation-Value (HSV), intensity is adjusted, and the result is converted back to RGB, making image details clearer and colours more vibrant [6]. Additionally, contrast-limited adaptive histogram equalization (CLAHE) combined with discrete wavelet transform (DWT) can be applied. First, the local contrast of the image is enhanced, followed by decomposition, filtering, and reconstruction, improving the image's contrast and sharpness. Finally, merging images processed with different enhancement techniques produces a better dehazing effect [5].

2.2. Image Restoration

Image restoration techniques are primarily based on the degradation model of hazy images, attempting to recover the haze-free original image. Commonly used methods include Wiener filtering, Kalman filtering, blind deconvolution, and regularization.

NAFSSR is a typical image restoration algorithm that extracts single-view features by stacking NAFNet blocks, with a Simple Cross-Attention Module (SCAM) inserted after each NAFBlock for cross-view feature fusion. During training, it uses random depth for regularization and channel shuffle for data augmentation, while during testing, it employs Test-Time Light Source Correction (TLSC) to address inconsistencies between training and testing [2].

Image restoration algorithms need to strike a balance between noise removal and detail preservation. Therefore, selecting appropriate image restoration methods and parameters depends on the specific application scenarios and requirements.

2.3. Deep Learning

Deep learning methods leverage deep neural network models for feature learning, enabling the automatic extraction of effective representations from input data and achieving efficient image feature extraction. This ultimately results in more accurate dehazing performance.

Among these methods, the Feature Fusion Attention Network (FFA-Net) is a network architecture specifically designed for image dehazing. It incorporates unique feature fusion and attention mechanisms, laying a solid foundation for subsequent tasks such as object detection. Meanwhile, YOLOv8n is a prominent model in the field of object detection, known for its high efficiency in detecting objects. It can quickly and accurately identify various objects in images.

When FFA-Net is combined with YOLOv8n, it significantly improves the detection performance for hazy images. FFA-Net preprocesses hazy images by removing haze interference, and then the

YOLOv8n model is used to train for object detection. During this process, Mosaic and Mixup algorithms are employed to further optimize the results [7].

Mosaic combines four images into one, enriching the background information of the training images, allowing the model to learn more features from different scenes, and enhancing its robustness.

Mixup, on the other hand, mixes two images and their corresponding labels in a certain proportion to generate new training samples, enabling the model to better adapt to different scenarios and improving its generalization ability.

These two algorithms are applied to the widely used KITTI dataset (a benchmark dataset in computer vision) for data augmentation. Additionally, a small-object detection layer is introduced, improving the accuracy of detecting small-sized objects in images. The model also incorporates a global attention mechanism (GAM) to focus the network on key areas of the image, enhancing detection accuracy [7]. These enhancements lead to an improved version of the YOLOv8n model.

Deep learning methods demonstrate remarkable effectiveness in processing large-scale, complex data. However, they also have certain limitations, such as the need for large amounts of training data and computational resources to ensure model performance. Moreover, deep learning models often require techniques like parameter tuning and regularization to avoid overfitting and improve generalization capabilities.

3. Recent Research Advances in Image Defogging Techniques

3.1. Image Enhancement - DDNet Density and Depth-Aware Network

3.1.1. Working Principle

DDNet is a density and depth-aware network designed for foggy scenarios. It improves target detection performance by leveraging depth and fog density information without requiring explicit image dehazing operations. The network consists of two components: the Density-Aware Attention Network (DAANet) and the Depth-Aware Non-Local Contextual Network (DNCNet) [1]. It has achieved notable results on foggy driving datasets.

As shown in Figure 1, DDNet employs specific network architecture. In DAANet, the Fog Density Perception Module (FDPM) is used to predict fog density. This is followed by the Density-Sensitive Attention Module (DSAM), which enhances features using fog density information. DNCNet, on the other hand, uses the Depth Prediction Module (DPP) to predict depth maps and then leverages depth information to explore spatial positional relationships, capturing long-range dependencies to enhance feature discriminative capabilities.

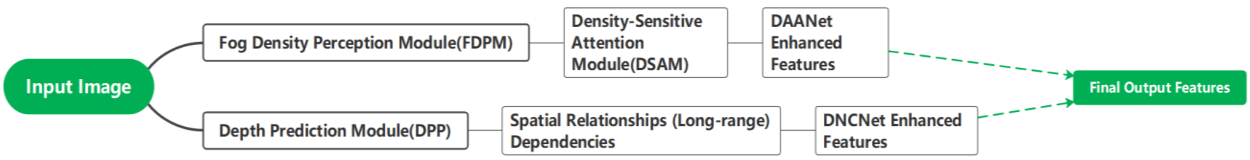

Fig 1. DDNet Working Principle.

3.1.2. Applications

The DDNet depth-aware network addresses some of the shortcomings of traditional image enhancement algorithms [4]. Traditional methods have made some attempts at dehazing, but they have numerous limitations. For example:

Histogram Equalization (HE) can improve brightness but struggles to distinguish between noise and useful information.

Methods based on specific theories, such as the Retinex theory and its related approaches, can decompose images to obtain the underlying normal light image. However, they face challenges such as insufficient brightness enhancement and difficulty balancing noise suppression with edge feature preservation.

Other approaches based on camera response models or atmospheric scattering models often fail to simultaneously meet the requirements of detail preservation, lighting enhancement, and computational efficiency.

In contrast, the DDNet depth-aware network overcomes these issues. However, it may lose some details during the dehazing process and, in certain cases, exhibit colour deviations. Therefore, the DDNet depth-aware network is particularly suitable for scenarios requiring real-time processing of low-light image enhancement.

3.2. Image Restoration - NAFSSR Network

3.2.1. Working Principle

NAFSSR is a network designed for stereo image super-resolution, which enhances images by incorporating depth information [6]. As shown in Figure 2, it employs stacked NAFNet blocks to extract single-view features, with a Simple Cross-Attention Module (SCAM) inserted after each NAFBlock to achieve cross-view feature fusion. During training, it uses random depth for regularization and channel shuffle for data augmentation. During testing, it adopts Test-Time Light Source Correction (TLSC) to address inconsistencies between training and testing.

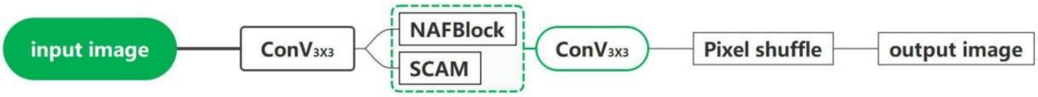

Fig 2. NAFSSR Working Principle.

3.2.2. Applications

NAFSSR has made advancements in the field of stereo image super-resolution and can be applied to related image restoration tasks. It inherits the strengths of NAFNet by incorporating a SCAM to fuse features between views, making it suitable for image restoration in stereo scenes. Additionally, it employs training strategies such as random depth and channel shuffle, while addressing inconsistencies between the training and testing phases. These approaches have led to notable performance improvements across multiple datasets.

However, NAFSSR has the drawback of high computational demands, which may make it unsuitable for real-time processing. Therefore, the NAFSSR network is best suited for scenarios involving stereo images that require super-resolution processing, have sufficient computational resources, and do not emphasize real-time processing speed.

3.3. Deep Learning — Combining FFA-Net and Improved YOLOv8n Model

3.3.1. Working Principle

The combination of FFA-Net and an improved YOLOv8n model is used for image dehazing [7]. As shown in Figure 3, Hazy images captured under foggy conditions are input into the system, where the FFA-Net network performs dehazing preprocessing, enhancing the clarity and contrast of the images [7]. The dehazed images are then output and prepared for object detection.

The improved YOLOv8n model is used for training in object detection. To enhance the model’s generalization ability and detection performance, the Mosaic and Mixup algorithms are applied to the KITTI dataset for data augmentation.

Mosaic enriches the training images by combining four images into one, providing more diverse backgrounds for training and enabling the model to learn features from a variety of scenes.

Mixup blends two images and their corresponding labels in a certain proportion to generate new training samples, helping the model adapt to different situations and improving its generalization.

In the YOLOv8n model, a small-object detection layer is added to improve its ability to detect small-sized objects. Additionally, a Global Attention Mechanism (GAM) is introduced to enhance the network's ability to perceive global information. These improvements collectively result in the final improved YOLOv8n model [8].

The enhanced model is then used to perform object detection on the dehazed images. The results of the object detection are analysed to evaluate the performance of the model.

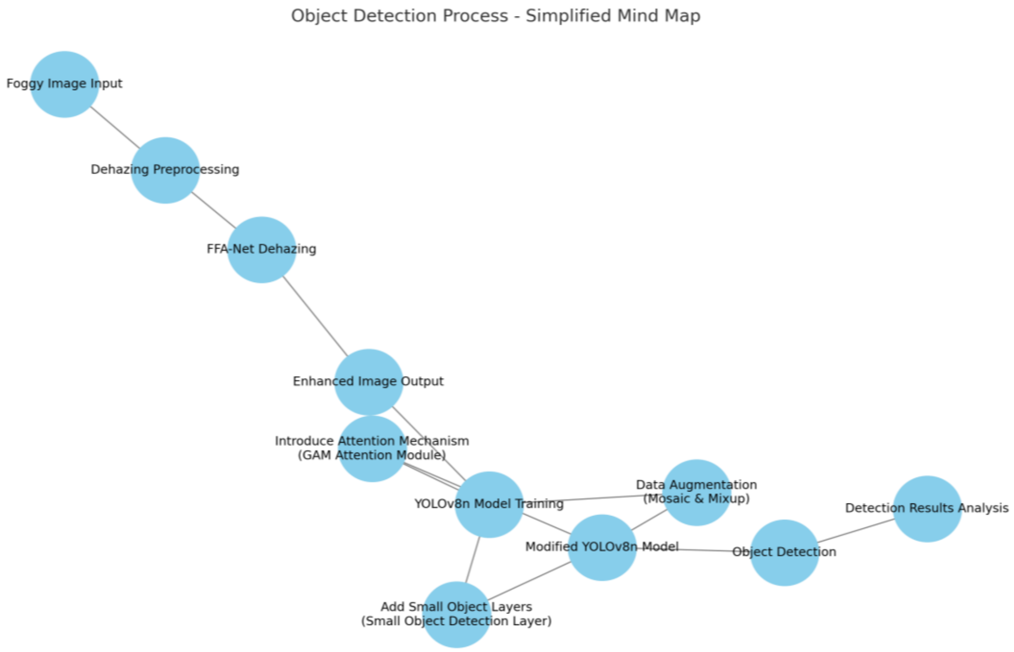

Fig 3. Object Detection Process.

3.3.2. Applications

The FFA-Net network features real-time capability and lightweight design, making it suitable for dehazing requirements during driving. The selected YOLOv8n model is the most lightweight in the YOLOv8 series. It employs a lightweight backbone network and design to reduce computational and memory overhead, enabling the model to achieve high inference speed on resource-constrained devices [4].

By performing data augmentation on the KITTI dataset and utilizing transfer learning strategies, the improved YOLOv8n model can better focus on detecting small objects, enhancing detection performance [9], which is crucial for real-time processing. However, the large number of parameters in the model increases computational demands, and the complex network structure adds to training difficulty [10].

Therefore, the combination of the FFA-Net network with the improved YOLOv8n model is well-suited for scenarios that require high detail recovery in images.

4. Conclusion

This paper introduces and analyses three types of images dehazing techniques to compare their primary application conditions. The main image dehazing techniques include traditional methods (image enhancement and image restoration) and deep learning methods.

Traditional image enhancement techniques, based on physical models and prior knowledge, offer certain advantages in computational efficiency. However, they require manually designed features and struggle to handle the diversity and complexity of images. Image restoration techniques rely on degradation models and use methods such as Wiener filtering to recover the original image. These techniques must balance noise reduction and detail preservation, often involving significant computational costs.

Deep learning methods utilize deep neural networks to automatically learn features, performing dehazing as a preprocessing step before detection. While these methods achieve excellent results and perform well in specific scenarios, they require large amounts of training data and computational resources, and their models are complex to fine-tune.

Although deep learning has achieved significant breakthroughs in the field of image dehazing, it still faces numerous challenges. Deep learning models have high hardware and computational resource requirements, limiting their application in resource-constrained environments. Additionally, their lack of interpretability makes it difficult to understand their decision-making processes, complicating optimization and improvement. Existing methods also perform sub-optimally in challenging scenarios such as dense fog or mixed rain and fog.

Future research could explore integrating the strengths of deep learning and traditional methods, for example, using traditional methods' prior knowledge to initialize deep learning models or embedding effective strategies from traditional algorithms into deep learning frameworks. Furthermore, more efficient neural network architectures should be designed to reduce computational resource demands and improve training and inference speeds.

In terms of dehazing evaluation, it is necessary to develop evaluation systems more aligned with human visual perception to guide model optimization. Additionally, a key focus should be enhancing the generalization capabilities of models to ensure more stable performance across diverse scenarios, thereby promoting the further development of image-dehazing technology in practical applications.

Acknowledgements

All the authors contributed equally, and their names were listed in alphabetical order.

Reference

[1] B. Xiao, J. Xie and J. Nie. DDNet: Density and depth-aware network for object detection in foggy scenes. International Joint Conference on Neural Networks (IJCNN), 2023, pp. 1-7.

[2] X. Chu, L. Chen and W. Yu. NAFSSR: Stereo Image Super-Resolution Using NAFNet. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2022, pp. 1238-1247.

[3] Zhang Mengwen, Sang Jianbin, Jia Mengbo, Wang Haiting, Yang Huicheng, Hu Yaocong, Xu Ke. Image Dehazing Method Based on Fast Fourier Transform and Deep Learning. Journal of Chongqing Technology and Business University,2024, pp. 1-10.

[4] H. Ullah, et al. Light-DehazeNet: A Novel Lightweight CNN Architecture for Single Image Dehazing . IEEE Transactions on Image Processing, 2021, 30:8968-8982.

[5] K. Kim, S. Kim, and K. Kim. Effective image enhancement techniques for fog-affected indoor and outdoor images. IET Image Processing, 2018, 12(4):465–471.

[6] Rong Wang and XiaoGang Yang. A fast method of foggy image enhancement. Proceedings of 2012 International Conference on Measurement, Information and Control, 2012, pp. 883–887.

[7] Xu Hui, Du Feng. Object detection algorithm for foggy driving conditions based on FFA-Net and improved YOLOv8n. Journal of Tianjin University of Technology and Education, 2024, 34(3): 40-49.

[8] Deng Xiang-yu, Li Jun-teng. Image dehazing algorithm based on adaptive sky region segmentation. Computer Engineering and Design, 2024, 45(9):2742-2748.

[9] Li, H., & Zhang, Y. "A Comprehensive Study on Traditional Image Enhancement Techniques in Image Dehazing". Journal of Image Processing Research, 2022, 35(2), 120 - 135.

[10] Zhang, J., Zhao, W., & Sun, M. "Deep Learning Based Image Dehazing: Advances and Challenges". IEEE Transactions on Pattern Analysis and Machine Intelligence,2021. 43(5), 1560 - 1575.