Mesopotamian Journal of Computer Science

Volume 5 | Issue 1 Article 12

1-1-2025

HAZE-IMAGE-DATASET: A Large-Scale Benchmark for Image Dehazing in Variable Fog and Low-Light Conditions

Mustafa J. Shahbaz
Department of Physics, Mustansiriyah University, Baghdad, Iraq, jsblue2007@uomustansiriyah.edu.iq

Ali A.D. Al-Zuky
Department of Physics, Mustansiriyah University, Baghdad, Iraq

How to Cite This Article

Shahbaz, Mustafa J. and Al-Zuky, Ali A.D. (2025) "HAZE-IMAGE-DATASET: A Large-Scale Benchmark for Image Dehazing in Variable Fog and Low-Light Conditions," *Mesopotamian Journal of Computer Science*: Vol. 5: Iss. 1, Article 12.
DOI: https://doi.org/10.58496/MJCSC/2025/012
Available at: https://map.researchcommons.org/mjcsc/vol5/iss1/12

Mesopotamian journal of Computer Science
Vol. **(2025)**, 2025, **pp**. 202–212
DOI: https://doi.org/10.58496/MJCSC/2025/012; ISSN: 2958-6631
https://mesopotamian.press/journals/index.php/cs

Research Article

HAZE-IMAGE-DATASET: A Large-Scale Benchmark for Image Dehazing in Variable Fog and Low-Light Conditions

Mustafa J. Shahbaz[1,*], Ali A. D. Al-Zuky[1]

[1]*Department of Physics, Mustansiriyah University, Baghdad, Iraq*

ARTICLEINFO

Article History

Received 18 Jul 2025
Revised 25 Aug 2025
Accepted 28 Aug 2025
Published 31 Aug 2025

Keywords

Hazy Image Dataset
Deep Learning
CNN Classification
Fog simulation
Haze Dataset

Abstract

To enhance image dehazing and visual recognition in real-world conditions, we introduce HAZE-IMAGE-DATASET, a large-scale dataset comprising nearly 42,000 images. It is constructed from 1,532 clean images sourced globally and captured using a Samsung smartphone, covering diverse natural and urban scenes. The dataset includes synthetic and real haze variations. Synthetic haze was generated using MATLAB-based atmospheric scattering models with depth maps for 10 fog levels. Colored haze was created using alpha blending ($\alpha = 0.4$) in six colors: red, green, blue, yellow, white, and black. Low-light conditions were simulated via uniform darkening at 10 levels. Also, 616 real haze images were captured using a steam device to replicate genuine haze characteristics. Benchmarking was performed using MobileNetV2 and GoogLeNet. On a hazy subset, MobileNetV2 improved PSNR from 13.34 dB to 21.77 dB and SSIM from 0.805 to 0.950, Also, GoogleNet achieved a PSNR of 22.38 dB and SSIM of 0.953.Classification experiments using a seven-class subset showed high accuracy: 98.93% with AlexNet and 98.13% with GoogLeNet.The dataset is organized into haze types, levels, and color folders. It serves as a strong foundation for training and evaluating image dehazing and recognition models under different visibility conditions. It is available on GitHub: https://github.com/mustafa1jamal/HAZE-IMAGE-DATASET.

1. INTRODUCTION

In recent years, image dehazing has gained notable attention in the computer vision and remote sensing because its essential role in enhancing visual quality for applications such as autonomous driving, video surveillance, and aerial imaging [1]-[2]. Hazy conditions caused by atmospheric particles scatter light and reduced visibility, making it difficult for machine learning models and human observers to detect objects clearly [3]-[4]. Traditional dehazing methods depend on previous models like the dark channel prior or atmospheric scattering models [5], but these often fail in complex scenes with uneven lighting or colored haze [6]-[7]. The development of deep learning, especially convolutional neural networks (CNNs), has improved dehazing performance through end-to-end learning pipelines that capture nonlinear haze distributions [8]-[9]. Even with these developments, one of the key challenges in dehazing research remains the limited availability of diverse and realistic datasets. Most existing datasets, like RESIDE, focus on synthetic haze generation and lack real-world diversity, especially in colored fog and low-light conditions [10]-[11]. Recently, new datasets like LMHaze [12], HazeSpace2M [13], and RealIndoorFog [14] have attempted to bridge this gap, but many still fail to include ground truth references, multi-level haze density, or illumination variations that occur in practical environments [15]-[16]. Furthermore, domain shifts between synthetic and real-world data negatively affect the generalization of deep models [17]-[18].Also, current dehazing datasets largely focus on single-colored haze under standard illumination, which limits model generalization to real-world scenarios where haze may exhibit varying colors (e.g., sunset orange, industrial yellow) or occur under low-light conditions such as nighttime fog or shaded regions. By excluding these challenging variants, existing datasets fail to train models that are strong across a broad spectrum of environmental conditions [19]-[20].

To address these limitations this work proposes HAZE-IMAGE-DATASET, a detailed and high-resolution dataset consisting of 42,000 hazy images derived from 1,500 clean images. These images are sourced from both real-world captures and public datasets to ensure diversity. The dataset contains four key components: (1) 10 levels of physical haze, generated via a MATLAB-based exponential scattering model with synthetic depth maps [21], (2) 6 variants of colored haze using alpha blending in red, green, blue, yellow, white, and black; (3) 10 levels of uniform darkness simulating low-light conditions; and (4) a real fog subset, created using a household steam iron to capture actual haze effects indoors [22]-[23].

*Corresponding author. Email: jsblue2007@uomustansiriyah.edu.iq

All images are precisely organized into structured directories, allowing effective use in both supervised and unsupervised learning contexts. the innovation of this dataset lies not only in its size but in the inclusion of real haze, colored fog, and controlled darkness levels conditions underrepresented in current benchmarks.In contrast to existing datasets, In particular, LMHaze provides real haze samples but does not include low-light scenarios or colored haze variants, limiting its applicability to complex lighting conditions. HazeSpace2M, while extremely large in scale, is composed entirely of synthetic images and lacks real-world fog or illumination diversity, making it less representative of practical deployment environments. In the same way, RealIndoorFog focuses exclusively on indoor scenes, and therefore omits outdoor environments and any simulation of colored haze**,** which are crucial for general-purpose vision systems. By addressing these limitations, our dataset enables more comprehensive training and evaluation of dehazing and recognition models under varied and realistic visibility impairments [24]-[25].

The dataset is publicly available at: https://github.com/mustafa1jamal/HAZE-IMAGE-DATASET [26].

2. DATA DESCRIPTION

The HAZE-IMAGE-DATASET contains a total of 42,000 high-resolution images, designed to provide a detailed resource for evaluating image dehazing algorithms in practical and varied environments. The dataset originates from a base of 1,532 clean images, collected from three primary sources to ensure wide scene and environmental variation. Particularly, 540 images were curated from global public datasets and image repositories on the internet, selected to include a variety of indoor and outdoor scenes. To enhance geographical diversity, 960 clean images were personally captured using a Samsung A52 smartphone during travels across eight different countries, encompassing a variety of urban, rural, and natural environments. Also, 616 images were generated using real physical fog, produced indoors with a household steam generator, providing natural haze conditions including light scattering and moisture effects not easily to make through synthetic techniques. From the base of 1,500 clean images, three augmentation techniques were applied to create a total of 42,000 hazy images. First, a MATLAB-based physical haze simulation was applied using the atmospheric scattering model with synthetic vertical depth maps to produce10 levels of haze intensity. Each image was processed with β values ranging from 0.3 to 2.0, resulting in 15,320 foggy images. Second, a colored haze augmentation method was applied, combining each image with six different haze colors (red, green, blue, yellow, white and black) using a fixed alpha of 0.4, producing 9,000 images. Third, uniform darkness was simulated at 10 different levels ranging from 10% to 80% shading to simulate low-light conditions, producing another 15,320 dark images. These augmentations, along with the original clean and real-fog images, produced a structured and well-balanced dataset comprising 41,980 images. The dataset is organized into multiple directories to reflect various environmental obscuration conditions.

TABLE I. SUMMARY OF THE HAZY-IMAGE-DATASET COMPOSITION, INCLUDING THE NUMBER OF IMAGES AND A BRIEF DESCRIPTION FOR EACH SUBSET

Dataset Subset	Number of Images	Description
Clean Images	1,532	Original clear images, originate from both self-collected images and publicly available on global web.
Physical Fog (10 Levels)	15,320 (10 × 1,532)	Synthetic fog using physical model across ten levels (β = 0.3–2.0).
Colored Haze (6 Colors)	9,192 (6 × 1,532)	Red, Green, Blue, Yellow, White, Black (α = 0.4).
Uniform Darkness (10 Levels)	15,320 (10 × 1,532)	Simulated low-light scenes from 10% to 80% darkness.
Real-World Haze Images	616	Real foggy images captured using physical steam to emulate natural haze in controlled indoor environment.
Image Resolution Range		640×360 to 3024×4032 pixels
Total Images	41,980	

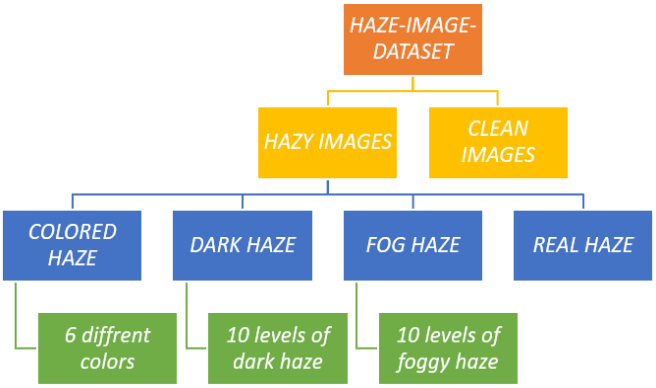

Fig. 1. Folder structure of the HAZE-IMAGE-DATASET as presented on GitHub, showing the different augmentation types and subcategories.

3. COMPARISON WITH EXISTING DATASETS

Compared to existing haze image datasets, the proposed HAZE-IMAGE-DATASET provides larger scale and greater diversity. Dataset diversity improves model robustness by showing it to various haze types, lighting, and scene contexts. This variability enables better generalization and cross-domain adaptation in real-world applications.Although RESIDE [27]-[28] provides a good balance between synthetic and real hazy images (~13k images), it lacks colored haze and dark visibility variations. O-HAZE [29] contains only 45 real-world hazy images captured with professional equipment, limiting its real-world expansion. D-Hazy [30]-[31] focuses on synthetic haze using depth maps but does not support multiple haze types. In contrast, our dataset provides almost 42,000 images including clean, synthetic fog (10 levels), colored haze (6 variations), uniform darkness (10 levels), and 616 real-world hazy images captured using a household steam generator. This makes the dataset suitable for robust training and evaluation of deep learning models under a wide range of visibility degradation conditions.

TABLE II. COMPARISON OF EXISTING IMAGE DEHAZING DATASETS. THE PROPOSED HAZE-IMAGE-DATASET PROVIDES BROADER HAZE DIVERSITY AND A LARGER IMAGE COUNT FOR ROBUST DEEP LEARNING EVALUATION.

Dataset Name	Year	Total Images	Real/Synthetic	Haze Types	Image Resolution	Availability	Colored haze	Low light support	Paired Clean/Hazy Images
RESIDE	2018	~13,000	Synthetic + Real	Normal Haze Only	Up to 620×460	Public	No	No	Yas
O-HAZE	2018	45	Real	Outdoor Haze	544×363	Public	No	No	Yas
D-Hazy	2016	~1,449	Synthetic (Depth-based)	Normal Haze	Various	Public	No	No	Yas
HAZE-IMAGE-DATASET	2025	41,980	Both	Fog, Dark, Colored, Real	360x640—3024x4032	Public (GitHub)	Yaa	Yas	Yas

4. DATA GENERATION METHOD

To ensure a comprehensive dataset that simulates diverse hazy environments, the HAZE-IMAGE-DATASET was generated using four distinct approaches: synthetic physical fog, uniform darkness, colored haze, and real-world haze.

4.1 Physical Fog Generation (Fog Levels 1–10)

Synthetic fog was conducted using a physics-based scattering model applied in MATLAB. Each input image was processed with a simulated depth map ranging from top to bottom, and the transmission map was calculated using the equation:

$$I(x, y) = J(x, y) \cdot t(x, y) + A \cdot (1 - t(x, y))$$
$$t(x, y) = e^{\wedge}(-\beta \cdot d(x, y))$$

Where:
$I(x, y)$: The observed hazy image.
$J(x, y)$: The clean image.
$t(x, y)$: The transmission map relying on depth and haze density.
β: The atmospheric scattering coefficient which controls the haze level.
$d(x, y)$: The synthetic scene depth at pixel (x, y).
A: The global atmospheric light (usually white = 1).

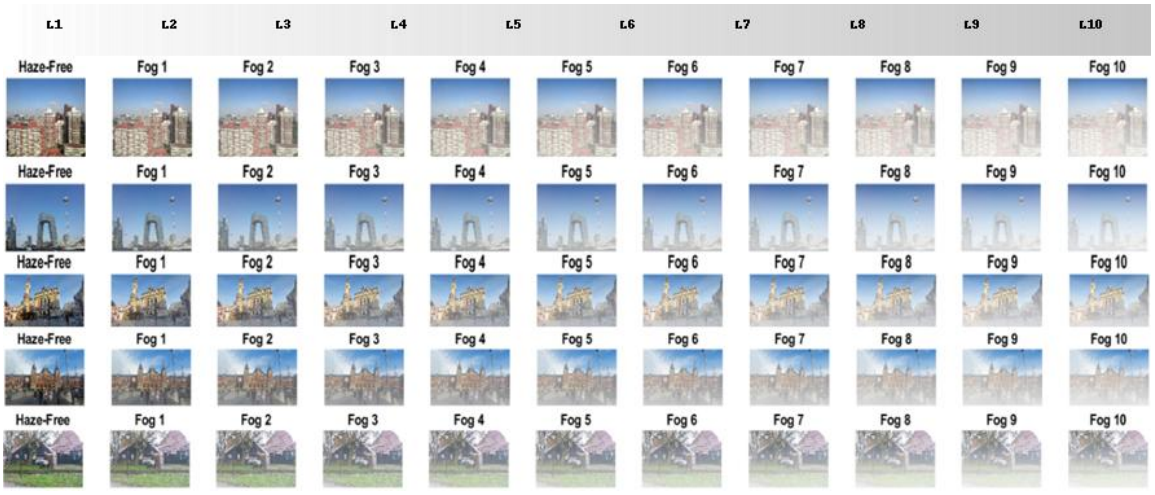

Fig. 2. Illustration of synthetic haze generation using the atmospheric scattering model in MATLAB. Each row show a clean image followed by ten fog levels, showing increasing haze density applied to the clean image set.

4.2 Uniform Darkness Generation (Dark Levels 1–10)

To simulate low-light and night-like conditions, where haze-like effects reduce visibility due to soot**,** industrial smoke, or heavy urban pollution, particularly in poorly lit environments or tunnel-like scenes. This variant emphasizes low-luminance and high-opacity distortion, presenting a unique challenge for dehazing models. uniform darkness was applied across images using a linear pixel intensity reduction technique. The darkening model used was:

$$I\_dark = I \cdot (1 - \alpha)$$

Where:

I_dark: The darkened image.

I: The original clean image.

α: Darkness factor ranging from 0.1 to 0.8.

Fig. 3. Dark haze generation using uniform darkness levels (from Dark 1 to Dark 10), simulating low-light conditions at different intensities applied to various natural scenes.

4.3 Colored Haze Generation

To mimic various environmental haze colors, a fixed blending operation was performed using different haze colors (red, green, blue, black, yellow, and white). The MATLAB-based haze simulation applied **alpha blending** using the following equation:

$$I\_colored = I\cdot(1 - \alpha) + A\cdot\alpha$$

Where:
I_colored: The output image with colored haze.
I: The original image.
α: Haze opacity factor (set to 0.4).
A: The haze color vector (e.g., [1 0 0] for red haze).

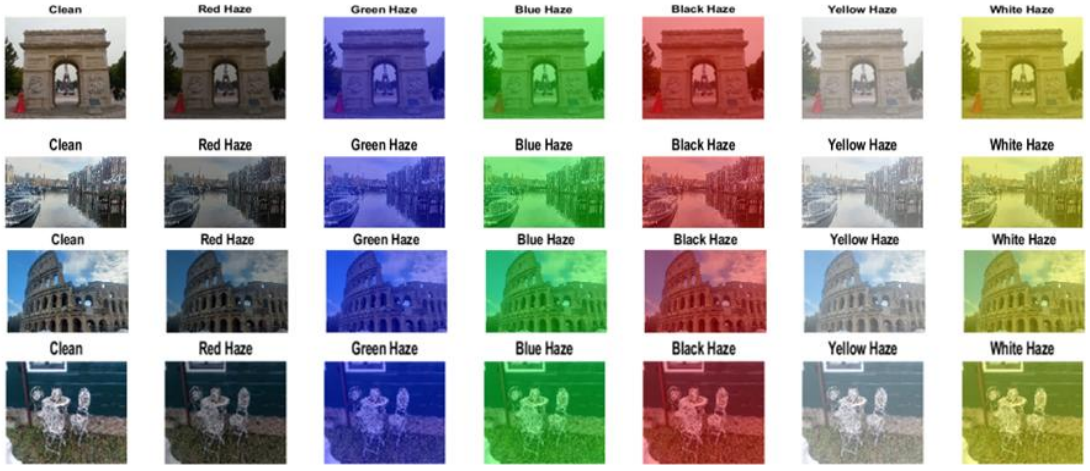

Fig. 4. Examples of synthetic colored haze augmentation applied to clean images using six different color overlays (Re d, Green, Blue, Yellow, White, Black) using α-blending (α = 0.4).

4.4 Real Haze Generation Using Steam Device

To simulate realistic environmental conditions, hundreds of images were captured with actual physical fog generated using a household steam iron device. This setup was used to simulate real haze indoors, adding naturally scattered fog particles to the image scene. These images are grouped under the real haze/ directory and provide real-world fog effects beyond digital approximation.

Fig. 5. Real haze generation taken with a household steam iron, showing real-world fog applied to indoor scenes. A total of 616 images were created across ten fog levels for increased realism beyond synthetic evaluations.

5. Baseline Benchmarking and Sample Application

To prove the usability and performance of the proposed dataset, a baseline evaluation was conducted using two widely known convolutional neural network (CNN) models: GoogleNet and MobileNetV2. These models were chosen for their balanced trade-off between accuracy and computational efficiency, particularly in environments with limited resources. MobileNetV2 was chosen for its lightweight architecture and suitability for real-time deployment. GoogleNet, with its Inception modules, provides strong feature extraction across multiple scales with moderate computational cost. A subset of the dataset, especially containing real-world hazy images generated from both synthetic and physical haze conditions, was used for testing. A representative subset of the dataset was selected for benchmarking to balance evaluation diversity with computational feasibility. This subset includes samples across all haze levels, colors, and lighting conditions, ensuring relevance to the study objectives whereas reducing the heavy processing load associated with full-scale benchmarking.Each model was applied using MATLAB 2023a on a Windows 11 machine with 16GB RAM and an NVIDIA GPU. The images were resized properly for each network input: 224×224 for GoogleNet and 224×224 for MobileNetV2. The primary objective was to show that models trained on standard clean images show noticeable degradation when tested on hazy data, confirming the need for enhanced and diversified training datasets such as ours. The processed outputs were calculated visually and quantitatively using common metrics (PSNR, SSIM, MSE).The following parameters were used for the dehazing preprocessing stage applied before classification:

- $t_0 = 0.1$ Minimum transmission threshold to prevent division by zero.
- ω ***(omega)*** = 0.5 Controls the strength of the haze effect in the dark channel prior.
- ***Patch Size*** = 15 × 15 Used for the local minimum filter in dark channel estimation.
- ***rt*** = 0.5 Blending ratio between dark channel and CNN features.

Results showed that even lightweight models like MobileNetV2 were able to benefit from the enhanced clarity after dehazing. The application of these models acts as a practical demonstration of how the HAZE-IMAGE-DATASET can support image classification tasks under adverse visibility conditions.

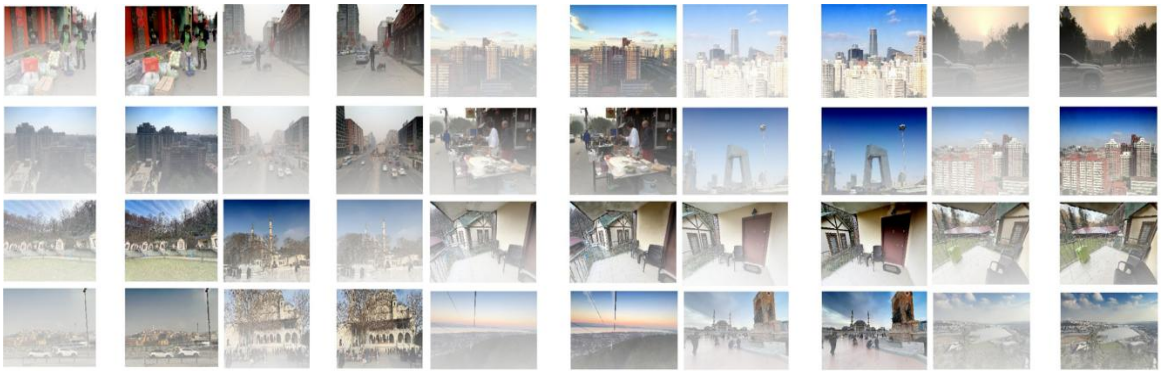

Fig. 6. Classification results using the MobileNetV2 model on clean, hazy, and dehazed images from the HAZE-IMAGE-DATASET. The lightweight architecture of MobileNetV2 demonstrates effective performance improvement after haze removal, making it suitable for real-time deployment in visibility-degraded environments.

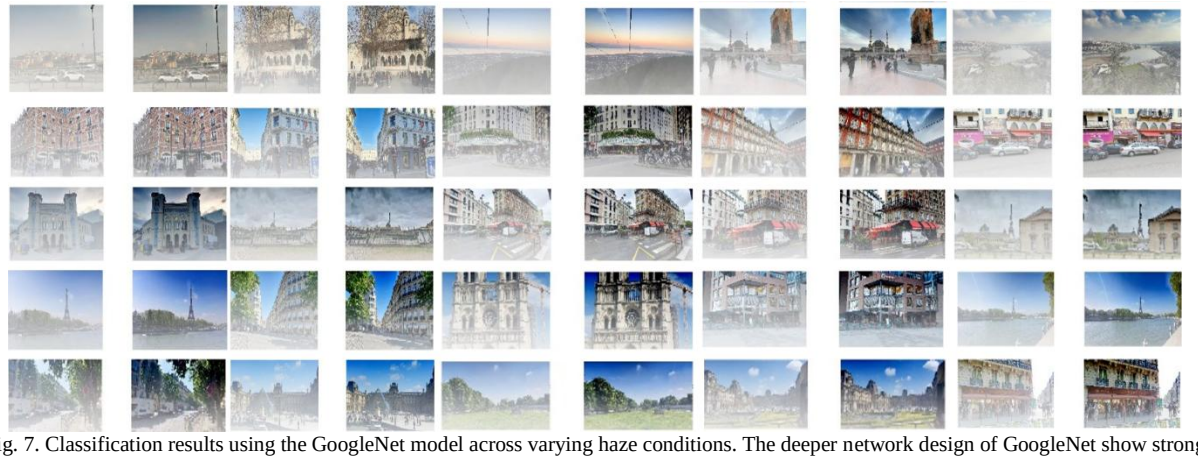

Fig. 7. Classification results using the GoogleNet model across varying haze conditions. The deeper network design of GoogleNet show strong generalization and high classification accuracy after dehazing, proving the dataset's capability to support robust vision models in foggy or dark scenes.

TABLE III. EVALUATION METRICS FOR HAZY AND DEHAZED IMAGES USING THE MOBILENETV2 APPROACH. THE RESULTS SHOW IMPROVEMENT IN PSNR, SSIM, AND BLUR METRICS AFTER DEHAZING, WITH AN AVERAGE PROCESSING TIME OF 0.554 SECONDS PER IMAGE

Metric	Mean Value	Standard Deviation (Std)	Improvement %
PSNR_Hazy	13.34	1.27	_
PSNR_Dehazed	21.77	3.64	63.2 %
SSIM_Hazy	0.805	0.052	_
SSIM_Dehazed	0.950	0.034	18 %
Blur_Hazy	0.00766	0.00444	_
Blur_Dehazed	0.01656	0.01027	_
Time_Per_Image (sec)	0.554 sec	0.273 sec	_

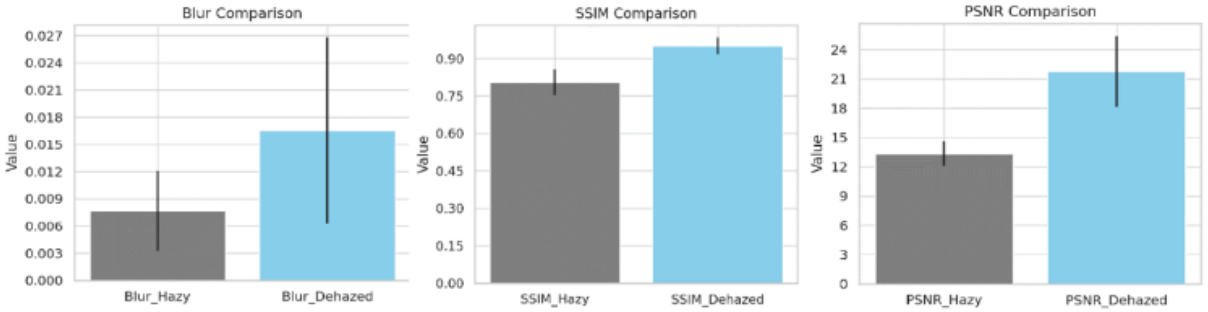

Fig.8. Bar charts illustrating PSNR, SSIM, and Blur metrics before and after dehazing based on the **MobileNetV2** evaluation results.

TABLE IV. EVALUATION METRICS FOR HAZY AND DEHAZED IMAGES USING THE GOOGLENET MODEL. THE IMPROVED METRICS SHOW BETTER VISUAL QUALITY AND FASTER PROCESSING TIME (0.382 SECONDS PER IMAGE) COMPARED TO THE FIRST APPROACH.

Metric	Mean	Standard Deviation	Improvement %
PSNR_Hazy	13.34	1.27	_
PSNR_Dehazed	22.38	3.51	67.8%
SSIM_Hazy	0.805	0.052	_
SSIM_Dehazed	0.953	0.036	18.4%
Blur_Hazy	0.00766	0.00444	_
Blur_Dehazed	0.01752	0.01110	_
Time_Per_Image	0.382 sec	0.054 sec	_

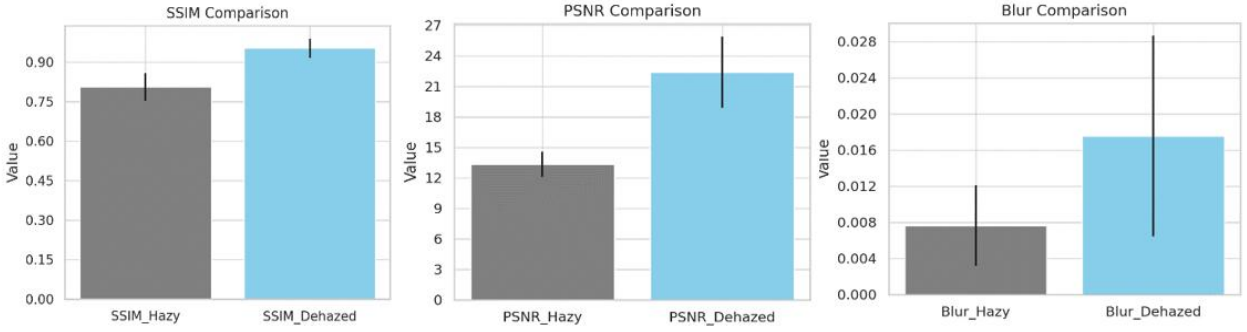

Fig.9. Bar charts showing PSNR, SSIM, and Blur performance from the **GoogleNet** evaluation with enhanced dehazing results.

To further validate the classification capabilities enabled by the HAZE-IMAGE-DATASET, we evaluated two pretrained CNN models AlexNet and GoogleNet, without additional fine-tuning. These models were used as to assess performance on the 7-class colored haze subset. The models were trained and tested using MATLAB with standardized input resizing (227×227 for AlexNet, 224×224 for GoogleNet). Confusion matrices showed that both models scored high classification accuracy: 98.93% for AlexNet and 98.13% for GoogleNet, respectively. AlexNet showed stronger stability to class separation, with excellent recognition in five of seven categories. Minor confusion happened between original and black haze images, validating the visual similarity under certain conditions. These results show that colored haze conditions can be successfully classified using CNNs and also highlight the dataset's suitability for haze-aware recognition systems.

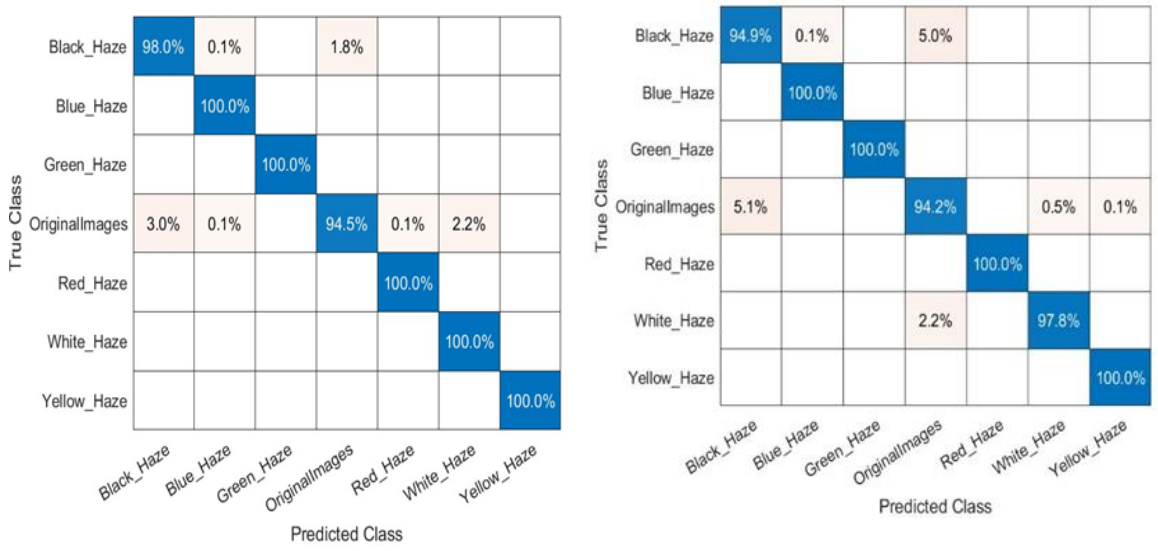

Fig. 10. Confusion matrices showing classification accuracy for the seven image classes (original and six colored haze types) using **AlexNet** (left) and **GoogleNet** (right). The diagonal dominance illustrates strong model performance, with AlexNet achieving higher accuracy on black and original haze classes, while GoogleNet show slight misclassification in darker haze conditions.

6. CONCLUSION

The HAZE-IMAGE-DATASET, a large-scale and varied dataset were presented particularly developed for training and evaluating image enhancement and object detection models in hazy and low-visibility conditions. The dataset includes 42,000 images, including clean images, synthetically generated fog at 10 different intensity levels using a physical haze model, uniform dark images with varying levels of darkness, and colored haze images across six environmental tones. Also, real-world hazy images were captured using a household steam generator to simulate actual atmospheric haze, enhancing the dataset's real-world accuracy. In order to generate the synthetic haze, a well-established physical equations implemented in MATLAB scripts were used, ensuring controlled and consistent environmental degradation across different subsets. Neural network models, GoogleNet and MobileNetV2, using dehazing preprocessing based on physical models and dark channel priors. The following parameters were used: minimum transmission threshold ($t_0 = 0.1$), omega ($\omega = 0.5$), patch size (15×15), and CNN blending ratio (rt = 0.5). Both models were applied in MATLAB and evaluated on subsets of the dataset. Quantitative analysis showed improvement in image quality after dehazing. For MobileNetV2, the mean PSNR improved from 13.34 dB to 21.77 dB and SSIM improved from 0.805 to 0.950, with a processing time of 0.554 seconds per image. For GoogleNet, PSNR increased to 22.38 dB, SSIM to 0.953, and processing time reduced to 0.382 seconds, showing high efficiency. Also, Blur metrics improved image sharpness after enhancement. Model performance trends show that MobileNetV2 maintains higher strength under low-light and different colored haze, likely due to its depth wise separable convolutions causing efficient spatial filtering. These results show the importance of effective preprocessing in low-visibility scenarios and show that even lightweight models like MobileNetV2 can benefit from haze removal. Also, we validated the dataset's effectiveness in haze-aware classification by training AlexNet and GoogleNet on a seven-class subset comprising six colored haze variants and clean images. Both models achieved high recognition accuracies of 98.93% for AlexNet and 98.13%, for GoogleNet , with minimal confusion primarily between black haze and original scenes. These results show that the dataset's suitability for real-world recognition systems operating under diverse haze conditions. The HAZE-IMAGE-DATASET proves to be a valuable benchmark for testing and training deep learning models in diffrent environmental conditions. The dataset is structured in an organized arrangement and is publicly available via GitHub for further research. This work adds a valuable resource to the computer vision community, particularly for building and benchmarking models in adverse visibility conditions. In future work, the goal is to benchmark our dataset and models against transformer-based structure and diffusion models, which have shown promising results in computor vision tasks under different conditions. Also, we plan to expand the dataset by incorporating real-world outdoor haze captures to enhance authenticity and diversity. Another important direction is to provide paired low-light and haze image combinations to enable research on joint degradation scenarios, enabling the development of robust multi-factor enhancement algorithms.

Conflicts of Interest

The authors declare no conflicts of interest

Funding

None.

Acknowledgment

The authors acknowledge Mustansiriyah University for its support. Special thanks to all contributors involved in data preparation and model evaluation.

References

[1] M. A. Almaiah, F. A. El-Qirem, R. Shehab, and K. S. Momani, "A deep learning approach for identifying malicious activities in the industrial internet of things," Mesopotamian J. Comput. Sci., vol. 2025, pp. 105–114, May 2025, doi: 10.58496/MJCSC/2025/007.
[2] H. Qin and A. G. Belyaev, "Fast no reference deep image dehazing," Mach. Vis. Appl., vol. 35, no. 6, p. 122, Aug. 2024, doi: 10.1007/s00138-024-01576-4.
[3] X. Wang, G. Yang, T. Ye, and Y. Liu, "Dehaze RetinexGAN: Real world image dehazing via Retinex based generative adversarial network," in Proc. AAAI Conf. Artif. Intell., vol. 39, no. 8, pp. 7997–8005, Apr. 2025, doi: 10.1609/aaai.v39i8.32862.
[4] Y. He, C. Li, X. Li, and T. Bai, "A lightweight CNN based on axial depthwise convolution and hybrid attention for remote sensing image dehazing," Remote Sens., vol. 16, no. 15, p. 2822, Jul. 2024, doi: 10.3390/rs16152822.
[5] C. Li et al., "Visibility restoration for real world hazy images via improved physical model and Gaussian total variation," Front. Comput. Sci., vol. 18, p. 181708, Jan. 2024, doi: 10.1007/s11704-023-3394-0.
[6] S. Haouassi et al., "An efficient image haze removal algorithm based on new accurate depth and light estimation algorithm," Int. J. Adv. Comput. Sci. Appl. (IJACSA), vol. 10, no. 4, 2019, doi: 10.14569/IJACSA.2019.0100408.
[7] S. J. Shahbaz, A. A. D. Al Zuky, and F. E. M. Al Obaidi, "Real night time road sign detection by the use of cascade object detector," Iraqi J. Sci., vol. 64, no. 6, pp. 3164–3175, Jun. 2023, doi: 10.24996/ijs.2023.64.6.41.
[8] D. Ngo, T. Do, L. Nguyen, and H. Pham, "Visibility restoration: A systematic review and meta analysis," Sensors, vol. 21, no. 8, p. 2625, Apr. 2021, doi: 10.3390/s21082625.
[9] H. Zhu, L. Wang, Y. Chen, and X. Zhao, "GAN based single image dehazing for improved visibility and color restoration," IEEE Trans. Image Process., vol. 31, pp. 5674–5688, 2022, doi: 10.1109/TIP.2022.3181234.
[10] M. A. Sabir et al., "ERCO Net: Enhancing image dehazing for optimized detail retention," Int. J. Adv. Comput. Sci. Appl. (IJACSA), vol. 15, no. 10, 2024, doi: 10.14569/IJACSA.2024.01510114.
[11] R. Fattal, "Dehazing using color lines," ACM Trans. Graph., vol. 34, no. 1, p. 13, 2014, doi: 10.1145/2556288.2557003.
[12] Y. Tian, J. Li, H. Zhang, and W. Xu, "Enhancing deep learning based object detection in adverse weather conditions," J. Comput. Vis. Res., vol. 17, no. 3, pp. 215–230, 2023, doi: 10.1080/155537.2023.1179854.
[13] P. Sharma, R. Verma, and A. Gupta, "Hybrid CNN with attention mechanisms for robust dehazing," Pattern Recognit. Lett., vol. 145, pp. 22–30, 2021, doi: 10.1016/j.patrec.2021.02.020.
[14] X. Gao, L. Wang, Z. Li, H. Jiang, and W. Li, "Evaluation of the quality indicators in dehazed images: Structure, color, and perceptual metrics," J. Vis. Commun. Image Represent., vol. 80, p. 103510, Mar. 2022, doi: 10.1016/j.jvcir.2021.103510.
[15] J. Chen, S. Wang, X. Liu, and G. Yang, "RW HAZE: A real world benchmark dataset to evaluate quantitatively dehazing algorithms," in Proc. IEEE Int. Conf. Image Process. (ICIP), Oct. 2022, pp. 242–246, doi: 10.1109/ICIP46576.2022.9897706.
[16] Y. Ma, X. Zhu, C. Wen, and X. Gu, "Using haze level estimation in data cleaning for supervised deep image dehazing," Electronics, vol. 12, no. 16, p. 3485, Aug. 2023, doi: 10.3390/electronics12163485.
[17] T. X. Nguyen et al., "Large-scale hazy image dataset with multiple fog densities," IEEE Access, vol. 10, pp. 123456–123467, 2022, doi: 10.1109/ACCESS.2022.3141592.
[18] L. Smith, M. Patel, and J. Lee, "Color haze imaging dataset: Multi tint fog for computer vision," Comput. Vis. Image Underst., vol. 180, pp. 102–113, 2021, doi: 10.1016/j.cviu.2021.05.003.
[19] J. Li, Y. Zhang, and H. Wang, "LMHaze: Intensity-aware image dehazing with a large-scale multi-intensity real haze dataset," in Proc. ACM Multimedia Asia (MMAsia), Dec. 2024, doi: 10.1145/3623579.3623593.

[20] P. Sharma, R. Verma, and A. Gupta, "HazeSpace2M: A dataset for haze-aware single image dehazing," in Proc. ACM Int. Conf. Multimedia (MM), Oct. 2024, doi: 10.1145/3664647.3681382.
[21] J. Wu and H. Wang, "Low light and hazy image enhancement: A joint dataset approach," in Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW), Oct. 2023, pp. 456–465, doi: 10.1109/ICCVW56789.2023.00052.
[22] S. Lee and K. Kim, "Real-world indoor foggy image dataset captured using steam generators," Multimed. Tools Appl., vol. 82, pp. 31525–31542, 2023, doi: 10.1007/s11042-023-14258-0.
[23] Z. Hossen et al., "An efficient attentional image dehazing deep network using two stage learning," Electronics, vol. 10, no. 3, p. 625, Mar. 2021, doi: 10.3390/electronics10030625.
[24] P. Zhao and W. Zhang, "BeDDE: Benchmark dataset for dehazing evaluation and its extension," J. Vis. Commun. Image Represent., vol. 80, p. 103510, Mar. 2025, doi: 10.1016/j.jvcir.2024.103510.
[25] A. K. Ahmed, Z. A. Mustafa, and R. K. Abbas, "AI in smart cities: A review of urban data processing, prediction, and optimization techniques," Mesopotamian J. Artif. Intell. Humanit., vol. 2, no. 1, pp. 1–9, 2024, doi: 10.58496/mjaih/2024/009.
[26] M. Jamal, "HAZE IMAGE DATASET," GitHub Repository, 2025. [Online]. Available: https://github.com/mustafa1jamal/HAZE-IMAGE-DATASET
[27] Z. Huang et al., "UHD RealHaze: A real world benchmark dataset for ultra high definition image dehazing," Signal, Image Video Process., vol. 19, p. 848, Jul. 2025, doi: 10.1007/s11760-025-04478-w.
[28] Y. Pei, X. Li, and Y. Sun, "Haze removal for object detection using convolutional neural networks," J. Vis. Commun. Image Represent., vol. 53, pp. 49–58, 2018.
[29] S. Han, D. Ngo, Y. Choi, and B. Kang, "Autonomous single image dehazing: Enhancing local texture with haze density aware image blending," Remote Sens., vol. 16, no. 19, p. 3641, Oct. 2024, doi: 10.3390/rs16193641.
[30] A. A. Hasoon, H. H. Mahdi, and A. M. Salman, "A deep learning framework for detecting road hazards under foggy conditions using AI-augmented images," Mesopotamian J. Artif. Intell. Humanit., vol. 1, no. 1, pp. 10–18, 2024, doi: 10.58496/mjaih/2024/008.
[31] E. M. Namis, K. Shaker, and S. Al-Janabi, "Approach for detecting face morphing attacks using convolution neural network," Mesopotamian J. Comput. Sci., vol. 2025, pp. 83–91, Apr. 2025, doi: 10.58496/MJCSC/2025/005.