

Article

A Foggy Weather Simulation Algorithm for Traffic Image Synthesis Based on Monocular Depth Estimation

Minan Tang ^{1,*} , Zixin Zhao ²  and Jiandong Qiu ² ¹ College of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730050, China² College of Electrical and Mechanical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China; 12221501@stu.lzjtu.edu.cn (Z.Z.); qiujd@mail.lzjtu.cn (J.Q.)

* Correspondence: tangminan@mail.lzjtu.cn; Tel.: +86-138-9368-8178

Abstract: This study addresses the ongoing challenge for learning-based methods to achieve accurate object detection in foggy conditions. In response to the scarcity of foggy traffic image datasets, we propose a foggy weather simulation algorithm based on monocular depth estimation. The algorithm involves a multi-step process: a self-supervised monocular depth estimation network generates a relative depth map and then applies dense geometric constraints for scale recovery to derive an absolute depth map. Subsequently, the visibility of the simulated image is defined to generate a transmittance map. The dark channel map is then used to distinguish sky regions and estimate atmospheric light values. Finally, the atmospheric scattering model is used to generate fog simulation images under specified visibility conditions. Experimental results show that more than 90% of fog images have AuthESI values of less than 2, which indicates that their non-structural similarity (NSS) characteristics are very close to those of natural fog. The proposed fog simulation method is able to convert clear images in natural environments, providing a solution to the problem of lack of foggy image datasets and incomplete visibility data.

Keywords: traffic safety; monocular depth estimation; atmospheric scattering model; fog simulation; natural scene statistics



Citation: Tang, M.; Zhao, Z.; Qiu, J. A Foggy Weather Simulation Algorithm for Traffic Image Synthesis Based on Monocular Depth Estimation. *Sensors* **2024**, *24*, 1966.

<https://doi.org/10.3390/s24061966>

Academic Editor: Sergio A. Velastin

Received: 16 February 2024

Revised: 13 March 2024

Accepted: 14 March 2024

Published: 20 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The continuous development of artificial intelligence technology has pushed self-driving cars to the forefront of societal discussion, raising significant concerns about their safety. According to data collected by the World Health Organisation (WHO) from 178 countries up to 2023, 1.19 million people worldwide lose their lives each year due to road traffic accidents [1]. Road traffic injuries have huge economic costs for individuals, families, and entire countries, including the cost of treatment for the dead and injured as well as the loss of labour for the deceased, those disabled by their injuries, and family members who need to take time away from work or school to care for the injured. Statistically, losses from road traffic accidents account for 3 percent of the gross domestic product of most countries. The main causes of road accidents around the world include human factors, road factors, vehicle factors, and environmental factors. Among these factors, adverse weather conditions play a crucial role. The focus on adverse weather conditions is of great significance in research directed towards road traffic safety. The data and background information highlight the potential benefits of self-driving cars, especially in terms of improving road traffic safety. As artificial intelligence technology continues to advance, self-driving cars can adapt to adverse weather conditions by sensing and responding in real time, reducing the number of accidents. The application of such technology is expected to reduce the number of road deaths globally and alleviate the heavy burden borne by individuals, families, and national economies. Overall, it is more realistic to strengthen the research and application of autonomous driving technology, especially in bad weather conditions, while concentrating on foggy target detection, which is more challenging due to

the problems of colour distortion and the loss of detailed features in foggy images, which bring many difficulties in extracting foggy image features. This will not only help reduce casualties and property damage but also promote sustainable socio-economic development.

In recent years, researchers in the field of target detection have shifted from traditional image processing methods to convolutional neural network (CNN)-based approaches [2–8] focusing on image defogging algorithms that improve scene visibility. This shift has triggered a significant amount of research aimed at improving image quality and enhancing the performance of target detection. Codruta et al. [7] proposed a new single-image defogging method for improving the visibility of foggy images. The method of the paper is a fusion-based strategy that extracts inputs from two original blurred image by applying white balance and contrast enhancement procedures. Filtering of luminance, chromaticity, and saliency features by calculating three metrics is used to efficiently blend the derived inputs in order to retain regions with good visibility; the multi-scale approach is designed using a Laplace pyramid representation to minimise artefacts introduced by the weight map. Liu et al. [9] proposed to add a dark channel fog removal algorithm based on the YOLOv7 algorithm and effectively improve the fog removal efficiency of the dark channel through the methods of downsampling and upsampling.

Although many algorithms have achieved the goal of removing haze, the introduction of defogging algorithms into foggy detection channels still faces high costs in practical applications. This means that deploying the algorithms on resource-constrained devices or applications that need to run in real time is expensive and infeasible. Therefore, it is particularly important to improve the accuracy of foggy road traffic target detection models without adding extra costs. For data-driven algorithms, the amount and diversity of training data directly affects the performance of the model. However, the lack of foggy weather datasets becomes a major obstacle to improve the performance of foggy weather target detection models. Therefore, in order to improve the accuracy of target detection in foggy weather conditions, the size of the dataset containing road traffic scenes must be significantly increased.

According to the above study, in order to scale up a traffic scene dataset, the dataset must meet certain criteria: the dataset should cover a wide range of application scenarios, show a variety of categories, contain a variety of environmental features, and preferably have a high frame rate. Well-known open-source annotated datasets commonly used for target detection in autonomous driving traffic scenarios include KITTI [10], Cityscapes [11], and BDD100K [12]. After careful evaluation, the BDD100K dataset emerged as the most suitable choice, as it contains a wider variety of data categories than similar datasets. In addition, the dataset is geographically, environmentally, and meteorologically diverse, which greatly facilitates the development of data-driven models. Extending the dataset to include foggy conditions is a feasible approach, where foggy traffic images are artificially synthesised from sunny day images. This approach allows trained models to improve performance in all scenarios. Similar approaches have been used to augment existing datasets for scenes such as rainy [13] days and nighttime [14]. For example, the literature [15] discusses the use of synthetic images for lane line detection, while previous studies have investigated synthetic fog images for automated driving, including FRIDA [16], FRIDA2 [17], Foggy Cityscape [18], and Multifog KITTI [19]; few studies have quantitatively evaluated artificially generated fog image enhancement datasets to determine their fidelity. Tarel et al. constructed digital image-based synthetic outdoor fog datasets, namely FRIDA (Foggy Road Image Database) and FRIDA2, using *SiVIC*TM software, but these datasets contain a limited number of images with low resolution and poor visual quality and are therefore insufficient to support the training of target detection models with higher accuracy and on a larger scale, especially in fog scenes. Sakaridis et al. developed Foggy Cityscape, and Mai et al. introduced synthetic fog into Cityscapes and KITTI images to create Multifog KITTI, but the scale of these datasets remains relatively small. In addition, these fog image synthesis methods rely on image perception and depth information acquired by sophisti-

cated sensors such as LiDAR, RGBD cameras, or binocular cameras, which adds additional cost in practice.

In this paper, we propose a new dataset enhancement method to extract depth information from monocular clear weather images using a self-supervised monocular depth estimation method, and then simulate foggy weather images based on an atmospheric scattering model. Our proposed framework estimates the depth map only from the original image, which means that the method can be easily applied to other datasets without the need for additional sensors to collect depth information. To improve the accuracy of the target detection model in foggy scenes, we simulate foggy images using the method in this paper on the BDD100K dataset, resulting in the FoggyBDD100K dataset. The quality assessment of our simulated FoggyBDD100K dataset verifies the authenticity of our obtained dataset and the effectiveness of this foggy sky simulation framework. The main contributions of this thesis include:

- (1) A new enhancement and simulation method for a foggy sky target detection dataset is proposed, which significantly improves the accuracy and generalisation of the foggy sky target detection model without introducing additional computational cost and complex algorithmic framework.
- (2) We have created a new dataset, FoggyBDD100K, which contains 16,702 foggy sky images. This will help the community and researchers to develop and validate their own data-driven target detection or defogging algorithms.
- (3) We quantitatively evaluated the obtained simulated foggy sky images and obtained quantitative evaluation results while validating the reliability of this foggy sky simulation framework.

Given the above research, traffic safety becomes a necessary factor in the study of severe weather target detection. About the rest of this article: Section 2 describes the methods used in this paper. Section 3 comparatively analyses and presents the reference-free image quality assessment algorithm used in this paper for fog map assessment. Section 4 describes the experimental procedure and analyses the experimental results. Finally, Section 5 summarises the work of this paper.

2. Methods

This study presents a foggy sky simulation algorithm developed for the BDDK100 dataset. The method uses a self-supervised monocular depth estimation method to extract relative depth information from monocular images. Next, a dense geometric constraint method is used to determine the absolute depth information in the image. A simulated haze image is then generated using an atmospheric scattering model. Finally, the simulated haze image is quantitatively assessed using the AuthESI image quality assessment algorithm.

2.1. Self-Supervised Relative Depth Estimation Network

The acquisition of accurate scale information directly from monocular images poses a common challenge for researchers, often necessitating the use of lidar [20–22] or stereo cameras [23–27] for depth estimation. Among stereo methodologies, Mo [24] proposed a fusion method that combines the advantages of tightly coupled depth sensors and stereo cameras to complete dense depth estimation, thereby achieving better depth estimation. Xiaofeng [27], on the other hand, studied accurate depth estimation under different mechanical parameters such as camera calibration and alignment errors. ManyDepth [28] leverages the matching geometric properties across time-series images to enhance the accuracy of single-image depth prediction. However, under these related methods, the cost increases significantly, which undoubtedly indirectly increases the difficulty of generalisation. In addition, the bulky structure reduces the speed of reasoning, which in turn reduces the applicability of the methods.

The focus of this study is to achieve accurate relative depth estimation using the Monodepth2 benchmark network, which employs a self-supervised monocular depth estimation network. The algorithmic framework consists of two main components: the depth

estimation network and the pose estimation network. The task of the depth estimation network is to estimate the depth value of each input image at the pixel level. By utilising the depth estimation function of Monodepth2, this network can accurately compute the depth map, thus enhancing the overall depth estimation process. Meanwhile, the pose estimation network is responsible for estimating the six-degree-of-freedom poses between neighbouring input image frames. Its main function is to construct self-supervised signals between images through geometric transformations. This network of the algorithm is crucial for building robust self-supervised learning signals that help the network to refine the depth estimation through geometric relationships in the image sequence.

This section delves into the intricacies of the unsupervised monocular depth estimation framework, with a focus on the depth estimation network responsible for obtaining depth information about an image. This network is the cornerstone of the research and the core of the article. The depth estimation network is implemented through a U-Net-based encoder–decoder network architecture, the architecture of which is crucial for understanding and optimising monocular depth estimation methods. The encoder employs multiple convolutional layers, each of which performs downsampling operations to generate feature maps at different scales. As shown in Figure 1, the Monodepth2 Depth Estimation Network Architecture Diagram visually summarises the intricate architecture and interrelationships between components in the depth estimation network. This analysis provides a comprehensive understanding of the network functionality and lays the foundation for advancing unsupervised monocular depth estimation methods.

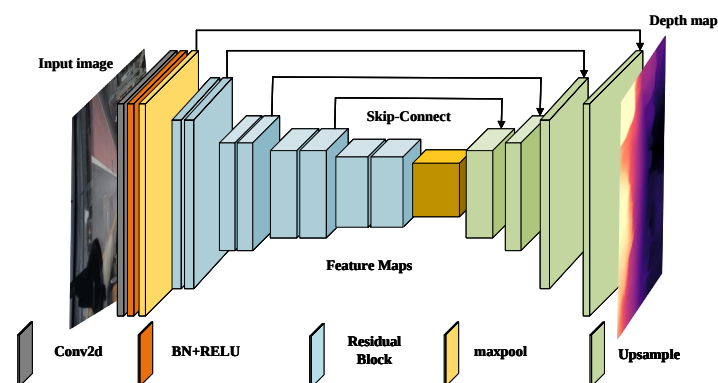


Figure 1. Depth estimation network structure used by Monodepth2.

In order to solve the challenges posed by bilinear sampling and local gradients caused by local minima in the current study, the baseline network DNet in the literature [29] was incorporated. The commonality with reference [30–33] is the use of a multi-scale depth prediction strategy. To further mitigate depth artefacts and enhance object-level depth inference, a novel Disparity Correction and Propagation (DCP) layer is introduced. The proposed DCP layer explicitly merges features at different scales in a hierarchical manner. The design rationale for the layered structure stems from the observation that low-resolution layers within the decoder network provide more reliable object-level depth inference, while high-resolution layers focus on capturing local depth details. The algorithm uses DCP to initiate the depth prediction process, perform relative depth estimation, and generate structural features.

Specifically, a convolutional neural network in the DCP layer is used to normalize the number of feature channels at different scales to eight, which is used to simplify the subsequent computation. Subsequently, features from the low-resolution layer are upsampled and feature fusion is performed with features from the high-resolution layer. This fusion introduces more accurate object-level inference to the higher-resolution depth prediction, thus improving the accuracy of the generated depth maps.

2.2. Absolute Depth Scale Recovery Based on Monodepth2

This module utilizes the geometric relationship between the ground and the camera in the dataset, focusing on extracting suitable ground points and then calculating the distance from the camera to the actual pixels in the image. It works by selecting ground points with stability and scale consistency as landmarks for the initial scale calculation. This assumption is in line with the general case of automated driving, where a sufficient number of ground points in a monocular image can be recognized.

After obtaining the relative depth map through the relative depth network, the module starts the scale recovery. The relative depth values are the basis for projecting the coordinates of each pixel point in the image into 3D space, and they follow the principles outlined in Equation (1). The result of this projection forms the corresponding 3D coordinates, paving the way for a robust ground-based scale recovery process.

$$D_t^{rel}(p_{i,j})p_{i,j} = KP_{i,j}, \quad (1)$$

where $p_{i,j}$ represents the pixel point in the i -th row and j -th column with homogeneous coordinates in two-dimensional space, $P_{i,j}$ is a 3D point in the corresponding 3D space, $D_t^{rel}(p_{i,j})$ is the relative depth corresponding to a pixel point in the 2D space, and K is the camera internal matrix.

Subsequently, in this paper, the eight-point method is used to calculate the surface normal vector. The eight-point method is a method to determine the surface normal vector of a pixel point by considering eight neighboring pixel points around the pixel. As shown in Figure 2, the red point in the center is the pixel point for which the surface normal vector is to be determined, and the eight pixel points around it are divided into four groups according to a specific rule, with each group containing two points. According to Equation (2), two sets of vectors are constructed starting from the red pixel point and ending at two pixel points in each group, ensuring that each group of points is perpendicular to the vector of the center pixel point, thus creating planes that remain orthogonal in three-dimensional space:

$$n_s = \overrightarrow{C_{ij}p_{1,s}} \times \overrightarrow{C_{ij}p_{2,s}}, \quad (2)$$

$$N(p_{i,j}) = \left(\sum_{s \in (1,4)} \frac{n_s}{\|n_s\|_2} \right) / 4, \quad (3)$$

where s denotes different planes ranging from 1 to 4, and $p_{1,s}$ and $p_{2,s}$ are the other two points in the plane apart from the centre pixel point $C_{i,j}$. The surface normal vector perpendicular to the plane n_s is obtained by cross-multiplying the vectors of the two points with the centre pixel. Finally, the surface normal vector of the centre point is obtained after normalization and averaging $N(p_{i,j})$.

Immediately thereafter, a method for confirming ground points using normalized ground point normals obtained in the camera's coordinate system is presented. In this system, the right, bottom, and front axes of the camera align with the X, Y, and Z axes in the world coordinate system. An ideal ground normal (denoted as ' $N = (0, 1, 0)^T$ ') is chosen as a reference for confirming ground points based on the obtained normalized ground point normals.

The uncertainty in the estimation of surface normals and the non-strict perpendicularity of normalized ground normals are taken into account. The study used the chosen ideal surface normal ($N = (0, 1, 0)^T$) with the normal vector of each 3D point in the image for the angle calculation. A threshold close to 0° is set to take into account the possible error introduced by the non-horizontal plane on which the vehicle is located in the image. This threshold controls the identification of "ground point clusters".

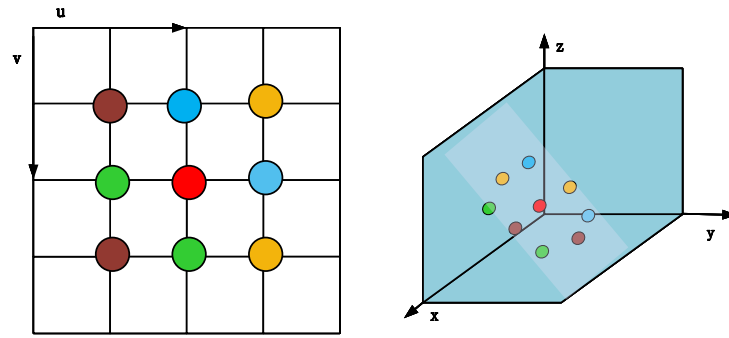


Figure 2. The above figure calculates the projection from 2D space to 3D space and the pairing of 8 neighbouring points. Points of the same colour are paired and then two vectors are formed separately from the centre point, and these two vectors are perpendicular to each other. Four surface normals can be calculated from the four vector pairs and used to form a surface normal.

Regarding the calculation of the angular difference, Equation (4) was chosen to calculate the similarity angle based on the absolute value similarity function of the cosine function:

$$S=s(P_{i,j}) = |\angle(\tilde{n}, N(P_{i,j}))| = \left| \arccos \frac{\tilde{n} \cdot P_{i,j}}{\|\tilde{n}\| \|P_{i,j}\|} \right|, \quad (4)$$

After the extraction of ground points from the monocular image and the creation of a dense ground point map, the height from each ground point to the camera is individually estimated. This process yields a substantial dataset of camera heights, and the actual camera height is subsequently determined through a probabilistic calculation method.

The camera height is the projection of the $\overrightarrow{OP_{i,j}}$ in the direction of the surface normal at grounding point $N(P_{i,j})$, denoting the distance from the camera to the ground point as $\overrightarrow{OP_{i,j}}$. The camera height can be calculated using the following equation:

$$h(P_{i,j}) = N(P_{i,j})^T \cdot \overrightarrow{OP_{i,j}}, \quad (5)$$

where $\overrightarrow{OP_{i,j}}$ denotes the vector from the camera to the grounding point.

Due to the large number of ground points, the anomalies hardly affect the estimated value of the true camera height. The obtained median value is determined as the estimated height of the camera, i.e., $h_M = \text{Median}(h(P_{i,j}))$. In addition to this, the actual height of the camera h_R is needed to construct the scale factor as shown in the following equation:

$$f_t = \frac{h_R}{h_M}, \quad (6)$$

After obtaining the corresponding scale factor, the final absolute depth value is obtained by combining the data obtained from the relative depth map. The following equation shows the method of absolute depth information acquisition by scale factor:

$$D_t^{abs} = f_t \cdot D_t^{rel}, \quad (7)$$

where D_t^{abs} is the calculated absolute distance, f_t is the scale factor, and D_t^{rel} is the relative distance.

2.3. Atmospheric Scattering Model

In foggy conditions, image quality degradation is attributed to the effect of suspended particles in the atmosphere that scatter sunlight and attenuate reflected light. The theory of the atmospheric scattering model proposed by McCartney (McCartney) provides insight into the specific reasons for the loss of colour information and contrast degradation in foggy images. As shown in Figure 3, the foggy day visual imaging model consists of two basic

components: an incident light attenuation model and an atmospheric light reflection interference model. The incident light attenuation model describes the energy change in reflected light propagating from the target scene to the acquisition point. Atmospheric light consists mainly of direct light, scattered light from the sky, and reflected light from the ground. As the model shows, the farther the light propagates, the greater the effect of atmospheric light on the fog imaging process. We can see that as the distance between the camera and the object increases, while the effect of atmospheric light increases, the light from the object decreases, leading to a decrease in contrast. The image captured by a typical three-channel RGB (red, green, and blue) camera can be represented by the following equation:

$$I_C = \int_{\lambda} G_{\lambda} h_{\lambda} R_{\lambda,C} + V_{\lambda} (1 - h_{\lambda}) R_{\lambda,C} d\lambda, \quad (8)$$

$$h_{\lambda} = e^{-d\alpha_{\lambda}}, \quad (9)$$

where λ is the wavelength, V_{λ} is the airborne light, h_{λ} is the transmittance, α_{λ} is the fog scattering coefficient, and d is the distance between the object and the camera, $C \in \{R, G, B\}$ is the image channel, and $R_{\lambda,C}$ is the spectral response of the camera for a particular image channel.

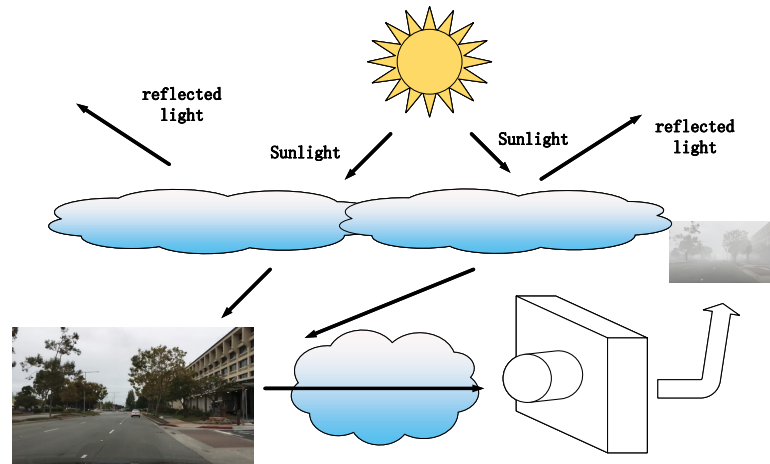


Figure 3. Physical modelling of atmospheric scattering. The light received in the camera sensor consists of light scattered by particles suspended in the air and light reflected by the imaging object, which is attenuated as it travels through the air.

This representation captures the intricate interrelationship between atmospheric light and object light, providing us with a comprehensive understanding of the phenomenon of fog imaging.

2.4. Simulated Fog Map for Absolute Depth Estimation

In conditions of dense fog or mist, the scattering coefficient is contingent on the prevailing weather conditions. Specifically, the scattering parameter is derived from the visible range R_m through the relationship expressed as $\beta = -\ln(\varepsilon)/R_m$, with its value determined based on an intuitive understanding of the visual range. However, within the visible spectral region, the scattering coefficient α_{λ} remains relatively constant. To streamline the algorithm, a constant value is assigned to the wavelength, denoted as $\alpha_{\lambda} = \alpha$ (commonly set to 0.02). Consequently, a simplified atmospheric scattering model is obtained as follows:

$$I_C(x, y) = G_C(x, y)h(x, y) + V_C(x, y)(1 - h(x, y)), \quad (10)$$

$$h(x, y) = e^{-\alpha d(x, y)}, \quad (11)$$

In Equation (10), (x, y) is the spatial location coordinates of the image, $C \in \{R, G, B\}$ is the image channel, $I_C(x, y)$ is the image obtained from one of the channels in the image after simulating fog, $G_C(x, y)$ is the fog-free image of a particular channel of the image, and $V(x, y)$ is the global atmospheric light value, which indicates the influence of other light paths in the atmospheric environment on the direction of perception, which is usually a global constant. In Equation (11), $h(x, y)$ is the transmittance, i.e., transmittance map, which is used to describe the ability of light to penetrate the suspended particles in the air on foggy days, and since the scene distances in the three channels of the image can be approximately regarded as the same and the scattering coefficients are almost constant, the transmittance is the same for all the three channels and the value range is roughly from 0 to 1. $V(x, y)$ is the scattering rate and $d(x, y)$ is the distance from the detected object to the vision sensor, i.e., the scene depth.

Hence, based on the foggy visual imaging model, the simulation of a foggy scene $I(x, y)$ can be conceptualized as a process of deriving solutions from an existing fog-free image $G(x, y)$. Essentially, solving for $I(x, y)$ entails determining the unknown transmittance $h(x, y)$ and atmospheric light value $V(x, y)$ with respect to $G(x, y)$. In other words, the solution for $I(x, y)$ involves identifying the unknown transmittance $h(x, y)$ and atmospheric light value $V(x, y)$, both of which are contingent on the characteristics of $G(x, y)$.

Hence, to enhance computational efficiency, we derived a simplified atmospheric scattering model. In the process of obtaining transmittance maps and atmospheric light values for foggy scene simulation, a depth map was acquired using a monocular depth estimation network and dense geometrical constraints. Subsequently, Equation (10) was employed to simulate the dataset image. The subsequent simulation process will be elucidated as shown in Figure 4.

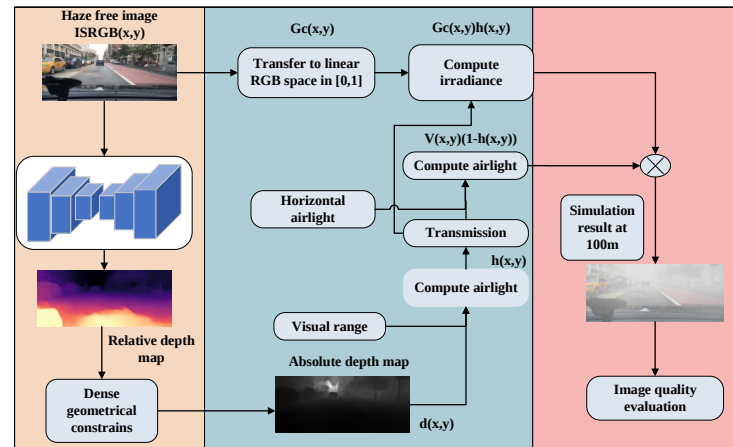


Figure 4. Fog simulation based on Equation (10).

3. Fog Day Simulation Image Feature Evaluation Algorithm

3.1. Introduction to Evaluation Algorithms

As a key information source for human subjective perception and machine processing systems, images play a decisive role in the accuracy and effectiveness of information acquisition. The quality of images greatly affects the speed and accuracy of subsequent algorithm processing, which highlights the importance of introducing image quality assessment. Considering the subjective nature of human vision, people's perceptions of fog images appear to be much the same, but quantitative analysis of the features of simulated fog images becomes extremely challenging. In order to verify the authenticity of fog simulated images, an in-depth analysis using a reference-free image quality assessment algorithm facilitates the quantitative assessment of the performance of fog simulated images.

Therefore, in this paper, the simulated fog image realism assessment tool is used to assess the quality of the simulated images in order to obtain the corresponding quan-

tative assessment indexes. Authenticity Evaluator for Synthetic Foggy/Hazy Images (AuthESI) [34] is a typical natural scene statistics (NSS) feature set based on real fog scenes. It takes the average contrast-normalized luminance coefficient of an image as the basic feature and fits the distribution of mean subtracted contrast normalized (MSCN) coefficients through a generalized Gaussian distribution model. The naturalness of the image is measured using the shape parameters and variance of the fitted Gaussian model. Finally, the quality of the image to be detected is determined using the difference in distance between the parameters of the model characterizing the image to be evaluated and the parameters of the pre-established model. The overall flowchart of AuthESI is shown below:

Figure 5 illustrates a comprehensive flowchart of the selected algorithm, dividing it into two parts: model learning and simulated fog map evaluation. In the model training stage, the distribution of MSCN coefficients for undistorted natural images is approximately Gaussian. Different forms of image distortion will lead to different degrees of distortion in the MSCN coefficient distribution. Taking advantage of this idea, we adopt the generalized Gaussian distribution (GGD) to simulate the distribution of MSCN coefficients. Subsequently, the logarithm of the MSCN coefficients is taken, a step that preserves the correlation between the data attributes and conveniently compresses the variable scales, thus improving the data processing efficiency.

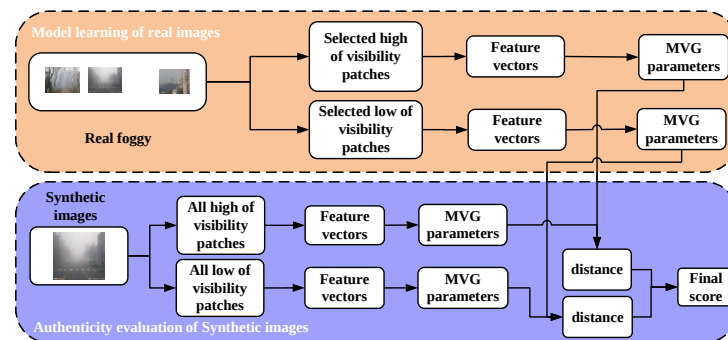


Figure 5. Process flow of AuthESI.

In order to capture the relationship between neighbouring pixels, the logarithmic difference of the MSCN coefficients in the horizontal, vertical, main diagonal, and sub-diagonal directions is derived. Finally, the quality of the evaluated image is determined by measuring the distance between the characteristic model parameters of the target image and the pre-established model parameters.

3.2. Comparison of Evaluation Algorithms

In order to accurately verify and evaluate the performance of simulated fog weather, we use four well-known no-reference image quality assessment (NR-IQA) algorithms to experiment and compare with the algorithm in this paper: BRISQUE [35], BLINDS-II [36], DIIVINE [37], NBIQA [38], and AuthESI [34]. The experiments were executed on the public DHQ database, which employs seven representative dehazing algorithms to generate 1750 dehazed images from 250 real foggy images with varying haze densities. Each image in the database is associated with a corresponding Mean Opinion Score (MOS). Evaluation metrics, including the Pearson Linear Correlation Coefficient (PLCC) and Spearman's Rank-Ordered Correlation Coefficient (SROCC), were employed to assess the performance of the algorithms. PLCC and SROCC measure the correlation between subjective scores and predicted scores, with a range of [0,1]. Higher values indicate better performance and greater alignment with human visual observations. SROCC primarily assesses monotonicity, while PLCC focuses on prediction accuracy. The experiments utilized a subjective quality database of real foggy images, randomly partitioned into 80–20%, 50–50%, and 20–80% training and test sets. The entire training and testing process

was iterated 1000 times, and the average SROCC and PLCC for the five methods were calculated, as illustrated in the Figure 6 below.

Mean SROCC and PLCC values of five NR-IQA methods for 1000 trials on a real fog map database

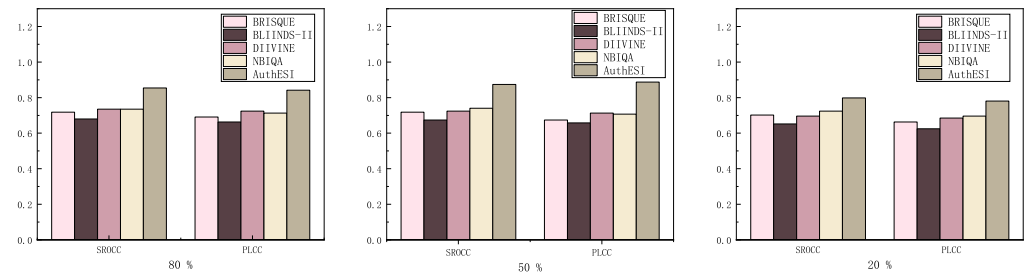


Figure 6. Mean SROCC and PLCC of the five NR-IQA methods across 1000 trials of the DHQ database.

As depicted in Figure 6, the AuthESI algorithm emerges as the top-performing model among all considered methods, affirming its efficacy. The figure illustrates that the objective quality scores predicted by AuthESI align more closely with human vision's perception of image quality in foggy conditions. Through an analysis of the preceding four non-reference image quality evaluation methods, it is evident that, in comparison with AuthESI, their extracted features fall short in capturing the distinctive characteristics of foggy conditions. Consequently, the first four methods are deemed less suitable for assessing the quality of foggy images. Figure 6 demonstrates that the AuthESI algorithm is well-suited for predicting the quality of real foggy images, and its predictions of the authenticity of simulated images exhibit a high degree of consistency with subjective judgments. This assessment tool proves effective in evaluating simulated foggy weather images. In addition, the data show that the algorithm performs relatively better when the training set contains about 50% of the original number of images, suggesting that the algorithm has a strong ability to evaluate foggy images.

AuthESI, the evaluation network for assessing foggy weather consists of two main components: the upper section covers the network training process, and the lower section pertains to the input and detection of simulated fog images. The network training involves utilizing real fog images, and for this purpose, the network is trained on the real fog dataset RTTS. The RTTS dataset, a publicly available real-world foggy day dataset, comprises 4322 real-world foggy images sourced from the Internet, including a substantial collection from traffic and driving cameras. Following dataset screening, 893 images featuring pedestrians, vehicles, and road traffic scenes were selected for training the evaluation network. Upon completing the training, the model's performance was assessed by calculating the PLCC (Pearson Linear Correlation Coefficient) and SRCC (Spearman's rank-Ordered Correlation Coefficient) values, which amounted to 0.8124 and 0.8414, respectively. These elevated correlation values signify that AuthESI's predictions regarding the authenticity of simulated images align more closely with subjective judgments.

4. Discussion of Experimental Process

This section provides a concise and precise description of the experimental results, their interpretation, and the experimental conclusions that can be drawn.

4.1. Experimental Dataset and Experimental Environment

The BDD100K dataset stands as the most expansive and diverse publicly available autonomous driving dataset to date, released by the Berkeley University AI Laboratory. Comprising 100,000 high-definition videos, each spanning approximately 40 s at 1280×720 resolution and 30 frames per second, this dataset serves as a valuable resource for autonomous driving research.

For our study, we selectively sample frames from the 10th second of each video, resulting in 100,000 high-resolution images (1280×720) capturing complex traffic scenes. The dataset covers 10 distinct tasks, which effectively amount to 13 classification categories when considering the nuanced distinction of traffic light colours. This multifaceted design aims to facilitate the comprehensive evaluation of the advancements in autonomous driving image recognition algorithms.

Noteworthy for its geographical, environmental, and weather diversity, the BDD100K dataset enables model evaluation across a spectrum of scenarios, fostering robust generalization capabilities. This diverse array of conditions ensures that recognition models are trained to effectively navigate and understand a wide range of real-world driving situations.

4.2. Fog Image Simulation Process Analysis

The experimental results are shown below. The image sequence from top to bottom includes the original image, relative depth map, absolute depth map, and the corresponding simulated fog map. The relative depth map uses colours of different brightness to represent different scene depths in the real world. Brighter colours indicate objects that are relatively close to the camera, while darker colours indicate objects that are farther away from the camera. In the absolute depth map, a greyscale scheme is used to represent the absolute depth of the scene, with darker colours indicating objects closer to the camera and lighter colours indicating objects farther away. Subsequently, an algorithm based on the atmospheric scattering model is applied to simulate the corresponding fog scene using the absolute depth information. The results show that the algorithm proposed in this study for simulating fog images achieves excellent simulation results.

As shown in Figure 7b,c, it can be seen from the experimental procedure that this depth estimation method successfully generates a depth map with the same size as the original image. The depth maps contain valuable depth information that accurately reflects the scene in the image, especially at the edges where depth changes are observed. This information is critical for integration with atmospheric scattering models. However, depth map results from monocular depth estimation have some limitations. The edge depth information in Figure 7b relative depth map is imprecise. The absolute depth map Figure 7c, obtained after scale recovery from Figure 7b, shows unexpected depth variations in regions that are expected to be stable. In addition, in some regions of the sky, the scale recovery is imprecise, as shown by the ellipse in Figure 7c, resulting in inaccurate prediction of depth information. These factors reduce the success of the algorithm to a large extent. The evaluation score of the fog map is shown in the upper right corner of Figure 7d, which is intuitively better while scoring lower.



Figure 7. Examples of depth estimation and foggy image synthesis, (a) original images from BDD100K, (b) relative depth maps predicted by Monodepth2, (c) absolute depth maps obtained by dense geometric constraints, and (d) synthetic fog (evaluation of the corresponding fog image in the upper right corner of the image).

Based on the experimental results, 69,863 clear images in the BDD100K dataset were processed for fog simulation, and 16,702 fog-simulated images were generated with significantly improved results. Although the depth map exhibits some defects, these defects do not explicitly affect the final simulation results of the fog map. However, given the low generation rate, there is a need to improve the depth map algorithm.

4.3. Comparison of Simulated Foggy Images under Different Visibility Conditions

Visibility serves as a fundamental meteorological parameter, offering a crucial metric for gauging atmospheric transparency. It signifies the maximum distance at which an observed object can be discerned from its surroundings. The varying levels of visibility correspond to distinct concentrations of fog, as outlined in the Table 1 below.

Table 1. The relationship between road visibility and fog concentration.

Category	Dense Fog	Heavy Fog	Moderate Fog	Light Mist
Visibility /km	<0.05	0.05–0.1	0.1–0.2	>0.2

The fog simulation is performed according to the above relationship between visibility and fog concentration. As shown in Figure 8 below, when the visibility is less than 50 m, it is dense fog; when the visibility is between 50 m and 100 m, it is heavy fog; when the visibility is between 100 m and 200 m, it is moderate fog; and when the visibility is light fog, the visibility is greater than 200 m. They correspond to the following figures.



Figure 8. Simulated fog images with scores at three different densities and examples of clear images, with fog thickness decreasing as visibility increases.

We select all images in normal scenes as the original input for foggy image generation and process them separately with different foggy densities. To enhance the diversity and complexity of fog density in real fog scenes, we try to simulate fog scenes more realistically and adapt the dataset to multi-density fog scenes. Moreover, it was concluded through experiments that the visibility in foggy days should be limited to a reasonable range, otherwise the artificially generated fog will be almost invisible or too thick to achieve the desired effect.

4.4. Night-Time Fog Image Simulation Assessment

Regarding the simulation of nighttime hazy scenes, there are still a number of scholars proposing research methods. Jin et al. [39] proposed a method to accurately simulate nighttime haze by using a 3D scattering formula. Zhang et al. [40] proposed a method to firstly reconstruct the scene geometry then simulate the light and object reflectivity and

finally render the haze effect. On the basis of the above, a method to generate realistic night haze images by sampling real-world light colours from a priori distributions was created.

Unlike others, our proposed simulated night fog images are based on self-supervised depth estimation, which is analysed as follows.

As can be observed in Figure 9, in the original nighttime image, nightfall causes the road, the surrounding buildings, and the sky to merge into one, which results in a significant reduction in visibility and a drastic decrease in object recognition. Because of this, there are multiple prediction failures in the generated relative depth map, as shown by the blue boxes in Figure 9. Further observation of the absolute depth map reveals that the objects in the image are almost unrecognisable, which clearly demonstrates the poor generation of the relative depth map, which directly affects the recovery of the image scale and has an impact on the generation of the depth map. The final simulated fog map and its quality score are shown in the upper right corner of the image. It is worth noting that the AuthESI scores of the images are generally high, which indicates that the NSS features of the nighttime fog differ significantly from those of the natural fog. This shows that the algorithm in this paper still has much room for improvement for the simulation of fog images at night.

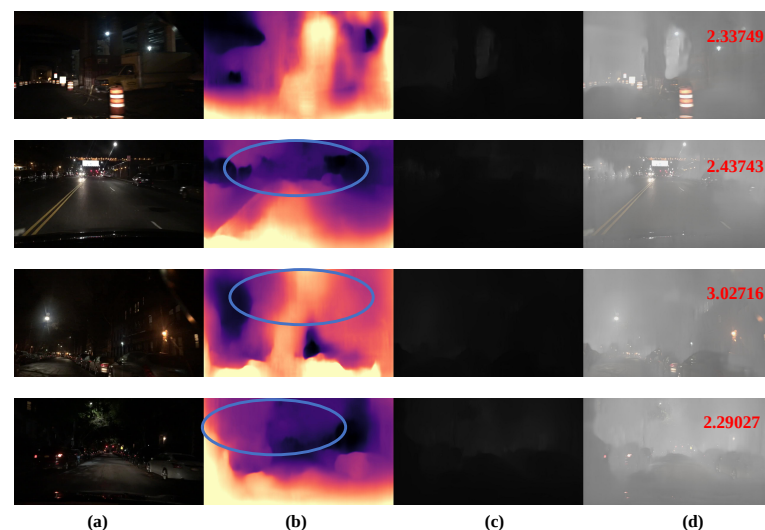


Figure 9. Example of nighttime fog image synthesis, (a) original nighttime image at BDD100K, (b) relative depth map (the parts circled in blue are significantly less effective relative to the depth map), (c) absolute depth map, and (d) synthesised fog (quality scores of the corresponding fog images are shown in the upper right corner of the image).

4.5. Evaluation of Foggy Simulated Image Features

In this section, we apply the foggy day simulation framework on the BDD100K dataset and then utilize an image evaluation algorithm to further assess the quality of the foggy images and verify the feasibility of the current foggy image simulation framework. We synthesized a total of 16,702 haze images with improved results from normal scenes in the BDD100K dataset. Given the large amount of data, we employed a Nightingale rose diagram to characterize the quantitative distribution of the data. Originally used to represent seasonal mortality data in military hospitals, the Nightingale rose diagram, also known as a polar plot, resembles the structure of a rose and is adept at conveying information about large amounts of data. The angle of each line segment to the centre of the circle remains constant, while the radius varies depending on the size of the data. In addition, each area can be coloured differently to indicate different data attributes. Essentially, the Nightingale rose chart converts the traditional bar chart into a visually appealing pie chart format. Unlike traditional pie charts, which use angles to represent values or percentages, the Nightingale rose chart uses the radius of a sector to represent

the size of the data while using colours to indicate the type of attribute, with each sector maintaining a consistent angle.

We arranged the 16,702 data in descending order, divided them into 60 equal intervals, and then found the percentage of the total number of images in each interval. Finally, we obtained the Nightingale rose diagram shown in Figure 10, which distributes the data from the starting point and completes the combination of the data in a clockwise direction. The graph presents the percentage distribution of images within each scoring interval. The sum of the percentages of data in each interval is 100%. Given the restricted space in the figure and the fact that smaller data had little impact on the analysis of the results, only distributions where the number percentage exceeded 3.5% are presented.

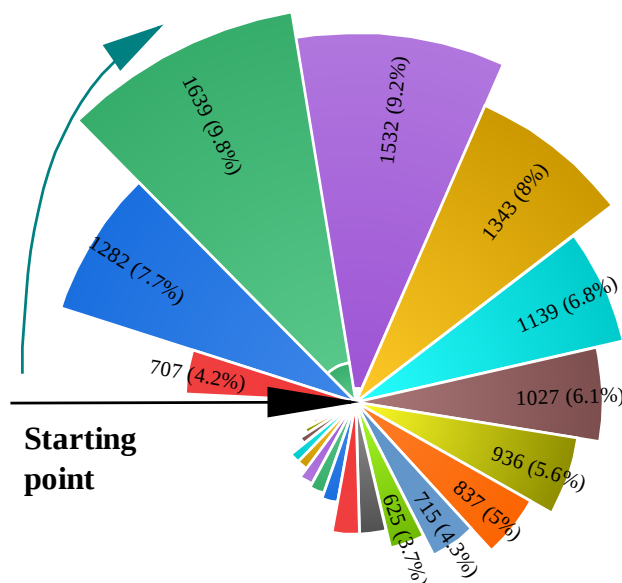


Figure 10. Visibility is the percentage distribution of the number of 100-meter simulated fog plots across the different scoring intervals (only cases where the percentage of the number of distributions is greater than 3.5 percent are shown)

Figure 10 shows the combined scoring results of the algorithm applied to simulated images with a visibility range of 100 m. As can be seen from the Figure 11, the AuthESI values for all images are concentrated in small intervals and the overall values are generally low, with more than 50% of the images scoring below 1.25. This observation suggests that the NSS characteristics of the algorithm are close to those of natural fog.

Evaluation of fog images using the AuthESI metric showed a clear trend towards lower scores overall. Specifically, more than 50% of the evaluated images had scores of no more than 1. This suggests that the scores for the simulated fog images were generally low. In addition, the vast majority (over 90%) of the image scores showed a gradient, indicating a concentration of scores and a relatively low misclassification rate.

After analysing the data with AuthESI scores over 2, we found 1122 images, of which night images accounted for 1001. Combined with the in-depth discussion of the problem in Section 4.4 and the fog map simulation process in Section 4.3, we can comprehensively summarise the reasons for the reduced success rate of the foggy simulation framework proposed in this paper. Firstly, the reduced visibility of the nighttime image leads to inaccuracy in the relative depth estimation, and this inaccuracy has a knock-on effect in the subsequent simulation process, which ultimately affects the quality of the simulated image. Secondly, when processing the sky region of the image, we observed that the relative depth images of the sky region did not have enough contrast to be significant compared to other regions. This lack of contrast has a significant impact in the scale recovery of the absolute depth, which affects the recovery and ultimately leads to the degradation of the fog image quality. Therefore, we believe that the robustness of the relative depth estimation

network and the scale recovery network still needs to be further improved to cope with the challenges in complex environments. This consistency in scoring highlights the reliability and stability of the estimation process.

In summary, the above analysis shows that the fog image simulation framework is effective and practical for image simulation under clear weather.

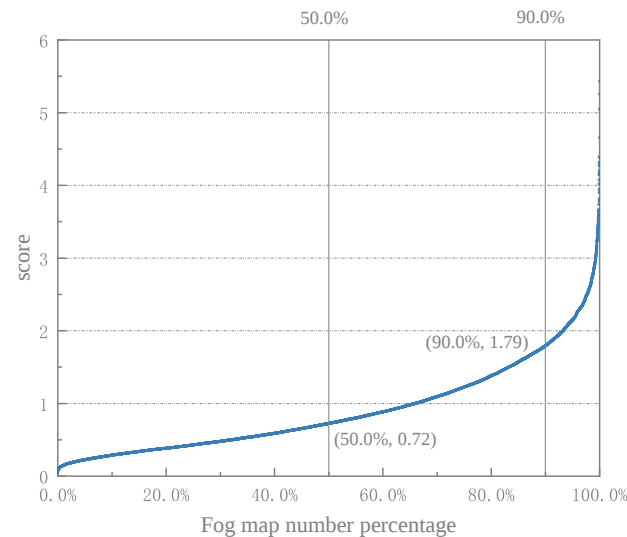


Figure 11. The curve of the data quantity percentage of foggy simulated images changing with the score (visibility 100 m).

5. Conclusions

This study focuses on extracting depth information from monocular clear weather images using a self-supervised monocular depth estimation method and then synthesising foggy weather images based on an atmospheric scattering model. An enhancement algorithm is proposed for generating foggy day image datasets from clear weather conditions.

A depth estimation network is constructed based on a self-supervised monocular depth estimation algorithm to obtain depth maps for the problem of the scarcity of datasets under multi-foggy weather conditions. An atmospheric scattering model is used to estimate transmittance maps from visibility, while dark channel maps are used to distinguish sky regions to estimate atmospheric light values. This approach combines the atmospheric scattering model and the depth maps generated by the depth estimation network to help simulate foggy weather scenarios. The BDD100K dataset was utilised to generate 16,702 (300 m visibility) fog images, which effectively solves the problem of insufficient target detection dataset in foggy weather. As a result, we built a novel foggy weather target detection dataset that can contain three foggy weather scenes with different densities.

In order to verify and assess the authenticity of the synthetic fog images, a non-reference image quality assessment algorithm, AuthESI, is employed. The assessment results show that the algorithm is generally consistent with the features of the natural scene fog data, proving the effectiveness of the proposed algorithm. The practical significance of the algorithm is noteworthy as it introduces a monocular depth estimation network for scene depth estimation without the need for additional sensors, a cost-effective solution for vision-based target detection datasets. Although this study focuses on target detection in foggy weather, it is expected that in future research the method can be extended to tasks such as object recognition and semantic segmentation in other complex scenes, thus broadening its applicability.

Author Contributions: Conceptualization, M.T.; data curation, Z.Z.; formal analysis, Z.Z.; funding acquisition, M.T.; investigation, Z.Z. and J.Q.; methodology, M.T., Z.Z. and J.Q.; project administration, M.T.; resources, M.T.; software, Z.Z.; supervision, M.T. and J.Q.; validation, J.Q.; visualization, Z.Z. and J.Q.; writing and original draft preparation, M.T. and Z.Z.; writing and review and editing, M.T. and J.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant numbers 62363022, 61663021, 71763025, and 61861025); Natural Science Foundation of Gansu Province (grant number 23JRRA886); Gansu Provincial Department of Education: Industrial Support Plan Project (grant number 2023CYZC-35).

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: This manuscript has not been published or presented elsewhere in part or in its entirety and is not under consideration by any other journal. We have read and understood your journal's policies, and we believe that neither the manuscript nor the study violates any of these. There are no conflicts of interest to declare.

References

1. Codur, M.Y.; Kaplan, N.H. Increasing the visibility of traffic signs in foggy weather. *Fresenius Environ. Bull.* **2019**, *28*, 705–709.
2. Schechner, Y.Y.; Narasimhan, S.G.; Nayar, S.K. Instant dehazing of images using polarization. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001, Kauai, HI, USA, 8–14 December 2001; p. 1.
3. Fattal, R. Single image dehazing. *ACM Trans. Graph.* **2008**, *27*, 1–9. [\[CrossRef\]](#)
4. Fattal, R. Dehazing using color-lines. *ACM Trans. Graph.* **2014**, *34*, 1–14. [\[CrossRef\]](#)
5. Berman, D.; Avidan, S. Non-local image dehazing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1674–1682.
6. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *33*, 2341–2353.
7. Ancuti, C.O.; Ancuti, C. Single image dehazing by multi-scale fusion. *IEEE Trans. Image Process.* **2013**, *22*, 3271–3282. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Zhu, Q.; Mai, J.; Shao, L. A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans. Image Process.* **2015**, *24*, 3522–3533.
9. Liu, X.; Lin, Y. YOLO-GW: Quickly and Accurately Detecting Pedestrians in a Foggy Traffic Environment. *Sensors* **2023**, *23*, 5539. [\[CrossRef\]](#)
10. Geiger, A.; Lenz, P.; Urtasun, R. Are we ready for autonomous driving? The KITTI vision benchmark suite. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 3354–3361.
11. Cordts, M.; Omran, M.; Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; Schiele, B. The cityscapes dataset for semantic urban scene understanding. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 3213–3223.
12. Yu, F.; Chen, H.; Wang, X.; Xian, W.; Chen, Y.; Liu, F.; Madhavan, V.; Darrell, T. Bdd100k: A diverse driving dataset for heterogeneous multitask learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 2636–2645.
13. Hu, X.; Fu, C.-W.; Zhu, L.; Heng, P.-A. Depth-attentional features for single-image rain removal. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 8022–8031.
14. Liu, T.; Chen, Z.; Yang, Y.; Wu, Z.; Li, H. Lane detection in low-light conditions using an efficient data enhancement: Light conditions style transfer. In Proceedings of the 2020 IEEE Intelligent Vehicles Symposium (IV), Las Vegas, NV, USA, 19 October–13 November 2020; pp. 1394–1399.
15. Nie, X.; Xu, Z.; Zhang, W.; Dong, X.; Liu, N.; Chen, Y. Foggy Lane Dataset Synthesized from Monocular Images for Lane Detection Algorithms. *Sensors* **2022**, *22*, 5210. [\[CrossRef\]](#)
16. Tarel, J.-P.; Hautiere, N.; Cord, A.; Gruyer, D.; Halmaoui, H. Improved visibility of road scene images under heterogeneous fog. In Proceedings of the 2010 IEEE Intelligent Vehicles Symposium, La Jolla, CA, USA, 21–24 June 2010; pp. 478–485.
17. Tarel, J.-P.; Hautiere, N.; Caraffa, L.; Cord, A.; Halmaoui, H.; Gruyer, D. Vision enhancement in homogeneous and heterogeneous fog. *IEEE Intell. Transp. Syst. Mag.* **2012**, *4*, 6–20. [\[CrossRef\]](#)
18. Sakaridis, C.; Dai, D.; Van Gool, L. Semantic foggy scene understanding with synthetic data. *Int. J. Comput. Vis.* **2018**, *126*, 973–992. [\[CrossRef\]](#)
19. Mai, N.A.M.; Duthon, P.; Khoudour, L.; Crouzil, A.; Velastin, S.A. 3D Object Detection with SLS-Fusion Network in Foggy Weather Conditions. *Sensors* **2021**, *21*, 6711. [\[CrossRef\]](#) [\[PubMed\]](#)

20. Bartoccioni, F.; Zablocki, É.; Pérez, P.; Cord, M.; Alahari, K. LiDARTouch: Monocular metric depth estimation with a few-beam LiDAR. *Comput. Vis. Image Underst.* **2023**, *227*, 103601. [[CrossRef](#)]
21. Eigen, D.; Puhrsch, C.; Fergus, R. Depth Map Prediction from a Single Image using a Multi-Scale Deep Network. *arXiv* **2014**, arXiv:1406.2283.
22. Wong, A.; Fei, X.; Tsuei, S.; Soatto, S. Unsupervised Depth Completion from Visual Inertial Odometry. *IEEE Robot. Autom. Lett.* **2008**, *5*, 1899–1906. [[CrossRef](#)]
23. Seo, B.-S.; Park, B.; Choi, H. Sensing Range Extension for Short-Baseline Stereo Camera Using Monocular Depth Estimation. *Sensors* **2022**, *22*, 4605. [[CrossRef](#)]
24. Mo, H.; Li, B.; Shi, W.; Zhang, X. Cross-based dense depth estimation by fusing stereo vision with measured sparse depth. *Vis. Comput.* **2022**, *39*, 4339–4350. [[CrossRef](#)]
25. Fu, S.; Safaei, F.; Li, W. Optimization of Camera Arrangement Using Correspondence Field to Improve Depth Estimation. *IEEE Trans. Image Process.* **2017**, *26*, 3038–3050. [[CrossRef](#)]
26. Jang, W.-S.; Ho, Y.-S. Disparity Fusion Using Depth and Stereo Cameras for Accurate Stereo Correspondence. In Proceedings of the Three-Dimensional Image Processing, Measurement (3DIPM), and Applications, SPIE/IS&T Electronic Imaging, San Francisco, CA, USA, 8–12 February 2015; Volume 9393, pp. 227–234.
27. Ding, X.; Xu, L.; Wang, H.; Wang, X.; Lv, G. Stereo depth estimation under different camera calibration and alignment errors. *Appl. Opt.* **2011**, *50*, 1289–1310. [[CrossRef](#)]
28. Watson, J.; Aodha, O.M.; Prisacariu, V.; Brostow, G.; Firman, M. The Temporal Opportunist: Self-Supervised Multi-Frame Monocular Depth. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 1164–1174.
29. Xue, F.; Zhuo, G.; Huang, Z.; Fu, W.; Wu, Z.; Ang, M.H. Toward Hierarchical Self-Supervised Monocular Absolute Depth Estimation for Autonomous Driving Applications. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Las Vegas, NV, USA, 24 October 2020–24 January 2021; pp. 2330–2337.
30. Casser, V.; Pirk, S.; Mahjourian, R.; Angelova, A. Depth Prediction without the Sensors: Leveraging Structure for Unsupervised Learning from Monocular Videos. In Proceedings of the AAAI Conference on Artificial Intelligence, Hilton Hawaiian Village, Honolulu, HI, USA, 27 January 2019–1 February 2019; pp. 8001–8008.
31. Yang, Z.; Wang, P.; Wang, Y.; Xu, W.; Nevatia, R. LEGO: Learning Edge with Geometry all at Once by Watching Videos. In Proceedings of the Name of the Conference, Salt Lake City, UT, USA, 18–23 June 2018; pp. 225–234.
32. Godard, C.; Aodha, O.M.; Michael, F.; Gabriel, J.B. Digging Into Self-Supervised Monocular Depth Estimation. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October 2019–2 November; pp. 3828–3838.
33. Zhou, T.; Brown, M.; Snavely, N.; Lowe, D.G. Title of presentation. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1851–1858.
34. Zhang, N.; Zhang, L.; Cheng, Z. Towards simulating foggy and hazy images and evaluating their authenticity. In Proceedings of the Neural Information Processing, Guangzhou, China, 14–18 November 2017; pp. 405–415.
35. Mittal, A.; Moorthy, A.K.; Bovik, A.C. No-Reference Image Quality Assessment in the Spatial Domain. *IEEE Trans. Image Process.* **2012**, *21*, 4695–4708. [[CrossRef](#)]
36. Saad, M.A.; Bovik, A.C.; Charrier, C. Blind Image Quality Assessment: A Natural Scene Statistics Approach in the DCT Domain. *IEEE Trans. Image Process.* **2012**, *21*, 3339–3352. [[CrossRef](#)]
37. Moorthy, A.K.; Bovik, A.C. Blind Image Quality Assessment: From Natural Scene Statistics to Perceptual Quality. *IEEE Trans. Image Process.* **2011**, *20*, 3350–3364. [[CrossRef](#)] [[PubMed](#)]
38. Ou, F.Z.; Wang, Y.G.; Zhu, G. A Novel Blind Image Quality Assessment Method Based on Refined Natural Scene Statistics. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019; pp. 1004–1008.
39. Jin, Z.; Feng, H.; Xu, Z.; Chen, Y. Nighttime Image Dehazing by Render. *J. Imaging* **2023**, *9*, 153. [[CrossRef](#)] [[PubMed](#)]
40. Zhang, J.; Cao, Y.; Zha, Z.-J.; Tao, D. Nighttime Dehazing with a Synthetic Benchmark. In Proceedings of the 28th ACM International Conference on Multimedia, New York, NY, USA, 10 August 2020; pp. 2355–2363.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.