Sea-Mie: Physically-Based Synthetic Fog for Maritime Image Defogging via Curriculum Learning

Anonymous CVPR submission

Paper ID *****

Abstract

Image dehazing in maritime environments is challenging due to highly variable fog densities, complex light–water interactions, and the lack of paired data that accurately reflects real conditions. We address these issues with a curriculum-based dehazing framework and a new synthetic maritime dataset generated in Blender using a full physical ray-scattering model. Unlike datasets such as RESIDE that apply haze through simplified atmospheric scattering, our dataset simulates volumetric light transport, multiple scattering, and realistic marine visibility across several fog densities for each scene. This structure enables a five-stage curriculum in which a dehazing network is trained from light to increasingly dense haze, improving convergence stability and reconstruction quality. Experiments show consistent gains over non-curriculum training. Together, our dataset and training strategy provide a physically grounded approach for robust, generalizable maritime image dehazing.

1. Introduction

Hazy and fog-degraded imagery poses a significant challenge for computer vision systems operating in maritime environments. Navigation, surveillance, and autonomous surface or underwater vehicles rely heavily on accurate visual perception, yet dense fog, sea spray, and aerosolized moisture can severely obscure scene geometry, attenuate contrast, and alter color fidelity. [9] [17]

Recent advances in image dehazing have shown promising performance under controlled conditions, but their effectiveness in real maritime settings remains limited by the scarcity of paired training data and the complexity of light transport in fog-laden coastal and offshore regions. [17] [16] Real maritime data is difficult to capture in a controlled manner, and synthetic datasets commonly used in dehazing research often rely on simplified atmospheric scattering models that do not fully capture the physics of marine fog. Existing synthetic datasets, such as RESIDE [11], generate fog by applying the atmospheric scattering model to clear images. While such datasets [2] [11] [1] are valuable and have supported considerable progress in learning-based dehazing, they are based on approximations of haze formation that assume homogeneous media and single-scattering effects. These assumptions rarely hold in maritime conditions, where fog composition, particle size distribution, multiple scattering, and strong light-water interactions produce highly nonlinear degradations.

As a result, networks trained solely on these approximations often struggle to generalize to real maritime images, leading to artifacts, color distortions, and incomplete haze removal. To address these limitations, we introduce a new synthetic maritime dehazing dataset created in Blender using a physically-based volumetric ray-scattering model. Instead of synthetically overlaying haze onto clean images, our dataset simulates full volumetric fog, multiple scattering events, depth-dependent extinction, and realistic maritime lighting conditions. Scenes include diverse vessel types, weather conditions, backgrounds, and fog densities, each rendered at multiple fog levels around a common ground-truth version. This enables the data to capture a richer set of visual degradations and provides a more realistic training distribution for data-driven dehazing models.

Beyond the dataset, we propose a curriculum learning strategy that exploits its multi-density structure. Curriculum learning has been shown to improve convergence stability by gradually increasing task difficulty, allowing neural networks to learn fundamental relationships before tackling more complex or ambiguous examples. [3] [14] [15] In our framework, the dehazing network is first trained using pairs of clear images and their least foggy counterparts. As training progresses, new stages introduce progressively heavier fog levels while retaining earlier, easier samples. This gradual progression encourages the network to first master basic structure and color reconstruction, then adapt its learned representations to increasingly severe degradations. Unlike multi-input schemes, each training example remains a simple pair of a single foggy image and its cor-

(a) Hazy Image

(b) One-shot Training

(c) Curriculum Learning

(d) Ground Truth

Figure 1. Comparison of one-shot direct training against curriculum learning approach. Curriculum learning is made possible using the sequence of increasing fog density in our dataset.

responding ground truth, ensuring architectural simplicity while benefiting from a well-designed difficulty schedule.

We demonstrate that the combination of physically-based synthetic data and a fog-density curriculum yields significant performance gains over traditional training regimes (Fig. 1). Models trained using our curriculum strategy exhibit improved PSNR, SSIM and perceptual quality. These results highlight the importance of both physically realistic training data and structured learning strategies in the development of robust dehazing systems for challenging maritime environments.

Overall, this work contributes:

- a novel physically-based synthetic maritime dehazing dataset generated using full volumetric ray scattering
- a fog-density curriculum learning framework that systematically increases difficulty during training
- extensive experiments demonstrating the benefits of combining physical realism and progressive training for achieving accurate results in maritime image dehazing.

2. Related Work

Early single-image dehazing methods rely on handcrafted priors to approximate the atmospheric scattering model. The most influential of these is the Dark Channel Prior (DCP) introduced by He et al. in 2009 [8], which exploits the statistical observation that non-sky patches typically contain pixels with very low intensity in at least one color channel, enabling effective estimation of scene transmission under haze. Although highly successful, prior-based methods assume homogeneous haze and limited scattering complexity, making them less suitable for challenging environments such as maritime scenes.

Deep neural networks have substantially advanced dehazing performance by learning direct mappings from hazy to clear images, but early CNN-based models [5] [12] remain constrained by limited receptive fields and difficulty modeling long-range dependencies. Recent work has shifted toward Transformer-based architectures that provide stronger global context modeling.

Song et al. introduced DehazeFormer [13], identifying several architectural limitations of Swin Transformers for low-level restoration and proposing improved normalization, activation, and spatial aggregation schemes. Their large-scale model achieves state-of-the-art performance on standard benchmarks, exceeding 40 dB PSNR on SOTS indoor. Guo et al. proposed DeHamer [7], a hybrid CNN–Transformer architecture that resolves feature inconsistency issues through feature modulation and introduces a transmission-aware 3D positional embedding, enabling the model to incorporate haze-density-related priors directly into the Transformer. Their method demonstrates state-of-the-art performance across multiple dehazing datasets. These works establish Transformers as powerful tools for dehazing, particularly when combined with haze-aware positional priors and hybrid CNN designs.

Most existing training datasets for dehazing—such as RESIDE [11]—apply haze synthetically using the atmospheric scattering model [6] [10]. While effective for large-scale supervised training, such datasets rely on simplified assumptions, including homogeneous media and predominantly single scattering. More recent directions use physically based rendering (PBR) to simulate volumetric scattering, multiple scattering events, and realis-

tic illumination, offering improved realism for outdoor environments. [18] [4] However, domain-specific PBR datasets—especially for maritime conditions—remain limited.

Curriculum learning [3] has shown benefits in several restoration tasks by structuring training from easy to hard samples, improving optimization stability and generalization. Although prior dehazing research has explored haze-aware priors [13] [7](e.g., density embeddings in Transformers), curriculum learning has not been explicitly applied to the dehazing problem. Notably, Transformer-based models such as DeHazeFormer and DeHamer focus on architectural improvements rather than data-level difficulty scheduling, leaving an opportunity to incorporate physically meaningful progression based on fog density.

Our work integrates insights from classical priors, recent Transformer-based models, physically based rendering, and curriculum learning. We introduce a physically simulated maritime dehazing dataset built with a volumetric ray-scattering model, together with a fog-density curriculum that progressively increases difficulty during training—an approach not explored in prior dehazing literature.

3. Method

3.1. Creation of Synthetic Data using Blender

We developed a fully randomized open-sea maritime scene generator using Blender. A pinhole camera with a resolution of 640×512, a fixed sensor size, and a $25mm$ focal length is placed at a randomly sampled height and assigned a random tilt angle. For each scene, a random subset of vessels is drawn from a predefined library and positioned on the ocean surface such that they lie within the camera's field of view and do not intersect with one another.

The ocean state—including wave height, wind direction, and additional surface parameters—is randomized to produce diverse water appearances and foam patterns. The background sky is selected at random from an extensive collection of HDR sky textures ranging from clear to heavily overcast conditions. Both the background illumination strength and azimuthal rotation about the global Z-axis are randomized, allowing the apparent sun position to vary freely over $360°$ around the vessels and the introduction of day/night effects.

To simulate fog, we enclose the scene in a large volumetric container and employ Blender's Mie scattering model. For each ground-truth (clear-weather) scene, we sample an aerosol particle diameter from the log-normal distribution shown in Fig. 3, centered around $5\mu m$ with a range of $1 - 25\mu m$. For this fixed aerosol size, we then generate 10 foggy variants using stratified random sampling of the fog density, where density values are drawn from a distribution proportional to $1/visibility$. This ensures a controlled progression from light to dense fog while maintaining physical consistency.

In contrast to most existing synthetic fog datasets [11] [7] [8], which rely on the classical atmospheric scattering model and therefore simulate only single-scattering events along each ray, our dataset incorporates full volumetric Mie scattering with multiple scattering per ray. As a result, light undergoes repeated scattering interactions as it travels through the fog volume, causing significantly stronger attenuation and yielding darker, denser, and more realistic fog for the same aerosol concentration. This behavior closely reflects real-world light transport in maritime environments, where multiple scattering dominates visibility loss. Consequently, our dataset provides a more physically faithful representation of fog, improving the reliability of supervised learning for defogging and visibility-restoration tasks. We mention here that for the time being we neglect absorption effects.

For every scene instance, we additionally output a depth map, a segmentation map, and a JSON file containing all metadata required for exact scene reconstruction. All ground-truth and foggy images are saved as 8-bit RGBA PNG files, while depth and segmentation maps are stored as 32-bit EXR files. Each iteration therefore yields one clear ground-truth scene together with 10 physically accurate fog realizations of identical geometry, lighting, ocean conditions, and camera parameters. This process enables the large-scale generation of paired (GT, foggy) image sets as shown in Fig. 2.

Ultimately, the goal is to produce thousands of such samples to form a comprehensive maritime dataset suitable for training and evaluating defogging and visibility-restoration models.

3.2. Curriculum Training

We adopt a state-of-the-art dehazing architecture [13] as the backbone for our framework and introduce a five-stage curriculum learning strategy that exploits the monotonic fog-density structure of our synthetic dataset. We start by testing our approach on 1 new fog density level being introduced at each stage, but the end goal is to introduce two per stage, for a total of 10 fog density levels. The models remain unchanged architecturally; the curriculum acts solely at the data-level, enabling staged exposure to increasingly severe fog conditions.

Training directly on heavily fogged images often leads to unstable gradients and suboptimal convergence, as dense haze severely obscures scene structure and attenuates high-frequency details. Our synthetic dataset provides five fog levels per ground-truth scene, ranging from light to extremely dense haze. This provides a natural ordering of difficulty, allowing the model to first learn basic restora-

Figure 2. Example of the sequential fogging proceedure. Fog is added to a clean image by gradually increasing the density parameter.

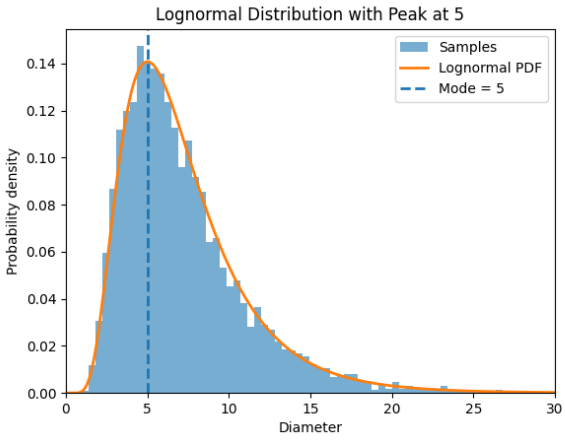

Figure 3. The log-normal distribution used for sampling the aerosol size. It is centered around $5\mu m$, has a range of $1 - 25\mu m$ and $\sigma = 0.5$

tion behavior under mild degradations before progressively addressing more complex scattering effects.

The curriculum consists of five stages, as shown in Fig. 4. Stage 1 uses only the lightest fog level. Subsequent stages incrementally add fog levels while retaining all previous ones, culminating in Stage 5, in which all five fog levels are included. Each training iteration draws uniformly from the available levels, and all samples remain in the form of independent (foggy image, ground-truth) pairs. No multi-input fusion or joint conditioning on multiple fog levels is used.

During Stage 1, the model learns fundamental contrast enhancement and low-level feature extraction. Stage 2 introduces moderate haze, prompting the network to generalize beyond lightly degraded inputs. In Stage 3, medium fog levels introduce stronger scattering and color distortions, requiring the model to infer missing details. Stage 4 further increases difficulty by adding heavy fog, while Stage 5 exposes the network to the full distribution of degradations. This progression maintains stability by anchoring the training signal with easier examples even as more challenging ones are introduced.

Training hyperparameters, including the optimizer and learning-rate schedule, remain fixed to ensure that improvements stem from the curriculum itself rather than modified optimization dynamics.

4. Results

We evaluate the effectiveness of the proposed five-stage curriculum training strategy by applying it to a state-of-the-art dehazing architecture [13] and comparing its performance against a series of non-curricular baselines. In the curriculum setting, the model is trained sequentially across five rounds, each incorporating progressively higher fog densities while retaining all lower-density samples. This exposes the network to an increasingly challenging set of degradations while ensuring stable learning dynamics anchored by earlier, easier examples.

To provide a fair comparison, we also train the same architecture independently on each individual fog-density subset of our dataset. This yields five separate baseline models, each optimized exclusively on hazy images of a single density level paired with their corresponding ground truth. Through this design, we compare two paradigms: a model that benefits from cumulative exposure to all hazy images up to a target density and a model trained only on the fog density of interest, without staged progression.

Quantitative evaluation reveals that the curriculum-trained models consistently outperform their single-density counterparts across all but the most extreme fog conditions (Figure 5). The curriculum approach delivers noticeable gains in both PSNR and SSIM (Figure 6), indicating that gradual exposure to increasing haze severity leads to more stable convergence and yields reconstructions with clearer structural details and improved color fidelity. These improvements demonstrate that the fog-density progression in our synthetic dataset effectively regularizes the learning process, enabling the network to build foundational representations before confronting highly degraded inputs.

However, both approaches show limitations at the highest fog densities, where severe information loss and heavy multiple scattering make accurate restoration inherently difficult. In these extreme scenarios, reconstructions become less consistent, suggesting that even staged training cannot fully compensate for the absence of recoverable image con-

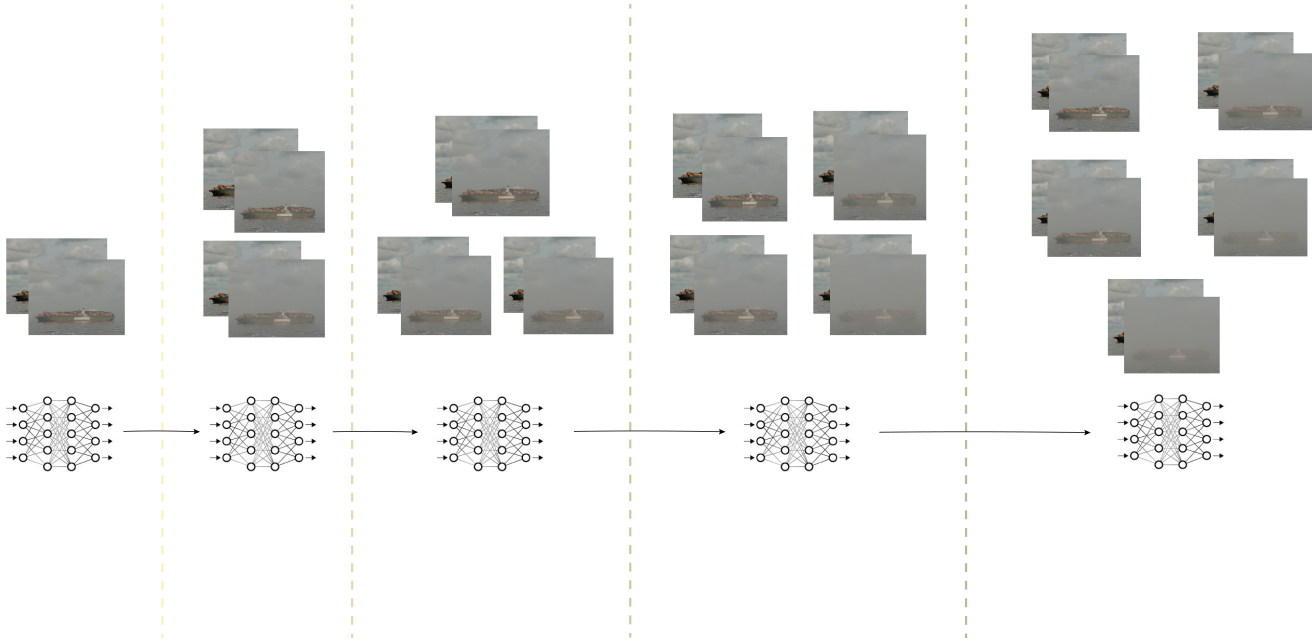

Figure 4. The proposed curriculum scheme. It consists of five stages, with each stage introducing more synthetic samples of increasing haze density.

Figure 5. Overview of test results for the hazy images of the highest fog density.

tent at the input level.

Future work will focus on further addressing these challenges by exploring additional curriculum stages with denser fog levels, evaluating a broader set of state-of-the-art architectures under the same multi-density framework, and developing a novel model explicitly designed to exploit the

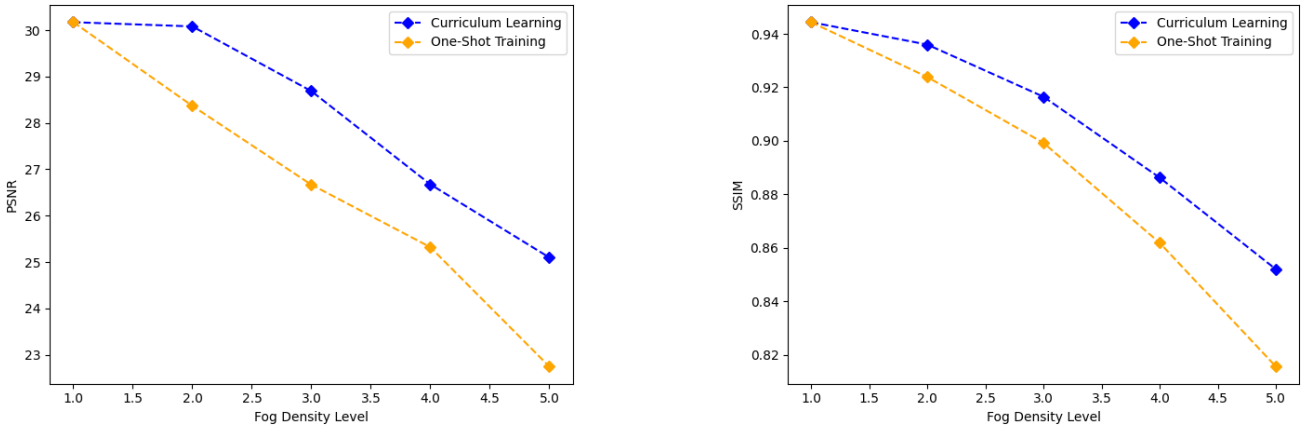

Figure 6. Difference of the two approaches in terms of PSNR and SSIM as fog density increases. We notice that curriculum learning yields more graceful decrease in both metrics as fog grows denser.

structured fog progression present in our dataset. By better leveraging cross-density cues and physical priors, such an architecture may further enhance robustness and generalization under real maritime visibility conditions.

5. Conclusion

We introduced a physically grounded synthetic maritime dehazing dataset generated using volumetric ray-scattering simulation, together with a five-stage curriculum learning strategy that exploits its structured fog-density progression. By training DehazeFormer from light to increasingly severe fog levels, the model acquires robust low-level priors in early stages and progressively adapts to complex scattering conditions in later ones. This staged formulation stabilizes optimization, improves generalization, and yields superior quantitative and perceptual performance compared to conventional training that exposes the network to all fog levels from the outset. Ablation results confirm that the full curriculum is critical: truncated or non-curricular schedules lead to degraded accuracy and slower convergence. Beyond providing a strong dehazing baseline, our findings highlight the importance of coupling physically realistic data with difficulty-aware training strategies. We believe this framework can serve as a foundation for future work in maritime vision, physically based simulation for restoration tasks, and curriculum mechanisms for low-level Transformer models.

References

[1] Codruta Orniana Ancuti, Cosmin Ancuți, and Radu Timofte. Nh-haze: An image dehazing benchmark with non-homogeneous hazy and haze-free images. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1798–1805, 2020. 1

[2] O. Codruta Ancuti, Ancuti Cosmin, Mateu Sbert, and Radu Timofte. Dense-haze: A benchmark for image dehazing with dense-haze and haze-free images. In *In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP)*, pages 1014–1018, 2019. 1

[3] Yoshua Bengio, Jérôme Louradour, Ronan Collobert, Yoesoep Rachmad, and Jason Weston. Curriculum learning. *Journal of the American Podiatry Association*, 60:6, 2009. 1, 3

[4] Zeyuan Chen, Yangchao Wang, Yang Yang, and Dong Liu. Psd: Principled synthetic-to-real dehazing guided by physical priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7180–7189, 2021. 3

[5] Qili Deng, Ziling Huang, Chung-Chi Tsai, and Chia-Wen Lin. Hardgan: A haze-aware representation distillation gan for single image dehazing. In *Computer Vision – ECCV 2020*, pages 722–738, Cham, 2020. Springer International Publishing. 2

[6] Raanan Fattal. Single image dehazing. *ACM Trans. Graph.*, 27(3):1–9, 2008. 2

[7] Chunle Guo, Qixin Yan, Saeed Anwar, Runmin Cong, Wenqi Ren, and Chongyi Li. Image dehazing transformer with transmission-aware 3d position embedding. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5802–5810, 2022. 2, 3

[8] Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1956–1963, 2009. 2, 3

[9] Xin He, Tong Jia, and Junjie Li. Learning degradation-aware visual prompt for maritime image restoration under adverse weather conditions. *Frontiers in Marine Science*, Volume 11 - 2024, 2024. 1

[10] Harald Koschmieder. Theorie der horizontalen sichtweite. *Beiträge zur Physik der freien Atmosphäre*, 1924. 2

[11] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single image dehazing and beyond, 2019. 1, 2, 3

[12] Wenqi Ren, Jinshan Pan, Hua Zhang, Xiaochun Cao, and Ming-Hsuan Yang. Single image dehazing via multi-scale convolutional neural networks with holistic edges. *International Journal of Computer Vision*, 128:240–259, 2020. 2

[13] Yuda Song, Zhuqing He, Hui Qian, and Xin Du. Vision transformers for single image dehazing. *IEEE Transactions on Image Processing*, 32:1927–1941, 2023. 2, 3, 4

[14] Petru Soviany, Radu Tudor Ionescu, Paolo Rota, and Nicu Sebe. Curriculum learning: A survey. *International Journal of Computer Vision*, 130:1526–1585, 2022. 1

[15] Xin Wang, Yudong Chen, and Wenwu Zhu. A Survey on Curriculum Learning . *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 44(09):4555–4576, 2022. 1

[16] Yongpeng Wei, Fang Liu, and Aruhan Yan. Unsupervised aerial maritime image dehazing via pooling transformer and multi-granularity attention optimization. In *2025 6th International Conference on Computer Vision, Image and Deep Learning (CVIDL)*, pages 599–602, 2025. 1

[17] Yuhang Xie, Meng Li, Siqi Wang, and Hongbo Wang. Overwater-haze: A large-scale overwater paired image dehazing dataset. *Processes*, 13(8), 2025. 1

[18] Shijun Zhou, Xing Xie, Baojie Fan, and Jiandong Tian. Physics-guided image dehazing diffusion, 2025. 3