CARLA-Haze: A Synthetic Benchmark for Outdoor Image Dehazing

Henry O. Velesaca[1,2] Leo Thomas Ramos[3,4] Angel D. Sappa[1,3,4]
[1]ESPOL Polytechnic University, Ecuador [2]University of Granada, Spain
[3]Computer Vision Center, Spain [4]Universitat Autònoma de Barcelona, Spain
hvelesac@espol.edu.ec; ltramos@cvc.uab.cat; sappa@ieee.org

Abstract

This paper presents CARLA-Haze, a synthetic dataset designed for outdoor image dehazing. CARLA-Haze contains 10,000 high-resolution paired images (clean and hazy), distributed across 10 different scenarios and 10 incremental haze intensity levels. The dataset includes diverse scenarios and visual elements, distance-based haze distribution consistent with realistic atmospheric conditions, and multiple viewpoints enhancing visual variability. Additionally, two preprocessed versions with standardized resolutions of 640×480 and 512×512 pixels are provided, each containing 40,000 images. CARLA-Haze also offers predefined splits for training, validation, and testing to facilitate its use. Experiments demonstrate that models trained with CARLA-Haze exhibit improved generalization in both synthetic and real-world settings, and that dehazing enhances downstream object detection performance under adverse visibility conditions. Dataset is available at https://leo-thomas.github.io/CARLA-Haze/

1. Introduction

Haze is an atmospheric phenomenon caused by fine particles suspended in the air, such as dust, smoke, or pollutants [6, 25, 35]. It reduces visibility and degrades image quality [16, 33], which significantly impacts image analysis and computer vision tasks such as segmentation, object detection, and classification [5, 33, 39]. These challenges are particularly critical in applications like autonomous driving, remote sensing, and surveillance [14, 26], where accurate scene interpretation is essential for decision-making .

Given the challenges that haze poses to various fields, image dehazing has been a hotspot of research in computer vision over the past decade [35]. Significant efforts have been dedicated to developing techniques that restore visibility in hazy images [15, 26], improving their usability in real-world applications. These methods, when trained with a sufficient number of image pairs, have demonstrated remarkable performance in haze removal [26]. However, acquiring high-quality data for this task is not trivial [22].

Collecting real-world images for image dehazing presents significant challenges. Capturing paired hazy and haze-free images requires maintaining identical conditions [5, 6], including illumination, scene composition, and atmospheric properties. However, natural variations in weather, light dynamics, and pollution make this nearly impossible in outdoor environments [15, 37]. Additionally, even in controlled environments, achieving consistency between hazy and reference images requires specialized equipment, such as vapor machines [39], and precise calibration, making the process both labor-intensive and expensive [12].

Furthermore, in most real-world scenarios, haze-free reference images simply do not exist [22, 37], as haze is an inherent part of the scene and cannot be selectively removed without altering other visual properties [22]. This scarcity of ground-truth data limits the availability of large-scale [27], diverse datasets that cover a wide range of haze intensities and scene complexities. As a result, the field lacks standardized benchmarks [21, 36], complicating the evaluation of dehazing models and hindering progress in the development of more robust approaches [19, 37].

Given the mentioned challenges, the generation of synthetic datasets has gained relevance in the field of computer vision [7, 17], proving especially useful for image dehazing. By simulating haze effects computationally, researchers can produce large volumes of data without the constraints associated with real-world image acquisition. These datasets serve as a scalable and controlled source of paired hazy and haze-free images, ensuring precise ground truth for supervised learning approaches. However, despite their advantages, there are still gaps to be addressed in existing synthetic datasets. Many of them contain a limited number of images, lack scene diversity, and often apply haze at fixed levels.

Based on the above, this work introduces CARLA-Haze (Fig. 1), a synthetic dataset for image dehazing. The dataset is generated using the CARLA simulator and consists of paired hazy and haze-free images across diverse scenarios. To provide a comprehensive range of haze intensities, it in-

Figure 1. CARLA-Haze and its main features.

cludes ten distinct levels, ranging from 10% to 100%. Each level contains 1,000 high-resolution images (1600×1200), resulting in a total of 10,000 hazy images along with their corresponding clear counterparts.

To evaluate the dataset's applicability, we design experiments assessing its effectiveness in two key areas. First, we examine the generalization capability of dehazing models trained on CARLA-Haze and tested on real-world data. Second, we explore its utility for data augmentation by integrating synthetic images into real-world training sets. The results indicate that CARLA-Haze is a valuable resource, offering a structured and scalable dataset that can contribute to advancing research in image dehazing. The contributions of this work can be summarized as follows:

- We introduce CARLA-Haze, a synthetic dataset for image dehazing composed of 10,000 high-resolution hazy and haze-free image pairs across diverse scenarios and ten haze levels. Two standardized versions are also provided at 640×480 and 512×512 resolutions, each with 40,000 images and predefined training, validation, and test splits.
- We establish a benchmark by evaluating sota dehazing models on CARLA-Haze, providing a reference for future comparisons.
- We show that CARLA-Haze is effective for data augmentation, enhancing the performance of models when combined with real-world training data.
- We demonstrate that dehazing with CARLA-Haze-trained models leads to notable improvements in downstream object detection on real-world hazy images.

2. Related work

2.1. Real-world datasets

Real-world dehazing datasets provide valuable resources for training and evaluating dehazing models. However, many of these datasets could be considered pseudo-real, as capturing the same scene with and without haze under identical environmental conditions is generally unfeasible, often requiring artificial haze generation methods.

One approach to constructing real-world dehazing datasets involves applying physical scattering models to haze-free images. D-HAZY [1] uses Koschmieder's model along with depth maps to simulate haze varying with object distance, while HazeRD [38] adopts Mie scattering to generate gradual haze variations under different weather conditions. RESIDE [18], one of the most widely used datasets, introduces synthetic haze across multiple subsets for indoor and outdoor tasks using an atmospheric model. Haze4K [20] follows a similar strategy, generating diverse haze levels by varying light and scattering parameters over real-world images.

Another method involves capturing outdoor scenes with artificially introduced haze using professional fog machines, enabling realistic light–haze interactions. O-HAZE [2] captures paired hazy and haze-free images under identical conditions, with careful control of environmental factors. DenseHaze [4] focuses on producing dense and homogeneous haze, while NH-HAZE [5] introduces non-homogeneous haze to better approximate real atmospheric variability, maintaining consistency through controlled recording settings. Other alternatives include manual haze generation, as in NH-RESIDE [22], which introduces layered cloud-like patterns with varying opacity using Adobe Photoshop to simulate non-homogeneous haze and increase variability.

2.2. Synthetic-world datasets

Given the challenges of acquiring real-world image pairs with and without haze under identical conditions, synthetic dehazing datasets provide a viable alternative. By leveraging computer-generated environments, these datasets allow for the creation of controlled and diverse scenarios where haze can be systematically introduced.

A common approach involves the use of traffic simulation environments. These tools allow for the capture of both haze-free and hazy images of the same scene while maintaining identical illumination, object positioning, and perspective. For instance, the datasets presented in [29, 30] were generated using the SiVIC simulator, a tool designed to model urban environments and vehicle sensors. It enables the creation of haze-free and hazy image pairs with consistent illumination and scene structure. Haze is simulated using Koschmieder's Law, introducing four levels of non-homogeneous haze to better mimic realistic visibility conditions.

More recent approaches leverage advanced simulation platforms and game engines to generate synthetic dehazing datasets with increased realism and flexibility. NT-HAZE [28] is a nighttime traffic dataset built with the CARLA

Dataset	Type	Image pairs	Resolution	Scene domain	Year	Ref.
SiVIC1	S	18	640×480	Outdoor	2010	[29]
SiVIC2	S	66	640×480	Outdoor	2012	[30]
D-HAZY	R	1,400	Varied	Indoor	2016	[1]
HazeRD	R	14	Varied	Outdoor	2017	[38]
O-HAZE	R	45	5,456×3,632	Outdoor	2018	[2]
DenseHaze	R	33	5,456×3,632	O/I	2019	[4]
RESIDE	R	72,135	Varied	O/I	2019	[18]
NH-HAZE	R	55	5,456×3,632	Outdoor	2020	[5]
GTA5	S	864	512×256	Outdoor	2020	[34]
Haze4K	R	4,000	Varied	Outdoor	2021	[20]
NT-HAZE	S	10,000	768×768	Outdoor	2022	[28]
CARLA-Haze	S	10,000	1,600×1,200	Outdoor	2025	Ours

Table 1. CARLA-Haze compared with other outdoor dehazing datasets. 'S' denotes synthetic datasets, while 'R' indicates (pseudo) real datasets.

simulator, where haze is modeled using Mie scattering to simulate varying intensities such as light, moderate, and dense fog. Similarly, the dataset in [34] exploits the GTA5 game engine to render urban and suburban scenes under diverse nighttime lighting and haze conditions, offering fine-grained control over environmental parameters like haze density and distribution.

A summary of the reviewed datasets is shown in Table 1. As observed, real-world datasets are grounded in natural environments, preserving realistic lighting and object layouts, but rely on the assumption of truly haze-free counterparts, an unfeasible condition in practice. Additionally, setups involving haze machines, while physically accurate, demand specialized equipment and strict environmental control, increasing the complexity and cost of data collection.

Meanwhile, interest in fully synthetic dehazing datasets has grown, though they remain less common than pseudo-real alternatives. These datasets enable precise control over haze and scene parameters, yielding perfectly aligned image pairs. Still, their effectiveness depends on covering a broad range of scenarios, object types, and atmospheric conditions with sufficient realism to support generalization to real-world settings. Without this, models trained on synthetic data may fail to perform reliably in real environments.

3. CARLA-Haze

3.1. Image generation

The CARLA [13] simulator (version 0.9.14), was used to generate the synthetic images for this study. Carla is an open-source platform designed for autonomous driving research, that provides high-fidelity 3D environments with configurable weather, lighting, and sensor setups, making it especially suitable for tasks that require controlled yet realistic visual conditions. Eight virtual towns were selected, details of which are shown in Table 2. These scenarios collectively represent a wide variety of settings, including

Town	Details
Town01	A small, simple town with a river and several bridges.
Town02	A small simple town with a mixture of residential and commercial buildings.
Town03	A larger, urban map with a roundabout and large junctions.
Town04	A small town embedded in the mountains with a special 'figure of 8' infinite highway.
Town05	Squared-grid town with cross junctions and a bridge.
Town06	Long many lane highways with many highway entrances and exits. It also has a Michigan left.
Town07	A rural environment with narrow roads, corn, barns and hardly any traffic lights.
Town10	A downtown environment with skyscrapers, residential buildings and an ocean promenade.

Table 2. Details of CARLA towns used for the generation of CARLA-Haze images.

simple towns, dense urban areas, highway networks, rural regions, and mountainous roads. Such variety allows the dataset to incorporate different object distributions, structural layouts, and visual elements, which is crucial for training dehazing models that can generalize effectively to diverse real-world conditions.

For image acquisition, a dual-camera system is configured on a single vehicle within the CARLA simulator. Both cameras operate at a resolution of 1600×1200 pixels and a field of view (FOV) of 100 degrees. The first camera collects clean, haze-free RGB images, while the second simulates fog conditions using the Rayleigh scattering model [24]. This model accounts for the scattering of light by particles significantly smaller than the wavelength of light [24, 32], enabling a physically realistic representation of atmospheric haze. The fog intensity is controlled through the fog density parameter, which is varied from 10% to 100% in steps of 10%, generating ten distinct haze levels.

To introduce variability, random perturbations are applied to the camera pose at each timestep. Translations vary within ± 2 meters along the X and Y axes and ± 1 meter along the Z axis. Rotations include pitch values between $0°$ and $7°$, and yaw and roll within $\pm 7°$. These spatial adjustments help introduce natural variation in the viewpoints while maintaining alignment between clean and hazy pairs. These variations, combined with the distinct elements present in each town, result in diverse visual contexts that ensure broader coverage in terms of structure, layout, and environmental appearance.

At each fog density level, 100 image pairs are generated across 10 scenes. Importantly, the number of scenes does not directly correspond to the number of towns, as some

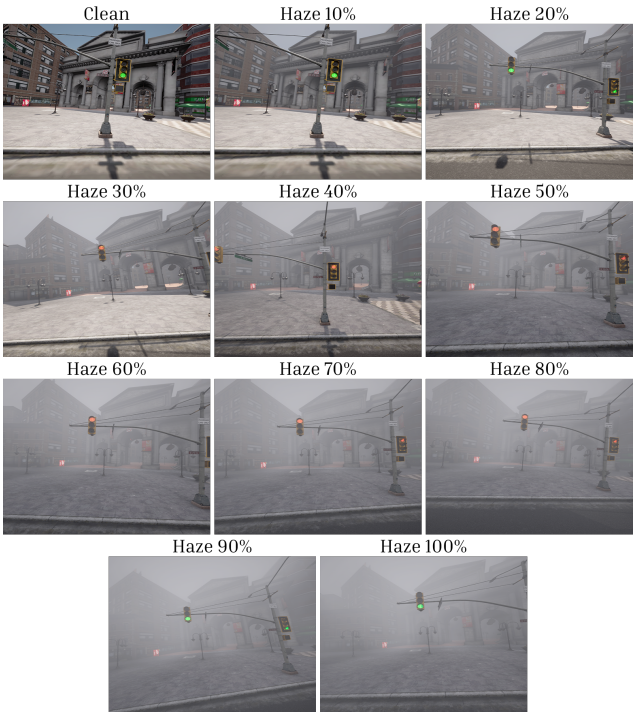

Figure 2. Examples of the different haze intensity levels included in CARLA-Haze.

towns are used more than once with different vehicle trajectories. This results in a total of 10,000 image pairs, with perfect alignment across haze levels and high variability in scene composition.

3.2. Dataset details

Haze intensity levels. One of the characteristics of CARLA-Haze is that it contains 10 haze intensity levels. These haze intensities range progressively from 10% (light haze) up to 100% (dense haze), as shown in Fig. 2, providing a comprehensive set of scenarios that simulate varying visibility conditions. This allows models to explicitly learn varying scattering coefficients, facilitating better generalization to real-world hazy conditions.

Distance-based haze distribution. Although CARLA-Haze images have specified haze intensity levels, the atmospheric scattering model employed distributes haze in a physically consistent manner based on object distance. Due to this, distant objects are more affected by haze, appearing more occluded due to increased scattering, while closer objects maintain higher visual clarity, as shown in Fig. 3. This distance-dependent behaviour aligns with real-world atmospheric physics, providing a perceptually realistic distribution rather than a uniformly dense haze layer.

Figure 3. Examples of distance-based haze distribution in CARLA-Haze. Note how objects closer to the camera exhibit higher visual clarity, while distant objects appear increasingly occluded due to atmospheric scattering, despite all scenes having 100% haze intensity.

Figure 4. Sample images illustrating the diversity of scenarios and visual elements in CARLA-Haze, including urban structures, rural areas, vegetation, vehicles, street infrastructure, and miscellaneous objects.

Diverse scene elements. CARLA-Haze contains a diverse set of visual elements, including buildings, vehicles, vegetation, street infrastructure, and recreational objects, as shown in Fig. 4, obtained from the multiple virtual scenarios provided by CARLA. These elements feature different shapes, textures, structures, and colors, providing high visual variability. Such diversity reflects distinct material properties and object geometries, allowing models to explicitly capture a broad range of visual patterns and reflectance characteristics under hazy conditions.

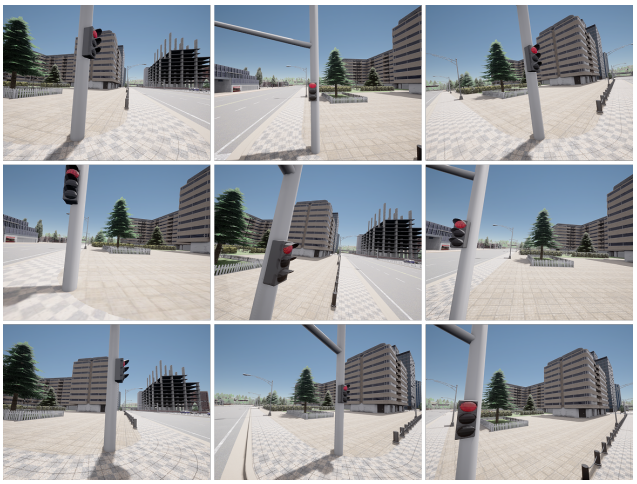

Figure 5. Examples showcasing variations in viewpoint, distance, and object scale captured across CARLA-Haze images. Note how the traffic light's position and apparent size vary across frames.

Visual perspective variability. CARLA-Haze contains images captured at different viewpoints and distances within the same scenario, resulting in variations of visual perspective. Consequently, scene elements exhibit positional and scale changes, as shown in Fig. 5, enhancing the dataset's visual richness. This intrinsic variability functions similarly to data augmentation, supporting the development of robust dehazing models capable of handling changes in viewpoints and scale.

Training, validation, and test splits. CARLA-Haze contains a total of 10,000 image pairs in its original form. These images are organized into 10 distinct scenes, each featuring 10 haze intensity levels, with 100 paired images (hazy and clean) per haze level. In addition to the original high-resolution images, we provide two processed versions of the dataset for easier use: one at 640×480 pixels and another at 512×512 pixels.

To generate the resized versions, a dedicated crop-and-resize procedure was used, rather than a direct resize, which would cause significant detail loss, or a direct crop, which could produce patches with insufficient contextual information. Images were first cropped using a window 75% larger than the target size to capture broader scene contexts and ensure higher variability. These cropping windows were strategically placed at each corner of the high-resolution images, producing overlapping regions among adjacent crops. Subsequently, each crop was resized to its target resolution. This approach ensures that each patch includes multiple objects and contextual elements, preserving broader visual diversity and representativeness.

Consequently, each resized dataset variant contains 40,000 patches, maintaining the structure of 10 scenes and 10 haze levels. Furthermore, each variant was partitioned into training, validation, and test subsets following an 80-15-5 ratio, allowing researchers to conveniently train and evaluate models. Users can thus flexibly select data by haze intensity level or create custom combinations according to their experimental requirements or objectives.

Haze level (%)	PSNR ↑	SSIM ↑	Training time (hours)
10	31.91	0.968	7.27
20	30.68	0.963	7.23
30	29.89	0.961	7.27
40	36.40	0.922	7.27
50	23.27	0.859	7.22
60	22.88	0.854	7.22
70	22.59	0.837	7.21
80	22.44	0.839	7.22
90	22.54	0.839	7.25
100	21.94	0.823	7.25

Table 3. CARLA-Haze benchmarking results for 300 epochs using the ConvIR [9] model.

4. Experiments

4.1. CARLA-Haze benchmarking

To establish a reliable performance baseline on the CARLA-Haze dataset, we conduct an initial benchmarking experiment. The goal of this stage is to provide a consistent reference point for subsequent experiments and comparisons. For this purpose, we employ ConvIR [9], a state-of-the-art model for image dehazing, training it independently on each haze level defined in the dataset. A total of 300 epochs is used in each case to ensure comparability across all configurations using an Nvidia A100 SXM4 40GB GPU.

The benchmarking results are presented in Table 3. As expected, the dehazing performance generally degrades as the haze density increases. Both PSNR and SSIM values exhibit a clear downward trend, reflecting the increasing difficulty of the task. Notably, performance is high and stable at low haze levels (10%-30%), with PSNR values above 29 dB and SSIM consistently above 0.96. This indicates that the model handles mild haze conditions effectively. A significant drop occurs beyond 40%, where PSNR and SSIM values decline more sharply. This is particularly evident from 50% onward, where PSNR falls below 24 dB and SSIM drops under 0.86.

These results provide a first validation of the CARLA-Haze dataset's capacity to represent a controlled and progressive haze simulation. The model's consistent decline in performance with increasing haze levels confirms that the dataset properly reflects the expected difficulty gradient. Moreover, the training time remains stable across all levels, indicating that the added haze complexity does not lead to

optimization instability or convergence issues under the current setup. Likewise, the figures highlight that considerable room for improvement remains, especially in high-density haze conditions. This reinforces the role of CARLA-Haze not only as a benchmarking tool but also as a foundation for developing and validating future dehazing approaches.

Additionally, the dataset allows for the combination of samples across different haze levels to increase training diversity and simulate more realistic conditions. Although this strategy is not explored in the current work due to computational resource constraints, it represents a promising direction for future research enabled by the dataset's structure.

4.2. CARLA-Haze for augmenting real-world training data

This experiment evaluates the use of CARLA-Haze as a source of synthetic data to augment limited real-world training sets. The goal is to assess whether combining synthetic and real data improves the generalization of dehazing models compared to training solely on real samples. O-HAZE [3] is selected for this experiment because it is an exclusively outdoor benchmark with paired hazy and haze-free images and the most representative of this category, which aligns with the outdoor focus of CARLA-Haze and ensures consistency in evaluation. We consider five subsets of it containing 5%, 10%, 20%, 25%, and 50% of the available training samples. For each case, we compare models trained only on real data against models trained on the same real subset combined with CARLA-Haze images from levels 8, 9, and 10. All configurations are trained for 500 epochs. This joint training strategy is applied across four different state-of-the-art architectures to ensure that the observed effects are consistent and architecture-independent.

The results presented in Table 4 show a consistent improvement when models are trained using both synthetic and real data, compared to training on real data alone. This trend is observed across all four architectures and for all real data subsets, confirming that the inclusion of CARLA-Haze provides complementary information that enhances model generalization.

For example, with only 5% of the real dataset, all models exhibit clear improvements in both PSNR and SSIM when trained jointly with CARLA-Haze. In particular, in EENet, the PSNR increases from 16.15dB to 17.93dB, while in ChaIR the improvement reaches 1.5dB approximately. This setting is particularly revealing, as the real data available is minimal, meaning that the model relies almost entirely on the synthetic domain for learning. The fact that the joint models outperform their real-only counterparts under these conditions suggests that CARLA-Haze provides sufficient structural and visual information to drive effective training even in near-absence of real supervision.

As the proportion of real data increases, the relative improvement becomes less pronounced. This is most likely due to the fact that all configurations are trained for the same number of epochs, which may limit the model's ability to fully exploit the larger datasets. Nonetheless, the joint training approach consistently leads to better results across all real data subsets. For instance, in the 50% setting, ChaIR improves from 18.08dB to 19.21dB in PSNR. This indicates that the synthetic data does not lose its relevance in more data-rich regimes, but continues to act as a beneficial complement that supports generalization.

Overall, the results confirm that the inclusion of CARLA-Haze leads to consistent improvements across all real data proportions and architectural variants. These gains demonstrate that the proposed dataset serves not only as an effective substitute in low-data regimes but also as a valuable complement in richer settings. This highlights its potential as a practical and scalable resource for supporting the development of more robust dehazing models. It is also worth noting that all models are trained under a fixed and relatively limited training schedule, and further gains could be expected with extended training, particularly for architectures that require longer convergence.

4.3. Usability on downstream tasks

This experiment investigates the impact of applying models trained with CARLA-Haze in downstream tasks, with the goal of evaluating their potential utility in real-world applications. Specifically, we focus on object detection, using the RTTS dataset[1], which contains real-world images exhibiting a variety of haze-related anomalies and conditions. To this end, we employ a YOLOv12-nano model [31], which is a state-of-the-art detection model widely used in different applications of object detection [23]. It was trained for 100 epochs on the RTTS training split, and evaluated on its corresponding test set. Then, we process the test images using a ChaIR model trained on CARLA-Haze (levels 8, 9, and 10 combined), and then re-run the object detector on the dehazed outputs.

The detection results are shown in Table 5. These indicate slight but consistent improvements across most metrics when using dehazed inputs. Specifically, the overall mAP@50 increases from 0.761 to 0.766, and mAP@50:95 rises from 0.550 to 0.555. These gains, while moderate, are achieved without retraining or fine-tuning the detector, demonstrating that the restoration process can directly benefit downstream performance in adverse visual conditions.

The improvements are particularly noticeable in the "Person" and "Motorbike" categories, which are often more affected by haze due to their fine-grained features and smaller visual footprints. The observed gains suggest that the dehazed inputs provide cleaner signals that facilitate ob-

[1] https://sites.google.com/view/reside-dehaze-datasets/reside-%CE%B2

Model	Data	Real data (5%)		Real data (10%)		Real data (20%)		Real data (25%)		Real data (50%)	
		PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
OKNet [10]	Only real	16.87	0.7527	16.42	0.7110	16.80	0.7272	17.54	0.7653	17.28	0.7581
	Joint	18.15	0.7592	17.21	0.7242	16.98	0.7400	18.02	0.7734	17.89	0.7645
EENet [11]	Only real	16.15	0.7334	17.43	0.7582	18.44	0.8002	16.67	0.7516	17.10	0.7618
	Joint	17.93	0.657	17.67	0.7591	18.88	0.8013	18.07	0.7761	18.72	0.8037
ConvIR [9]	Only real	17.23	0.7648	17.62	0.7669	17.53	0.7637	17.68	0.7290	17.24	0.7589
	Joint	17.87	0.7785	17.73	0.7613	17.73	0.7650	17.80	0.7749	17.41	0.7602
ChaIR [8]	Only real	15.99	0.7150	17.21	0.7292	16.57	0.7550	17.12	0.7663	18.08	0.7648
	Joint	17.50	0.7650	17.73	0.7469	17.51	0.7669	18.90	0.8043	19.21	0.8090

Table 4. Quantitative results on the O-HAZE dataset using different proportions of real training data, with and without augmentation using CARLA-Haze. Joint training with CARLA-Haze consistently improves performance across all data regimes.

Model	Class	Precision	Recall	mAP@50	mAP@50:95	Speed (ms)
YOLOv12n	Bicycle	0.807	0.586	0.725	0.545	4.2
	Bus	0.745	0.590	0.692	0.504	
	Car	0.865	0.782	0.862	0.624	
	Motorbike	0.780	0.607	0.704	0.485	
	Person	0.866	0.720	0.819	0.589	
	All	0.812	0.657	0.761	0.550	
YOLOv12n (dehazed images)	Bicycle	0.828	0.563	0.722	0.546	4.4
	Bus	0.729	0.617	0.696	0.506	
	Car	0.871	0.778	0.866	0.628	
	Motorbike	0.783	0.617	0.719	0.498	
	Person	0.876	0.721	0.828	0.597	
	All	0.817	0.659	0.766	0.555	

Table 5. Detection results on the RTTS evaluation set (1,296 images). Comparison between predictions on the original hazy images and those obtained after dehazing with a model trained on CARLA-Haze.

ject localization and classification, especially in cases where haze may have otherwise reduced detector confidence.

To better visualize the impact of dehazing on object detection, qualitative results are presented in Fig. 6. As shown, the dehazed versions of the images exhibit noticeably clearer scenes, with enhanced contrast and sharper edges that reveal more structural detail. This visual improvement translates into more reliable predictions by the detection model, as reflected in the increased confidence scores associated with the detected objects (Fig. 6a).

Beyond improved confidence, the processed images also lead to a reduction in misclassifications (Fig. 6b). The dehazed inputs help the model maintain more coherent predictions, minimizing overlapping bounding boxes and reducing confusion between object instances. These qualitative observations reinforce the quantitative gains discussed earlier and illustrate the practical value of using CARLA-Haze-based restoration in real-world visual pipelines.

In sum, the experimental results provide strong evidence of the practical value of CARLA-Haze. On one hand, the use of CARLA-Haze for augmenting limited real-world datasets leads to consistent improvements across a range of architectures and data regimes, confirming its utility as a complementary source of supervision. On the other hand, its application in downstream tasks such as object detection shows that the visual enhancements produced by models trained on CARLA-Haze can translate into more reliable predictions and fewer misclassifications in real-world scenarios. These findings position CARLA-Haze as a versatile and impactful dataset for both model training and deployment under adverse visibility conditions.

5. Limitations

While CARLA-Haze provides a diverse and well-structured collection of synthetic images across multiple scenarios and haze intensities, it is important to acknowledge certain limitations. First, it does not include certain visibility conditions, such as nighttime scenes or alternative outdoor environments that are less represented in CARLA. These gaps suggest promising directions for extending the dataset toward broader outdoor variability. Another limitation of this study is that it presents a primary evaluation of CARLA-Haze's capabilities. Naturally, additional experimentation would help further understand the dataset's utility when combined with other architectures, and reveal how well its benefits generalize across other real-world datasets.

6. Conclusions and Future Work

This work introduces CARLA-Haze, a synthetic dataset developed to support research in outdoor image dehazing. The dataset is constructed using CARLA simulator under 10 distinct urban, semi-urban, and rural scenarios, and 10 incremental haze levels. The dataset comprises a total of 10,000 clean-hazy image pairs. The scenes are built to capture a range of visibility conditions with realistic depth-dependent haze distribution and diverse object and texture compositions. Additionally, two preprocessed versions at standardized resolutions (640×480 and 512×512) are released, ex-

(a) Improvement in confidence scores.

(b) Reduction of misclassifications.

Figure 6. Object detection results using original and processed images from the RTTS dataset, performed by YOLOv12. Results show that processing images with a model trained on CARLA-Haze leads to: (a) improved confidence scores, resulting in more reliable predictions; and (b) reduced misidentifications and increased recognition of challenging instances.

panding the dataset to 40,000 samples and enabling flexible usage across training pipelines.

The conducted experiments validate the utility of CARLA-Haze across multiple dimensions. Benchmarking results using a state-of-the-art model demonstrate the progressive difficulty associated with increasing haze intensity, confirming that the dataset reflects meaningful visibility degradation. Furthermore, joint training with CARLA-Haze and real data improves performance across all tested architectures and data regimes, even under limited supervision. Finally, its impact on downstream tasks is confirmed through object detection experiments, where preprocessing with CARLA-Haze-trained models yields measurable gains in detection accuracy, confidence, and consistency.

Future work may explore extending the dataset to additional conditions, such as nighttime visibility or alternative outdoor environments, not currently represented in CARLA. Likewise, investigating its usefulness in broader pipelines or additional downstream applications, such as segmentation, could provide further insight into its versatility. Overall, CARLA-Haze establishes a foundation for continued research in outdoor image restoration under controlled synthetic variability.

Acknowledgment

This work was partially supported by Grant PID2024-162815NB-I00 funded by MICIU/AEI/ 10.13039/501100011033 and by ERDF/EU, and Grant PID2021-128945NB-I00 funded by MICIU/AEI/ 10.13039/501100011033 and by ERDF/EU. The authors acknowledge the support of the Generalitat de Catalunya CERCA Program to CVC's general activities, and the Departament de Recerca i Universitats from Generalitat de Catalunya to the SGR Research Group 2021 MACO (reference 2021 SGR 01499).

References

[1] Cosmin Ancuti, Codruta O. Ancuti, and Christophe De Vleeschouwer. D-hazy: A dataset to evaluate quantitatively dehazing algorithms. In *ICIP*, 2016. 2, 3

[2] Codruta O. Ancuti, Cosmin Ancuti, Radu Timofte, and Christophe De Vleeschouwer. O-haze: A dehazing benchmark with real hazy and haze-free outdoor images. In *CVPRW*, 2018. 2, 3

[3] Codruta O. Ancuti, Cosmin Ancuti, Radu Timofte, and Christophe De Vleeschouwer. O-haze: a dehazing benchmark with real hazy and haze-free outdoor images. In *CVPRW*, 2018. 6

[4] Codruta O. Ancuti, Cosmin Ancuti, Mateu Sbert, and Radu Timofte. Dense-haze: A benchmark for image dehazing with dense-haze and haze-free images. In *ICIP*, 2019. 2, 3

[5] Codruta O. Ancuti, Cosmin Ancuti, and Radu Timofte. Nh-haze: An image dehazing benchmark with non-homogeneous hazy and haze-free images. In *CVPRW*, 2020. 1, 2, 3

[6] Jiyou Chen, Shengchun Wang, Xin Liu, and Gaobo Yang. Rw-haze: A real-world benchmark dataset to evaluate quantitatively dehazing algorithms. In *ICIP*, 2022. 1

[7] Richard J. Chen, Ming Y. Lu, Tiffany Y. Chen, Drew F. K. Williamson, and Faisal Mahmood. Synthetic data in machine learning for medicine and healthcare. *Nature Biomedical Engineering*, 5(6):493–497, 2021. 1

[8] Yuning Cui and Alois Knoll. Exploring the potential of channel interactions for image restoration. *Knowledge-based Systems*, 282:111156, 2023. 7

[9] Yuning Cui, Wenqi Ren, Xiaochun Cao, and Alois Knoll. Revitalizing convolutional network for image restoration. *IEEE TPAMI*, 46(12):9423–9438, 2024. 5, 7

[10] Yuning Cui, Wenqi Ren, and Alois Knoll. Omni-kernel network for image restoration. In *AAAI*, 2024. 7

[11] Yuning Cui, Qiang Wang, Chaopeng Li, Wenqi Ren, and Alois Knoll. Eenet: An effective and efficient network for single image dehazing. *Pattern Recognition*, 158:111074, 2025. 7

[12] Jin-Ting Ding, Yong-Yu Peng, Min Huang, and Sheng-Jun Zhou. Agrigan: unpaired image dehazing via a cycle-consistent generative adversarial network for the agricultural plant phenotype. *Scientific Reports*, 14(1):14994, 2024. 1

[13] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. CARLA: An open urban driving simulator. In *CoRL*, 2017. 3

[14] Bhawna Goyal, Ayush Dogra, Dawa Chyophel Lepcha, Vishal Goyal, Ahmed Alkhayyat, Jasgurpreet Singh Chohan, and Vinay Kukreja. Recent advances in image dehazing: Formal analysis to automated approaches. *Information Fusion*, 104:102151, 2024. 1

[15] Xiaojie Guo, Yang Yang, Chaoyue Wang, and Jiayi Ma. Image dehazing via enhancement, restoration, and fusion: A survey. *Information Fusion*, 86-87:146–170, 2022. 1

[16] Md Tanvir Islam, Nasir Rahim, Saeed Anwar, Muhammad Saqib, Sambit Bakshi, and Khan Muhammad. Hazespace2m: A dataset for haze aware single image dehazing. In *ACMMM*, 2024. 1

[17] Indu Joshi, Marcel Grimmer, Christian Rathgeb, Christoph Busch, Francois Bremond, and Antitza Dantcheva. Synthetic data in human analysis: A survey. *IEEE TPAMI*, 46(7):4957–4976, 2024. 1

[18] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single-image dehazing and beyond. *IEEE TIP*, 28(1):492–505, 2019. 2, 3

[19] Bo Liu, Si-Bao Chen, Jia-Xin Wang, Jin Tang, and Bin Luo. An oriented object detector for hazy remote sensing images. *IEEE TGRS*, 62:1–11, 2024. 1

[20] Ye Liu, Lei Zhu, Shunda Pei, Huazhu Fu, Jing Qin, Qing Zhang, Liang Wan, and Wei Feng. From synthetic to real: Image dehazing collaborating with unlabeled real data. In *ACM MM*, 2021. 2, 3

[21] Priya Narayanan, Xin Hu, Zhenyu Wu, Matthew D. Thielke, John G. Rogers, Andre V Harrison, John A. D'Agostino, James D Brown, Long P. Quang, James R. Uplinger, Heesung Kwon, and Zhangyang Wang. A multi-purpose realistic haze benchmark with quantifiable haze levels and ground truth. *IEEE TIP*, 32:3481–3492, 2023. 1

[22] Vinay P, Abhisheka K S, Lithesh Shetty, Kushal T M, and Shylaja S S. Non homogeneous realistic single image dehazing. In *WACVW*, 2023. 1, 2

[23] Leo Thomas Ramos and Angel D. Sappa. A decade of you only look once (yolo) for object detection: A review. *IEEE Access*, 13:192747–192794, 2025. 6

[24] Lord Rayleigh. On the transmission of light through an atmosphere containing small particles in suspension, and on the origin of the blue of the sky. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 47(287):375–384, 1899. 3

[25] Mohit Singh, Vijay Laxmi, and Parvez Faruki. Visibility enhancement and dehazing: Research contribution challenges and direction. *Computer Science Review*, 44:100473, 2022. 1

[26] Yuda Song, Zhuqing He, Hui Qian, and Xin Du. Vision transformers for single image dehazing. *IEEE TIP*, 32:1927–1941, 2023. 1

[27] Hang Sun, Zhiming Luo, Dong Ren, Bo Du, Laibin Chang, and Jun Wan. Unsupervised multi-branch network with high-frequency enhancement for image dehazing. *Pattern Recognition*, 156:110763, 2024. 1

[28] Chunming Tang and Wenzhe Yao. Ndpc-net: A dehazing network in nighttime hazy traffic environments. In *ICPICS*, 2022. 2, 3

[29] Jean-Philippe Tarel, Nicolas Hautière, Aurélien Cord, Dominique Gruyer, and Houssam Halmaoui. Improved visibility of road scene images under heterogeneous fog. In *IEEE IV Symposium*, 2010. 2, 3

[30] Jean-Philippe Tarel, Nicholas Hautiere, Laurent Caraffa, Aurelien Cord, Houssam Halmaoui, and Dominique Gruyer. Vision enhancement in homogeneous and heterogeneous fog. *IEEE ITSM*, 4(2):6–20, 2012. 2, 3

[31] Yunjie Tian, Qixiang Ye, and David Doermann. Yolov12: Attention-centric real-time object detectors. *arXiv preprint arXiv:2502.12524*, 2025. 6

[32] Xingang Wang, Junwei Tian, Yalin Yu, Qin Wang, Xin Yao, Yupeng Feng, and Haokai Gao. A modified atmospheric scattering model and degradation image clarification algorithm for haze environments. *Optics Communications*, 560: 130489, 2024. 3

[33] Haiyan Wu, Yanyun Qu, Shaohui Lin, Jian Zhou, Ruizhi Qiao, Zhizhong Zhang, Yuan Xie, and Lizhuang Ma. Contrastive learning for compact single image dehazing. In *CVPR*, 2021. 1

[34] Wending Yan, Robby T. Tan, and Dengxin Dai. Nighttime defogging using high-low frequency decomposition and grayscale-color networks. In *ECCV*, 2020. 3

[35] Tian Ye, Yunchen Zhang, Mingchao Jiang, Liang Chen, Yun Liu, Sixiang Chen, and Erkang Chen. Perceiving and modeling density for image dehazing. In *ECCV*, 2022. 1

[36] Jing Zhang, Yang Cao, Zheng-Jun Zha, and Dacheng Tao. Nighttime dehazing with a synthetic benchmark. In *ACM MM*, 2020. 1

[37] Shengdong Zhang, Xiaoqin Zhang, Shaohua Wan, Wenqi Ren, Liping Zhao, Li Zhao, and Linlin Shen. Diverse hazy image synthesis via coloring network. *IEEE TAI*, 5(7):3703–3713, 2024. 1

[38] Yanfu Zhang, Li Ding, and Gaurav Sharma. Hazerd: An outdoor scene dataset and benchmark for single image dehazing. In *ICIP*, 2017. 2, 3

[39] Shiyu Zhao, Lin Zhang, Shuaiyi Huang, Ying Shen, and Shengjie Zhao. Dehazing evaluation: Real-world benchmark datasets, criteria, and baselines. *IEEE TIP*, 29:6947–6962, 2020. 1