

Data Composition Strategy for FFA-Net Image Dehazing

Chae-bin Park¹, Hyo-jae Lee¹, Jae-Ho Nah^{1*}

Abstract

We propose a generalized training strategy for FFA-Net to achieve robust image restoration in diverse haze conditions. By combining synthetic and real-world datasets (e.g., ITS, OTS, Haze4K, and NH-HAZE), we enhance restoration stability across varying atmospheric scenarios. Furthermore, we leverage performance benchmarks to motivate the choice of FFA-Net alongside detailed dataset overviews. Experimental results demonstrated that the proposed data composition strategy—integrating Haze4K and NH-HAZE at a 1:1 ratio—significantly improves cross-domain generalization and stabilizes dehazing performance under non-homogeneous real-world haze. This provides a practical and generalizable solution for reliable real-world dehazing without requiring manual adjustment according to input haze conditions.

Key Words: Image Dehazing, FFA-Net, Training Strategy, Dehazing Network, Deep Learning.

I. INTRODUCTION

Haze is caused by the presence of atmospheric particles such as dust, fine particulates, and water droplets, which introduce complex noise into images captured by cameras. As a result, the visibility of outdoor scenes is significantly degraded, leading to reduced contrast, blurred object details, and distorted surface colors. This degradation not only lowers the visual quality of images but also severely affects the performance of various high-level computer vision tasks, including object detection and semantic segmentation. Consequently, image dehazing has been extensively studied as a crucial topic in the field of image restoration and enhancement, and it is widely regarded as a key problem for improving the reliability and performance of computer vision algorithms [1].

With the rapid development of deep learning techniques, CNN-based dehazing networks such as FFA-Net [2] have achieved promising results on synthetic haze datasets by leveraging effective feature fusion mechanisms. However, most existing dehazing models are primarily trained on synthetic datasets due to the difficulty of obtaining paired hazy–clear images captured in real-world scenarios.

Due to this data constraint, experiments are typically conducted by splitting a single synthetic dataset into training and test sets. As a result, models tend to overfit to the characteristics of a specific synthetic dataset, and their performance is evaluated only within these domain-

restricted settings.

However, models that achieve high performance on a specific dataset inevitably suffer from limited generalization ability—commonly referred to as the domain gap—when applied to environments with different characteristics (e.g., transitions from indoor to outdoor scenes, or discrepancies between synthetic and real-world data), often resulting in color distortion and loss of fine details [3]. In particular, haze encountered in real-world scenarios such as autonomous driving, drones, and surveillance systems is typically non-homogeneous, dynamically changing across frames, and time-varying, making stable and consistent restoration a highly challenging task. Nevertheless, models trained on a single dataset struggle to adapt promptly to such dynamic conditions and require manual model switching depending on environmental variations (e.g., weather conditions, haze density, and scene characteristics), leading to inefficiency and potential operational risks in real-world deployment.

The main contribution of this paper is a generalized dehazing training strategy based on a balanced combination of synthetic and real-world hazy datasets. Instead of training multiple dehazing models specialized for specific haze conditions, we propose training a single model using a carefully designed mixture of synthetic and real-world data. The proposed strategy is validated using FFA-Net as the base model, which provides a relatively simple architecture while achieving strong dehazing performance.

Manuscript received May 13, 2026; Revised July 07, 2026; Accepted July 27, 2026. (ID No. JMIS-26M-05-023)

Corresponding Author (*): Jae-Ho Nah, +82-2-2287-6169, jaeho.nah@smu.ac.kr

¹Department of Computer Science, Sangmyung University, Seoul, Korea, 2003dolmen@naver.com, gywo0627@gmail.com, jaeho.nah@smu.ac.kr

Specifically, we train FFA-Net using a balanced 1:1 combination of the synthetic hazy dataset Haze4K [4] and the real-world hazy dataset NH-HAZE [5]. This training configuration effectively reduces the domain gap between synthetic and real-world hazy images, enabling the model to achieve stable and reliable dehazing performance under dynamically changing real-world environments.

This paper is organized as follows. Section 2 reviews related work, and Section 3 describes the characteristics of the datasets used in this study. Section 4 discusses the proposed data composition strategy, the experimental results, and their subsequent analysis. Finally, Section 5 concludes the paper.

II. RELATED WORK ON SINGLE-IMAGE DEHAZING

One of the most representative classical approaches for single image dehazing is the dark channel prior (DCP) method [1]. This method is based on the empirical observation that, in outdoor images free of sky regions, local patches typically contain pixels with very low intensity in at least one color channel. Using this statistical prior, DCP estimates the transmission map and the global atmospheric light. Driven by recent advancements in artificial intelligence, early deep learning studies, primarily focused on CNN-based approaches.

FFA-Net [2] is a representative CNN-based dehazing network designed to effectively capture deep haze-related features. The network consists of three feature-attention groups, each comprising 19 basic blocks. Each basic block incorporates local residual learning and a feature attention (FA) module, which enables the network to bypass thin haze or low-frequency information while focusing more on dense haze regions and high-frequency details. In contrast, Light-DehazeNet (LD-Net) [6] adopts a lightweight CNN architecture to jointly estimate the transmission map and atmospheric light, and it further introduces a color visibility restoration (CVR) module to alleviate color distortion. More recent work, DEA-Net [7], enhances conventional CNN-based approaches by integrating detail-enhanced convolutions for boundary and fine-detail restoration with content-guided attention mechanisms that account for both spatial and channel-wise importance.

Although these CNN-based models are generally effective at extracting local features, they exhibit inherent limitations in capturing global contextual information. This limitation becomes particularly pronounced in environments where haze is spatially non-homogeneous. To overcome these limitations, subsequent studies have explored alternative architectures, including Transformer-based and U-Net-based approaches, to enhance contextual understanding and multi-scale feature representation.

Transformer-based approaches offer a strong capability for modeling global dependencies; however, they can be inefficient for high-resolution image processing due to the quadratic computational complexity $O(n^2)$ of self-attention. To solve this issue, MB-TaylorFormer [8] significantly improves computational efficiency by linearizing self-attention using a Taylor approximation, reducing the complexity to $O(n)$. In addition, it employs a multi-branch architecture with deformable kernels to extract features from images of varying scales. Subsequently, DehazeFormer [9], built upon the Swin Transformer framework, enhances fine texture restoration by refining normalization layers and activation functions. Nevertheless, its performance tends to degrade in complex real-world outdoor environments, indicating limitations in robustness under diverse haze conditions.

Next, U-Net-based models exhibit a structural characteristic in which global information is extracted during the encoder stage and local details are restored in the decoder stage, making them well suited for image restoration tasks such as dehazing. Accordingly, various studies have explored the application of U-Net architectures to haze removal. MixDehazeNet [10] introduces a Mix Structure Block based on a U-Net backbone, in which self-attention is replaced by multi-scale parallel large-kernel convolutions (MSPLCK) and an enhanced parallel attention (EPA) module. In MSPLCK, large kernels are employed to capture global features, while small kernels are used to extract fine-grained local details. The EPA module is specifically designed to efficiently handle spatially non-homogeneous haze distributions. HazeFlow [11] further improves real-world generalization performance through a pretraining strategy based on the multi-component blending model (MCBM), which synthesizes non-homogeneous haze. Han et al. [12] proposed a lightweight two-stage U-Net-based network, in which the computationally expensive operations of the Dark Channel Prior are replaced by efficient convolutional operations. In this framework, the first stage estimates the transmission map, while the second stage performs haze removal.

Meanwhile, there also exist approaches that combine different paradigms or introduce novel architectural designs. ConvIR [13], for instance, aims to harness the strengths of both CNNs and Transformers by integrating local representations from convolutional layers with global contextual information from Transformer-based modules. To this end, it incorporates multi-shape attention and frequency restoration modules. Curricular Contrastive Regularization [14] performs contrastive regularization in a consensus contrastive space by categorizing negative samples into easy (input hazy images) and hard or ultra-hard samples (restored results from other models such as FFA-Net), and re-weighting them in a curriculum learning manner. Alternatively, USID-Net [15], an unsupervised learning–

based approach, learns disentangled representations by separating content and haze components for real-world image restoration, and it demonstrates effective dehazing performance without paired supervision through multi-scale feature attention. However, its complex network architecture and multiple loss functions lead to unstable training behavior.

Overall, these studies aim to balance strong local feature extraction with effective global representation learning. Among them, FFA-Net [2] stands out as a representative CNN-based model that maintains architectural simplicity and stable training characteristics while effectively integrating global contextual information through attention mechanisms. Therefore, FFA-Net is adopted as the baseline model to validate the proposed training methodology.

III. TRAINING DATASETS

In this study, the training and evaluation of FFA-Net are conducted using three representative datasets: RESIDE [16], NH-HAZE [5], and Haze4K [4] as summarized in Table 1 and Fig. 1. These datasets encompass synthetic,

Table 1. Quantitative comparison and specifications of the benchmark dehazing datasets.

Dataset	Type	Images	Description
ITS	Training	13,990	Indoor synthetic
OTS	Training	72,135	Outdoor synthetic
SOTS	Testing	500	Indoor/outdoor synthetic
NH-HAZE	Training/testing	55	Real outdoor
Haze4K	Training/testing	3,000/1,000	Indoor/outdoor synthetic

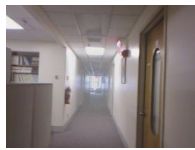
high-resolution, and real-world dehazing environments, thereby facilitating a comprehensive assessment of the model's generalization capability and practical applicability.

3.1. RESIDE Dataset

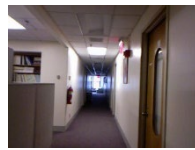
Realistic single image dehazing (RESIDE) [16] is a large-scale synthetic dehazing dataset consisting of both training and testing sets. By providing a systematic benchmark based on large-scale synthetic images, RESIDE ensures data diversity and experimental reproducibility. The dataset is composed of the indoor training set (ITS), outdoor training set (OTS), and the synthetic objective testing set (SOTS).

The ITS contains 13,990 synthetic hazy images, generated from 1,399 indoor clear images collected from the NYU-Depth V2 and Middlebury datasets. The OTS consists of 72,135 large-scale outdoor synthetic hazy images, which were constructed to overcome the limitations of indoor-centered synthetic datasets. The OTS is based on 2,061 outdoor clear images collected from the Internet. Since outdoor images lack ground-truth depth information, the deep depth estimation model proposed by Liu et al. [17] is employed to estimate scene depth for haze synthesis. By generating hazy images using estimated depth maps, this approach effectively improves the generalization capability of the trained dehazing models.

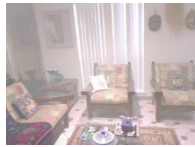
The testing set, synthetic objective testing set (SOTS), is composed of hazy–clear image pairs synthesized from 500 clear images selected from NYU-Depth V2. The same atmospheric scattering model used for the training sets is applied to generate these pairs. Owing to the availability of paired ground truth, full-reference evaluation metrics such as peak signal-to-noise ratio (PSNR) and structural simi-



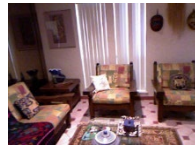
(a) Indoor training set (ITS)



(b) Outdoor training set (OTS)



(c) SOTS (indoor)



(d) SOTS (outdoor)



(e) NH-HAZE



(f) Haze4K

Fig. 1. Example image pairs from the dehazing datasets used in this study. In each pair, the left and right images represent the hazy and clear inputs, respectively.

arity index measure (SSIM) [18] can be employed for objective performance evaluation.

In this study, our goal is to build a stable and general-purpose dehazing model, and thus the large-scale synthetic RESIDE dataset, encompassing both indoor (ITS) and outdoor (OTS) scenes, is utilized for training. By jointly learning from the structural diversity of indoor scenes and the complex outdoor environments, the model can effectively capture features related to diverse atmospheric conditions, illumination variations, and depth changes. This comprehensive training strategy provides a crucial foundation for improving the model's adaptability to real-world environments.

3.2. NH-HAZE Dataset

NH-HAZE [5] is the first real-world dehazing dataset composed of paired clear and non-homogeneous hazy images captured under identical outdoor scene conditions. It consists of 55 outdoor scenes, each recorded as a hazy–clear image pair, and is specifically designed to represent realistic non-homogeneous haze. Previous dehazing studies have been constrained by the lack of datasets that adequately reflect the spatially varying haze distributions encountered in real-world environments. NH-HAZE was introduced to address this limitation.

The dataset comprises outdoor conditions collected over a period of approximately two months to realistically reproduce natural haze scenarios. Non-homogeneous haze is generated using an LSM1500 PRO haze generator, which disperses fine particles with sizes ranging from 1 to 10 μm throughout the scene. This setup enables the reproduction of realistic haze distributions with varying densities across different depths, ranging from near-field regions (1–2 m) to far-field regions (20–30 m). Consequently, NH-HAZE captures critical real-world characteristics that are absent in conventional homogeneous-haze datasets, including depth-dependent haze density variations, the spatial non-uniformity of haze layers, and illumination artifacts in outdoor scenes.

In this study, NH-HAZE is incorporated to construct a general-purpose dehazing network that is robust across diverse haze conditions. By complementing the synthetic RESIDE dataset—which cannot fully capture realistic haze scattering effects—NH-HAZE enables the model to learn from physically realistic haze distributions. Training FFA-Net on these non-homogeneous, real-world scenes significantly enhances its adaptability to diverse domains and depth variations, thereby improving its robustness in real-world dehazing scenarios.

3.3. Haze4K Dataset

Haze4K [4] is a large-scale synthetic hazy–clear image

dataset constructed to overcome the limitations of existing synthetic dehazing datasets and to enable more realistic performance evaluation. The dataset consists of 4,000 paired hazy–clear images, generated from 1,000 clean images, including 500 indoor images from the NYU-Depth V2 dataset and 500 outdoor images from the RESIDE-OTS dataset. From each clean image, four hazy images are synthesized to form the complete dataset. Specifically, the test set comprises 1,000 hazy images synthesized from 125 indoor and 125 outdoor clean images (250 clean images in total), while the training set consists of the remaining 3,000 hazy images generated from 750 clean images. In addition, Haze4K provides higher-resolution images compared to conventional synthetic datasets, allowing for more detailed structural restoration and high-quality dehazing training.

In this study, Haze4K is incorporated into the training process of FFA-Net to improve the model's generalization capability. Its large-scale composition, covering both indoor and outdoor environments, along with diverse variations in haze intensity, enables FFA-Net to learn richer and more diverse haze distribution patterns. Furthermore, the attention-based feature extraction architecture of FFA-Net effectively captures the complex haze characteristics present in Haze4K, leading to stable and high-quality dehazing results even in real-world scenarios.

IV. TRAINING STRATEGY AND EXPERIMENTAL RESULTS

In this section, we will first describe the results from the initial experiments that were trained solely by each dataset and analyze their limitations. Next, we will present our augmentation and joint training strategy that utilizes two different datasets together to overcome the generalization limitations.

The experimental environment comprised the following hardware and software setups. Two hardware environments were used for training and evaluating the FFA-Net model: (1) Intel i5-12400F with RTX 4060 Ti (16 GB) and (2) Xeon Gold 5220R with RTX A6000 (48 GB). All experiments were conducted based on the official FFA-Net implementation released on GitHub. For quantitative performance evaluation, the two most widely used metrics in the dehazing literature—PSNR and SSIM [18]—were adopted.

To ensure computational efficiency, all input images were randomly cropped to a resolution of 240×240 pixels. The initial learning rate was set to 0.0001, and the batch size was fixed at 2. The model was trained for a total of 500,000 steps, and performance was periodically evaluated to monitor training progress.

In the initial experiments, FFA-Net models were trained separately on the indoor training set (ITS) and the outdoor training set (OTS) of RESIDE and evaluated on the SOTS test set. These trained models exhibited characteristics nearly identical to the two pre-trained models provided with FFA-Net. As shown in Table 2, the quantitative results showed that

the model trained on ITS achieved satisfactory performance in indoor scenarios but suffered from significant performance degradation in outdoor environments. Conversely, the model trained on OTS performed well in outdoor scenes but demonstrated noticeably reduced restoration capability in indoor tests. The visual results in Figs. 2 and 3 also show

Table 2. Quantitative comparison of dehazing performance. Column headers and the first column denote the training and test datasets, respectively. The best and second-best results are highlighted in **bold** and underlined, respectively. In the table, “H4K:NH-H $n:1$ ” ($n=1, 5, 10, 20$) represents the mixing ratio of the datasets used for training the Haze4K and NH-HAZE models.

Test dataset	Metric	ITS	OTS	NH-HAZE	NH-HAZE (aug.)	Haze4K	H4K:NH-H 1:1	H4K:NH-H 5:1	H4K:NH-H 10:1	H4K:NH-H 20:1
SOTS indoor	PSNR	37.3917	18.9664	10.1916	19.3082	25.0418	<u>28.9253</u>	23.7989	27.2537	26.1518
	SSIM	0.9886	0.8564	0.6378	0.8328	0.9013	<u>0.9651</u>	0.9184	0.9526	0.9203
SOTS outdoor	PSNR	20.3100	32.1402	13.4879	14.8892	24.5510	26.8505	25.1574	<u>27.0491</u>	26.7615
	SSIM	0.8934	0.9802	0.7495	0.7828	0.9395	<u>0.9582</u>	0.9402	0.9558	0.9502
NH-HAZE	PSNR	10.3890	10.6145	16.5224	10.0211	12.9237	18.0261	13.7316	16.0509	<u>16.7354</u>
	SSIM	0.3793	0.4223	0.6595	0.5154	0.5017	0.6289	0.3864	0.6004	<u>0.6300</u>
Haze4K	PSNR	25.4130	16.8676	20.1907	22.3377	25.3711	28.7900	24.7848	<u>27.2806</u>	26.6296
	SSIM	<u>0.9625</u>	0.8717	0.9197	0.9428	0.9400	0.9664	0.9394	0.9621	0.9291

Scene 1: Outdoor scene (0001.png in SOTS Outdoor)

 Hazy	 ITS (20.2105 / 0.8971)	 OTS (31.9019 / 0.9855) [Best]
 NH-HAZE (14.7874 / 0.8484)	 NH-HAZE aug (11.9919 / 0.7209)	 Haze4K (28.3709 / 0.9726) [2nd]
 H4K:NH-H 1:1 (26.0100 / 0.9608)	 H4K:NH-H 5:1 (24.4703 / 0.9564)	 H4K:NH-H 10:1 (24.6310 / 0.9629)
 H4K:NH-H 20:1 (23.0161 / 0.9637)	 Clear (Ground Truth)	

Fig. 2. Qualitative comparison of dehazing results on Scene 1 (outdoor scene). Values in parentheses indicate the PSNR/SSIM.

Scene 2: Indoor scene (1403.png in SOTS Indoor)

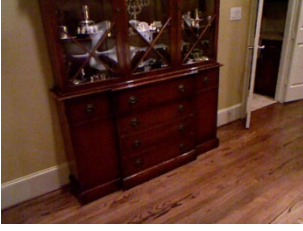
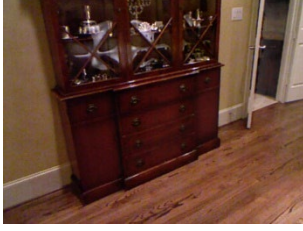
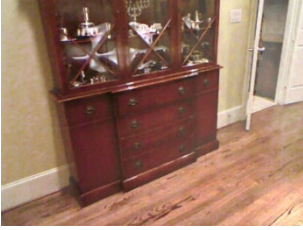
 Hazy	 ITS (35.0502 / 0.9902) [Best]	 OTS (23.4428 / 0.8746)
 NH-HAZE (9.0937 / 0.5326)	 NH-HAZE aug (7.7317 / 0.496)	 Haze4K (32.3078 / 0.982) [2nd]
 H4K:NH-H 1:1 (31.230 / 0.9783)	 H4K:NH-H 5:1 (25.458 / 0.9177)	 H4K:NH-H 10:1 (29.822 / 0.974)
 H4K:NH-H 20:1 (28.1734 / 0.9242)	 Clear (Ground Truth)	

Fig. 3. Qualitative comparison of dehazing results on Scene 2 (indoor scene). Values in parentheses indicate the PSNR/SSIM.

trends consistent with the findings in the table. These results indicate that domain gaps exist even among synthetic datasets, and models trained on a single dataset have limited generalization ability across different environments.

To verify FFA-Net’s performance on other datasets, which are not included in the FFA-Net paper [2], we additionally trained the model on NH-HAZE and Haze4K and measured the trained models’ performance across four datasets. NH-HAZE contains real-world, non-homogeneous haze characteristics, which are very different from the ITS, OTS, and SOTS. As a result, the ITS- and OTS-trained models produced very low-quality results (PSNR around 10 dB) in NH-HAZE. As shown in Fig. 4, they failed to effectively reduce non-homogeneous haze. In contrast, the NH-HAZE-trained model showed the best and worst results in NH-HAZE and SOTS, respectively. As shown in Figs. 2 and 3, this model unevenly recovered the homogeneous haze, and as a result, some regions are too dark or bright. These results

also confirm the presence of a severe domain gap and that models trained on a single domain have limited generalization ability.

The above degradation could be attributed to the limited size of the NH-HAZE dataset, which consists of only 55 image pairs, leading to training instability when used alone. To mitigate this issue, data augmentation techniques—including rotation, horizontal and vertical flipping, and scaling—were applied to substantially increase the training data volume. As shown in Table 2, the increased sample size enabled the model to improve the SOTS dehazing quality. However, this improvement came at the cost of reduced restoration quality on the NH-HAZE dataset. This degradation occurred because geometric and scale transformations—such as vertical flipping and scaling—distort the physical flow patterns and inherent spatial priors of the real-world, non-homogeneous haze in the NH-HAZE dataset. Consequently, these transformations introduce physically implausible artifacts (e.g., ground-level

Scene 3: Non-homogeneous scene (55.png in NH-HAZE)

		
Hazy	ITS (10.1408 / 0.2968)	OTS (13.1950 / 0.4178)
		
NH-HAZE (17.5515 / 0.610) [2nd]	NH-HAZE aug (16.5616 / 0.5917)	Haze4K (12.6135 / 0.4754)
		
H4K:NH-H 1:1 (17.0000 / 0.6040)	H4K:NH-H 5:1 (16.5937 / 0.6207)	H4K:NH-H 10:1 (18.0375 / 0.6387) [Best]
		
H4K:NH-H 20:1 (16.1677 / 0.5933)	Clear (Ground Truth)	

Fig. 4. Qualitative comparison of dehazing results on Scene 3 (non-homogeneous scene). Values in parentheses indicate the PSNR/SSIM.

haze appearing in the sky), thereby contradicting the physical setup of real scenes and disrupting contextual continuity [19].

Finally, when trained on Haze4K, the model demonstrated moderate and balanced performance across the SOTS and NH-HAZE datasets. Specifically, while its metrics fall within an intermediate range on SOTS Indoor and Outdoor, it relatively outperformed the ITS- and OTS-trained models on the NH-HAZE dataset. However, a qualitative limitation was observed in Scene 3 (Fig. 4), where the model failed to completely remove the dense white haze in the center, merely shifting its hue to a bluish tint. This insufficient restoration indicated that while the model achieves competitive baseline numbers, it still faces challenges in fully capturing complex non-homogeneous haze patterns.

These findings led to a broader research question: Can the characteristics of real-world haze and the diversity of large-scale synthetic data complement each other? We

hypothesized that combining NH-HAZE (for its realistic, non-homogeneous priors) with Haze4K (for its diverse indoor/outdoor scenes) could alleviate single-dataset domain biases. Crucially, since NH-HAZE suffers from a limited sample size, data augmentation remains indispensable during this integration. This combination ultimately enables FFA-Net to capture more physically plausible features while broadening its robustness against diverse haze patterns.

Accordingly, this study investigated several Haze4K-to-NH-HAZE mixing ratios—specifically, 1:1, 5:1, 10:1, and 20:1—to examine the influence of dataset composition on dehazing performance. As presented in Table 2, the 1:1 setup generally yielded the best results, followed closely by the 10:1 setup. Although the absolute PSNR and SSIM values vary, consistent trends were observed across the different datasets for all combinations. As shown in Figs. 2 and 3, models utilizing our composition strategy achieved

comparable or slightly lower quality compared to the Haze4K-only baseline, while delivering much higher quality than the NH-HAZE-only baseline. Conversely, as illustrated in Fig. 4, our models demonstrated competitive restoration quality compared to the NH-HAZE-only baseline, both with and without data augmentation. A similar trend was evident in the quantitative results presented in Table 2. This indicated their clear superiority over models trained solely on the ITS, OTS, or Haze4K datasets.

Naturally, subtle quality trade-offs existed between the configurations. In Scene 1, the 10:1 setup outperformed the 1:1 setup, as the latter introduced noticeable artifacts (stains) in the sky region. In Scene 2, however, the 1:1 setup was superior due to the absence of visible artifact patterns on the wallpaper. In Scene 3, the 10:1 setup proved most effective in terms of thorough haze removal. Nevertheless, the qualitative gaps among our mixed-dataset models were remarkably minor compared to the severe degradation or overfitting exhibited by models trained solely on a single dataset. Therefore, we recommend using a mixing ratio of 1:1, while the 10:1 ratio also serves as a highly commendable alternative.

V. CONCLUSION

This study investigated the training and performance of the FFA-Net architecture under various combinations of synthetic haze datasets (ITS, OTS), real-world non-homogeneous haze datasets (NH-HAZE), and Haze4K, which reflect realistic image capture conditions. Experimental results showed that a data composition strategy employing a Haze4K-to-NH-HAZE ratio of 1:1 achieved robust and reliable performance in most cases. These results empirically confirm that incorporating ground-truth-based real-world data during training substantially improves generalization performance in real hazy environments.

Rather than relying on a single dataset, this work proposes a balanced data composition strategy that integrates synthetic and real-world datasets, as well as large-scale and realistic-environment data. We believe this approach provides a practical direction for training dehazing models capable of handling diverse haze conditions. Although the experiments are limited to the FFA-Net architecture, the proposed training data configuration serves as a transferable guideline for other deep learning-based image dehazing models. Importantly, the resulting model demonstrates reduced overfitting to synthetic haze patterns, highlighting its strong potential for real-time video processing applications such as live streaming, autonomous driving assistance systems, and drone-based visual enhancement.

ACKNOWLEDGEMENT

We would like to thank the authors of FFA-Net for making their source code publicly available, and the anonymous reviewers for their constructive feedback, which helped improve this manuscript. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean Government [Ministry of Science and ICT (MSIT)] under Grant RS-2025-00521436.

SUPPLEMENTARY MATERIALS

We attach the source code as supplementary material. It includes scripts for training and testing FFA-Net, along with a GUI application for comparing image results.

REFERENCES

- [1] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341-2353, 2011.
- [2] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia, "FFA-Net: Feature fusion attention network for single image dehazing," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, 2020, pp. 11908-11915.
- [3] C. Zheng, W. Ying, and Q. Hu, "Comparative analysis of dehazing algorithms on real-world hazy images," *Scientific Reports*, vol. 15, p. 10822, 2025.
- [4] Y. Liu, L. Zhu, S. Pei, H. Fu, J. Qin, and Q. Zhang, et al., "From synthetic to real: Image dehazing collaborating with unlabeled real data," in *Proceedings of the ACM International Conference on Multimedia*, 2021, pp. 50-58.
- [5] C. O. Ancuti, C. Ancuti, and R. Timofte, "NH-HAZE: An image dehazing benchmark with non-homogeneous hazy and haze-free images," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 1798-1805.
- [6] H. Ullah, K. Muhammad, M. Irfan, S. Anwar, M. Sajjad, and A. S. Imran, et al., "Light-DehazeNet: A novel light-weight CNN architecture for single image dehazing," *IEEE Transactions on Image Processing*, vol. 30, pp. 8968-8982, 2021.
- [7] Z. Chen, Z. He, and Z. M. Lu, "DEA-Net: Single image dehazing based on detail-enhanced convolution and content-guided attention," *IEEE Transactions on Image Processing*, vol. 33, pp. 1002-1015, 2024.
- [8] Y. Qiu, K. Zhang, C. Wang, W. Luo, H. Li, and Z. Jin, "MB-TaylorFormer: Multi-branch efficient transformer expanded by Taylor formula for image dehazing," in

Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 12802-12813.

- [9] Y. Song, Z. He, H. Qian, and X. Du, "Vision transformers for single image dehazing," *IEEE Transactions on Image Processing*, vol. 32, pp. 1927-1941, 2023.
- [10] L. Lu, Q. Xiong, B. Xu, and D. Chu, "MixDehazeNet: Mix structure block for image dehazing network," in *Proceedings of the International Joint Conference on Neural Networks*, 2024, pp. 1-10.
- [11] J. Shin, S. Chung, Y. Yang, and T. H. Kim, "HazeFlow: Revisit haze physical model as ODE and non-homogeneous haze generation for real-world dehazing," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2025, pp. 6263-6272.
- [12] Y. Han, J. Kim, J. Lee, J. H. Nah, Y. S. Ho, and W. C. Park, "Efficient haze removal from a single image using a DCP-based lightweight U-Net neural network model," *Sensors*, vol. 24, p. 3746, 2024.
- [13] Y. Cui, W. Ren, X. Cao, and A. Knoll, "Revitalizing convolutional network for image restoration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, pp. 9423-9438, 2024.
- [14] Y. Zheng, J. Zhan, S. He, J. Dong, and Y. Du, "Curricular contrastive regularization for physics-aware single image dehazing," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 5785-5794.
- [15] J. Li, Y. Li, L. Zhuo, L. Kuang, and T. Yu, "USID-Net: Unsupervised single image dehazing network via disentangled representations," *IEEE Transactions on Multimedia*, vol. 25, pp. 3587-3601, 2023.
- [16] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, and W. Zeng, et al., "Benchmarking single-image dehazing and beyond," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 492-505, 2019.
- [17] F. Liu, C. Shen, G. Lin, and I. Reid, "Learning depth from single monocular images using deep convolutional neural fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 10, pp. 2024-2039, 2016.
- [18] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.
- [19] P. Shyam and H. Yoo, "Data efficient single image dehazing via adversarial auto-augmentation and extended atmospheric scattering model," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2023, pp. 227-237.

AUTHORS



Chae-bin Park received her B.S. degree in Computer Science from Sangmyung University, Seoul, Korea, in 2026. Her research interests include image processing, computer vision, and machine learning.



Hyo-jae Lee received his B.S. degree in Computer Science from Sangmyung University, Seoul, Korea, in 2026, where he focused on image processing and computer vision. He is currently interested in extending this work into machine learning applications.



Jae-Ho Nah received the B.S., M.S., and Ph.D. degrees in computer science from Yonsei University, Seoul, South Korea, in 2005, 2007, and 2012, respectively. He is currently an Assistant Professor at Sangmyung University, Seoul. Prior to joining academia, he was with LG Electronics. He has co-authored more than 40 technical papers published in international journals and conferences, including ACM SIGGRAPH, ACM SIGGRAPH Asia, ACM Transactions on Graphics, IEEE Transactions on Visualization and Computer Graphics, Computer Graphics Forum, High-Performance Graphics, and Computers & Graphics. His research interests include ray tracing, rendering algorithms, and graphics hardware.

