DEVELOPMENT OF AN EFFICIENT DEFOGGING FRAMEWORK FOR ROAD SAFETY

A thesis submitted in partial fulfilment of the requirement
for the award of degree of

Doctor of Philosophy
In
Electrical and Instrumentation Engineering Department

by

AKSHAY JUNEJA

(902004020)

under the guidance of

Dr. Sunil Kumar Singla

Professor, Electrical and Instrumentation Engineering Department,

Thapar Institute of Engineering and Technology, Patiala, Punjab, India

Dr. Vijay Kumar

Associate Professor, Information Technology Department,

Dr. B.R. Ambedkar National Institute of Technology, Jalandhar, Punjab, India

ELECTRICAL AND INSTRUMENTATION ENGINEERING DEPARTMENT
THAPAR INSTITUTE OF ENGINEERING AND TECHNOLOGY
(Deemed-to-be-University)
PATIALA – 147004 (PUNJAB)

July 2024

DECLARATION

I, hereby declare that the work being presented in the thesis entitled "**DEVELOPMENT OF AN EFFICIENT DEFOGGING FRAMEWORK FOR ROAD SAFETY**" in partial fulfilment of the requirement for the award of the degree of **Doctor of Philosophy** submitted in Electrical and Instrumentation Engineering (EIE) department of TIET, Patiala, is an authentic record of my own work carried out under the supervision of **Dr. Sunil Kumar Singla, Professor and Head, EIE Department, TIET, Patiala, Punjab** and **Dr. Vijay Kumar, Associate Professor and Head, IT Department, Dr. B. R. Ambedkar NIT, Jalandhar, Punjab**, and refers other researchers' work which are duly listed in the reference section.

The matter presented in this thesis has not been submitted in any other University/Institute for the award of any degree in India or abroad.

Place: Patiala

Date: 01st July, 2024

(Akshay Juneja)

(Registration Number: 904002040)

It is certified that the above statement made by the candidate is correct to the best of my knowledge and belief.

(Dr. Sunil Kumar Singla)
Professor and Head
EIE Department
TIET, Patiala, Punjab

(Dr. Vijay Kumar)
Associate Professor and Head
IT Department
NIT Jalandhar, Punjab

CERTIFICATE

This is to certify that the thesis entitled "**DEVELOPMENT OF AN EFFICIENT DEFOGGING FRAMEWORK FOR ROAD SAFETY**" being submitted by **Akshay Juneja**, bearing **Registration Number 902004020** at this institute in the partial fulfilment of the requirement for the award of the degree of **Doctor of Philosophy** in Electrical and Instrumentation Engineering (EIE) Department is a record of student's own work carried out under my supervision and guidance. This thesis work is of desired standard and has not been submitted in any other University/Institution for the award of any other degree.

(Dr. Sunil Kumar Singla)
Professor and Head
EIE Department
TIET, Patiala, Punjab

(Dr. Vijay Kumar)
Associate Professor and Head
IT Department
NIT Jalandhar, Punjab

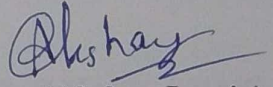
ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisors **Dr. Sunil Kumar Singla**, Professor and Head, EIE Department, TIET, Patiala, Punjab and **Dr. Vijay Kumar,** Associate Professor and Head, IT Department, Dr. B. R. Ambedkar NIT, Jalandhar, Punjab, for their excellent guidance, research attitude, encouragement, and co-operation throughout the course of research work. I consider myself fortunate to have worked under their guidance.

I would like to thank **Dr. Padmakumar Nair**, Director TIET, Patiala, and **Dr. N. Kalaiselvi**, Director Council of Scientific and Industrial Research (CSIR), New Delhi for providing facilities for current research work. This research work is supported by CSIR, New Delhi, Government of India under the scheme of Extra Mural Research - II. The sanction number of the scheme is 22(0801)/19/EMR-II.

I am very obliged to the **Ph.D. Coordinator** of the department **Dr. Mandeep Singh** (Professor, EIE) and doctoral committee members **Dr. Gagandeep Kaur** (Associate Professor, EIE), **Dr. Mandeep Singh – 2** (Associate Professor, EIE), and **Dr. Rinkle Rani** (Professor, Computer Science), for their intellectual support and encouragement throughout the course of this work.

I wish to extend my sincere thanks to my family and friends for their positive encouragement they showered on me throughout this research work.



(Akshay Juneja)

(902004020)

ABSTRACT

In the early morning and the late night of winter season, the majority of road accidents occur every year despite less traffic, due to the poor visibility of drivers in the presence of air pollutants such as fog, haze, smog, etc. By improving the visibility of drivers, precious lives can be saved. The existing visibility restoration algorithms available in the literature suffer from one or more poor qualitative measures such as edge distortion, color distortion, texture distortion, halo effects, excessive lighting, gradient artefacts, etc. Present thesis is an attempt to propose new visibility restoration algorithms and improvement in existing algorithms, for enhancement of drivers' visual range to save human lives. Initially, a new real time roadside database called hazy unpaired dataset for road safety (HUDRS) has been proposed. After the dataset acquisition, four visibility restoration algorithms namely, Wavelet based multiscale convolution neural network (Hybrid CNN); a combination of deep learning architecture called residual regression network (RRNet) and morphological erosion; a combination of gush enhancer-based autoencoder (Aethra-Net) and vessel enhancement-based filter; and a bounding function for gray-world kernel prior (BFGKP), have been proposed for the estimation of image restoration parameters, i.e. atmospheric light and transmission map. A conventional optimization technique called grey wolf optimization has also been utilized for single image dehazing. Moreover, a hardware set-up has been presented to accomplish real-time video dehazing using Aethra-Net. In addition to that, BFGKP has been used to obtain results for real-time dehazing using MATLAB Mobile Application designed by Mathwork, Inc. Various visibility restoration challenges such as computation time, fog density, and parameter tuning have been considered while implementing this framework for real-time application.

Most of the existing hazy datasets such as FRIDA, FRIDA2, O-HAZE, I-HAZE, NH-HAZE, RESIDE, etc., consists of synthetic foggy images produced using either software or artificial fog generating machine. A real-time database called HUDRS has been prepared which consists of foggy and fog-free images captured during morning/evening and day time, respectively. All the images have been captured with manual adjustment of Canon Power Shot SX400 IS camera under natural environmental conditions. It depicts 522 hazy and 1050 haze-free real-time roadside images with a resolution of 4608×3456. Further, each clear image has been processed with haze of different intensities, producing a total of 2088 synthetic hazy images. The performance

of the existing and proposed datasets has been evaluated in terms of several parameters namely, blind/referenceless image spatial quality evaluator ($BRISQUE$), brightness (β), percentage of saturated pixels ($\mathbb{P}$), structure similarity index metric ($SSIM$), perception-based image quality evaluator ($PIQE$), peak signal-to-noise ratio ($PSNR$), and image entropy (IE), by using existing prior-based, metaheuristic-based (i.e., grey-wolf optimization), and deep learning-based dehazing algorithms. Experimental results reveal that they work differently in the real-time scenarios due to the presence of uncertainty such as area of frame and fog density in the real-world.

A novel single image dehazing algorithm called hybrid CNN has been proposed that uses a multi-scale convolution neural network and various filters namely, retinex, wavelet, and inverted wavelet, to process hazy images. The wavelet filter has been used to enhance the features of non-sky regions. Similarly, the inverted-wavelet filter has been used to enhance the features of sky regions. The wavelet-transformed and inverted wavelet-transformed images have been processed with fine-scale network and coarse-scale network, respectively, to obtain two separate feature maps. The airlight has been estimated by using 2D order statistic filter. The combined maps and estimated airlight have been used to produce the optimized transmission map for hazy image restoration. The proposed approach has been found superior to existing dehazing techniques in terms of fog aware density evaluator ($FADE$) and naturnalness image quality evaluator ($NIQE$).

A new software-based video dehazing architecture called Aethra-Net has been proposed. It consists of gush enhancer-based autoencoder and a vessel enhancement-based filter, to estimate the transmission map. The multiple blocks of ResNet-101 have been employed to overcome vanishing gradient problem. The vessel enhancement-based filter has been used to emphasize the objects in tubular structures. The refined Aethra-Net has been obtained by including all significant features from both feature maps (i.e., estimated transmission map and vessel enhancement-based map). The airlight has been estimated using mean filter. The processed frames have been restored in the same sequence as the original video. The performance analysis demonstrates that the proposed Aethra-Net has outperformed the existing techniques in terms of $BRISQUE$ by 15.29%. Moreover, in the proposed system, improved results for $SSIM$ and $PSNR$-based performance metrics have been obtained. The average dehazing time of 0.11 seconds has been achieved along with the updated airlight for each frame.

In another work, a novel bounding function for gray-world kernel prior (BFGKP) has been proposed to obtain the clear images from real-time foggy images. The gray-world kernel prior has been used to estimate transmission map, and 2D order statistic filter has been used to estimate airlight. The flickering effects and edge distortion have been minimized using the proposed algorithm. The proposed BFGKP has been found superior than existing algorithms in 92.59% images when $FADE$ is calculated. Further, BFGKP has been used in a real-time dehazing framework comprising a GoPro Hero 7 camera to capture the images, a display screen or a system to view the output, a Camlink to transfer the frames captured by GoPro in 4k resolution, and a bluerigger chord to set-up connectivity. In addition to that, the proposed algorithm has been used in a cloud computing-based real-time defogging process, using MATLAB mobile application. This application is accessible through the Google play store for android-powered devices. The computation time of the proposed model is 7.68 times less than the existing techniques.

Further, a new algorithm to overcome the effect of smog (fog + smoke) on the roadside images has been proposed. In the proposed algorithm, to estimate the transmission map for single image desmogging, a residual regression network (RRNet) has been designed. In RRNet, the shallow features have been extracted using convolution function, which is followed by reduction in the thickness of smog to find the depth of image. To make the network stable and faster, batch-normalization function has been used. The halo effect has also been reduced in the proposed algorithm. The transmission map has been refined by using morphological erosion operation. The atmospheric light has been estimated using a 2D order statistic filter. A comparative analysis has been drawn that reveals the supremacy of proposed RRNet over existing desmogging techniques in terms of $PIQE$, $BRISQUE$, and $FADE$, by 8.05%, 2.64%, and 2.02%, respectively.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
AECL	Autoencoder-based Contrastive Learning
AME	Absolute Mean Error
BA	Block Artifacts
BFGKP	Bounded Function for Gray-world Kernel Prior algorithm
BPTT	Backpropagation Through Time
$BRISQUE$	Blind/Referenceless Image Spatial Quality Evaluator
BSDS	Berkeley Segmentation Dataset
CAP	Color Attenuation Prior
CCDI	Color Cast Dependent Image
CD	Color Distortion
CDP	Channel Difference Prior
CEP	Color Ellipsoid Prior
CIE	Commission Internationale de l'Éclairage
CNN	Convolution Neural Network
CRNN	Convolution Recurrent Neural Network
CV	Computer Vision
DCP	Dark Channel Prior
ED	Edge Distortion
EL	Excessive Lighting
$FADE$	Fog Aware Density Evaluator
FFA-Net	Feature Fusion Aggregation Network
FRIDA	Foggy Road Image DAtabase
GAN	Generative Adversarial Network
GCANet	Gated Context Aggregation Network
GGF	Gain Gradient Filter
GWO	Grey Wolf Optimization
HE	Halo Effect
ICAP	Improved Color Attenuation Prior
ICP	Illumination Channel Prior
IDE	Image Dehazing and Exposure
IE	Image Entropy
IGBF	Information Gain-based Bilateral Filter

Abbreviation	Description
IIR	Infinite Impulse Response
KITTI	Karlsruhe Institute of Technology and Toyota Technological Institute
LHG	Large Haze Gradients
LSTM	Long Short-Term Memory
MAMLP	Multi Axis Multi-layer Perceptron
MRFID	Multiple Real Fog Image Dataset
MSBDN	Multi-scale Boosted Dehazing Network
MSE	Mean Square Error
NH-HAZE	Non-Homogeneous Haze
NHTSA	National Highway Traffic Safety Administration
$NIQE$	Naturalness Image Quality Evaluator
NLBF	Non-Linear Bounding Function
NLD	Non-local Dehazing
NYU	New York University
OTM-AAL	Optimal Transmission Map and Adaptive Atmospheric Light
$PIQE$	Perception-based Image Quality Evaluator
$PSNR$	Peak Signal-to-Noise Ratio
RC	Reconciling Color
RESIDE	REalistic Single Image DEhazing
REVIDE	REal-world VIdeo DEhazing
RNN	Recurrent Neural Network
RRNet	Residual Regression Network
SA	Scene albedo
SCP	Smoggy Channel Prior
SPD	Structural Patch Decomposition
$SSIM$	Structure Similarity Index Metric
US	United States
VM	Variational Minimization
WF	Weiner Filter
WIMF	Window-based Integrated Means Filter

LIST OF SYMBOLS

Symbol	Description
μm	Micro meter
x, y	Pixel location
I and J	Hazy and Dehazed Images, respectively
t and A	Transmission Map and Global atmospheric light, respectively
β and d	Scattering coefficient and Scene depth, respectively
km	Kilometre
g and p	Ground truth and Reconstructed images, respectively
μ_g and μ_p	Sample means of g and p, respectively
σ_j	Sample cross co-variance
σ_g^2 and σ_p^2	Sample variance of g and p, respectively
c_1 and c_2	Constants
M and N	Length and Breadth of image, respectively
ΔE_{00}	Color deviation
$\Delta L'$, $\Delta C'$, and $\Delta H'$	Lighting, chroma, and hue components of the pair of reference and computed image, respectively
k_L, k_C, and k_H	Parametric factors for lighting, chroma, and hue components, respectively
S_L, S_C, and S_H	Weighting functions for lighting, chroma, and hue components, respectively
R_T	Rotation of different planes of color spaces
C_G	Contrast gain
$\mathbb{P}$	Percentage of saturated pixels
$\mathcal{E}$	Visible edge ratio
$\beta 1$	Brightness

CHAPTER-1
INTRODUCTION

This chapter presents a brief overview of various weather degradation techniques, significance of work, various deep learning techniques and performance evaluation metrics.

1.1 Preamble

The image quality of scenarios captured during poor weather conditions like smog, fog, haze, downpour, and cloud is degraded. The climatic conditions greatly affect the visibility of human beings and machines. The visibility of scene is enormously harmed by haze and fog when contrasted with the other climate conditions. In the present era, there are diverse computer vision algorithms such as object detection [1], agricultural monitoring [2], [3], medical imaging and security [4], [5], early classification of multivariate data [6], image encryption [7], [8], text understanding [9], automated cars [10], and urban navigation [11]. The foremost area of research in computer vision is visibility restoration. The visibility is restored in satellite images using different image processing techniques such as image enhancement [12] and cloud removal [13]. In [14], the feature map of remote sensing images was obtained by fusion of two prior knowledge-based techniques namely, oblique gradient prior and luminance channel prior. It preserved temporal and spatial features, and also retain the color, edges, and texture of the original images.

Under foggy weather conditions, the performance of human and system vision is degraded due to the phenomenon of scattering and absorption. The light traveling from the object towards the observer gets scattered when it passes through water droplets or dust particles. It reduces the direct transmission of light and whiteness effect due to airlight. Researchers have made remarkable progress in transmission map refinement and visibility restoration algorithms in the last few decades [15]. It is achieved by using machine learning techniques such as random forest and support vector machine [16].

The poor visibility while driving any vehicle also leads to the loss of human life in some cases and permanent injury in others. Approximately 4.37 lakh incidents were reported on Indian roadways in 2019; out of which 1.54 lakh resulted in fatalities and 4.39 lakh caused serious injuries having victims with permanent disability [17]. In 2019,

two-wheelers accounted for the majority of road fatalities with 58,747 deaths. Trucks and lorries came in second with 22,637 deaths, followed by automobiles with 21,196 deaths, and buses with 9,192 deaths. In India, 44% of all accident deaths are caused by road accidents. National Highway Traffic Safety Administration (NHTSA) stated that 8,730 people died in road accidents in US in first quarter of 2021. The fatality rate increases by 10.5% as compared to first quarter of 2020. In 2021, there is 9% increase in number of accidents during foggy and misty weather in India [18]. In 2022, the Motor Vehicle Accident Fund Trust settled 87 death and 8 injury claims. The total compensation amount paid is Rs. 1,78,00,000 [19], [20]. Approximately 1.3×106 people are killed in vehicle accidents each year. Nearly 25% of those accidents occur on foggy days in the winter season because fog reduces the range of vision [21]. Therefore, the motivation behind this research is to obtain real-time clear scene radiance so that fatal road accidents can be avoided or reduced, and human lives can be saved.

In the early years, the researchers estimate depth from atmospheric components such as power intensity and sky intensity [22], structure using chromatic decomposition, and dichromatic atmospheric scattering [23]. Thus, several single-image and multi-image defogging/dehazing based methods have been introduced to enhance visibility under undesirable weather conditions.

1.2 Weather Degradation Techniques

In this section, the basic concepts of different weather degradation scenarios are presented (see Figure 1.1) [24]–[30].

1.2.1 Fog

The suspension of water droplets (i.e., size 5 μm) is formed by condensation of water vapors occurring near the earth's surface is called fog [31], [32]. The formation and disappearance of fog depends on the temperature and humidity of the atmosphere. Fog is broadly classified into six categories such as radiation fog, valley fog, upslope fog, freezing fog, advection fog, and coastal fog [31], [32].

Radiation fog occurs in winters when the wind is not blowing. The land becomes cool, and the air holds less moisture that leads to condensation of water. It disappears as the sunlight reaches the surface of Earth. Valley fog is found over the valley. The topographical layout creates a hindrance in its path due to which it lasts for few days.

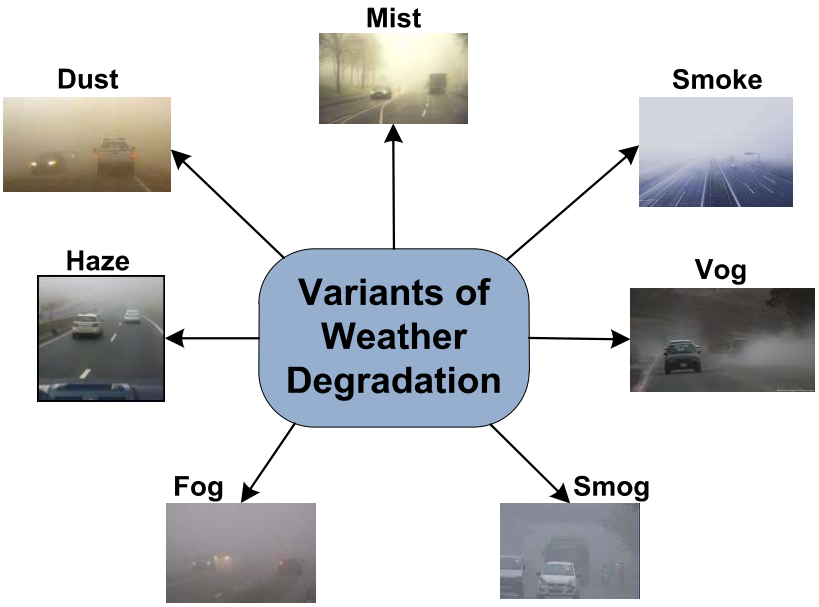


Fig. 1.1 Variants of weather degradation technique

Upslope fog occurs when air flows in an upward direction over rising terrain. It is usually found on hillsides. Freezing fog usually occurs over hot bathtubs and backyard swimming pool when cold air blows over the warm wetlands. Due to this, the ice crystals are formed over the surface. The horizontal movement of air is called advection. Advection fog occurs when air holds the moisture and blows over a cool surface. The cold air causes the condensation of water. Coastal fog is a type of advection fog and found in coastal areas [33]. It usually occurs in the spring or summer seasons when the sea is comparatively colder than land.

The most common visibility restoration technique is defogging. The color of fog gradients is white, which is same as the natural airlight. The main focus is to estimate the atmospheric light and transmission map of the image. Prior based defogging techniques are enough to achieve the task.

1.2.2 Haze

Haze is defined as the obfuscation caused in the atmosphere by the reflection of light from aerosols present. The particle size is 2-10 μm. When sunlight passes through the aerosols, some particles absorb the light while others cause scattering of light before it reaches a person, thus causing haze pollution. Some natural sources of haze pollution are soot from wildfires and dust carried by the wind, while some human-made sources are industrial fuel burning and automobiles. Haze removal techniques equally

concentrate on the estimation of transmission maps and atmospheric light. Deep learning-based architectures are utilized to obtain the desired results.

1.2.3 Smog

Smog is defined as the combination of smoke and fog. It is harmful to the human body as it causes a burning sensation in the eyes and causes throat issues [34]. To obtain clear scene radiance for images captured in the presence of smog gradients, gradient prior techniques are fused with learning-based models to obtain the desired results. A lot of research has been done in the field of fog-removal and haze-removal techniques. Hence, smog-removal techniques are still in their infancy stage.

1.2.4 Other Variants

Dust is a group of solid and liquid particles of diameter 1μm to 30μm, such as ash, pollen, or sandstorm, suspended in the air because of wind erosion or other mechanical processes [31]. Based on the particle size, the dust particles are classified into primary and secondary. The primary particles are generated from soil erosion, disintegration, and combustion. The secondary particles form industrial activities, automobiles, and thermal power plants [35]. The suspension of liquid droplets or solid particles in any gas is known as an aerosol. Larger-sized particles settle in a few hours. While smaller-sized particles stay in the atmosphere, the aerosols help in the formation of clouds in the sky and the absorption or scattering of sunlight [36].

Mist is lighter than fog. It is the suspension of liquid droplets in a gas at or near the Earth's surface. These droplets condense to form clouds when there is a change in temperature and humidity.

Smoke is defined as the presence of fine particles due to the combustion of coal and other substances. When chemicals and plastic products are burned, harmful cancer-causing pollutants are released into the atmosphere. Similarly, highly toxic substances such as hydrogen cyanide, ammonia, and acrolein are released when furniture is burned [37].

Vog is defined as pollution occurring due to volcanic eruptions, and releasing harmful gases [32]. It consists of sulphur dioxide that pollutes the air and hydrogen sulphide, which pollutes the water when lava reaches the sea [38].

The comparison of the above discussed variants is summarized in Table 1.1 [33].

1.3 Fog Imaging Model

Once the air pollutant is detected, the next task is to process the image to eliminate it and make the scene clear. The generalized procedure to remove haze, fog, or smog from an image is almost identical. The optical model is defined by various researchers in computer vision problems using the Koschmieder atmospheric scattering model [39] (see Figure 1.2). The mathematical formulation of the optical model is given as follows:

$$I(x,y) = J(x,y)t(x,y) + A[1 - t(x,y)] \tag{1.1}$$

where $I(x,y)$ represents the hazy input image. $J(x,y)$ is the clear scene radiance. The global atmospheric light in all the three channels, i.e., RGB is represented by A. It signifies the intensity of light. The part of the light that is not scattered during the transmission is represented by $t(x,y)$.

The transmitted map $t(x,y)$ is defined as:

$$t(x,y) = exp[-\beta(d(x,y)] \tag{1.2}$$

where β is the scattering coefficient of atmosphere, and $d(x,y)$ is the depth of the object present in the image.

The visibility restoration is achieved using an estimated airlight and a refined transmission map. Wang et al. [40] proposed gray projection-based airlight estimation method in which pixels were counted in the segmented sky region. Both horizontal and vertical crop operations are performed on grayscale images. The mean value of the 10% pixels with the largest intensities in the cropped region is treated as the airlight. Similarly, the transmission map affects the visibility of the recovered image. The two physical parameters that affect the optimization of the map are fog density and scene depth [41]. In [42], the color balance and saturation of defogged images were modified with variations in the sky segmented region of the optimized transmission map.

1.4 Significance of Computer Vision in Weather Degradation

Computer vision (CV) techniques are proposed by various researchers for the visibility restoration of weather-degraded images. The distorted images are captured in different illumination conditions by using various image capture equipment, such as stereo camera, visible light polarization camera, and single shot detector.

Table 1.1 Comparison of different weather degradation variants

Technique	Source	Composition	Particle Size (in µm)	Visibility Distance	Color
Dust	Wind erosion	Ash + Pollen + Sandstorm	1 - 30	-	Grey
Fog	Condensation of Water vapours	Moisture	5	$\leq 1km$	White
Mist	Condensation of Water vapours	Moisture	70 - 350	$> 1km$	Grey + Blue
Haze	Absorption and scattering of sunlight by tiny pollution particles	Dust + Smoke + Moisture	2-10	$> 1km$	White + Blue
Smoke	Fires, cooking devices, pest control devices, cigarettes Condensation of Water vapours	Carbon + Tar + Oil + Ash	0.01-50	-	Black
Smog	Automobiles, pollution from power plant	Smoke + Fog	1	$\leq 1km$	Black
Vog	Volcanic Eruption	Sulfer dioxide + Hydrogen sulfide	2.5	-	Yellow

Images captured with different cameras produce different results, even if the same algorithm is implemented. Some specifications of the camera, such as focus and viewing geometry are adjusted whenever a new site is selected. The CV algorithms can be used to adjust the focus, saturation, lighting, and other parameters of the images. It is difficult to analyze an image captured in distorted weather conditions. Sometimes there is an unavoidable resemblance in the weather conditions such as fog, haze, and smog, which makes it difficult to determine the air pollutant with the human eye. These techniques can be used to extract the features of the image and classify them into different weather conditions. For instance, researchers proposed deep learning architectures to detect smoke in outdoor images [43], [44].

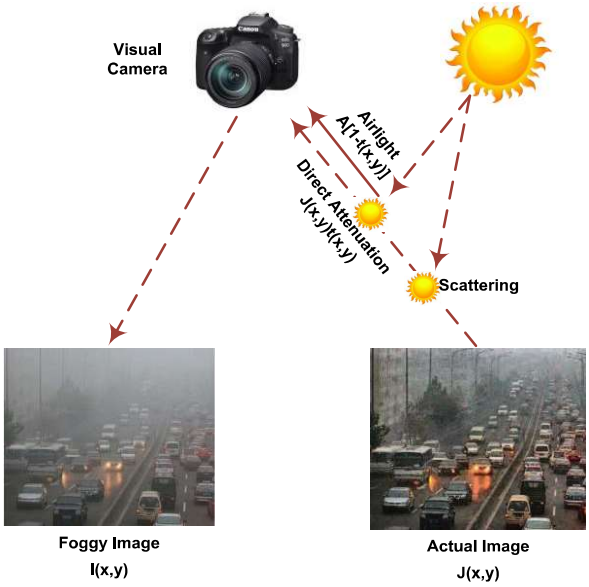


Fig. 1.2 Fog imaging model

In poor weather conditions, the atmospheric light that reaches the surface of Earth is mainly scattered by vapours in the atmosphere. So, it is difficult to distinguish between the sky and non-sky regions. Also, the objects are mainly covered in white light, which makes it difficult to detect them. This drawback can be overcome by using object detection-based CV algorithms. The image processing techniques are extended to be utilized for video processing tasks such as object tracking and real-time processing. These techniques include frame extraction, processing, and video reconstruction. For instance, the object detection algorithm is implemented on sequentially extracted frames to achieve object tracking. In the past few years, the researchers have shown interest in video processing techniques for real-world applications.

1.5 Classification of Computer Vision Techniques

The computer vision techniques for image dehazing are classified into three categories namely, prior-based techniques, deep learning-based techniques, and hybrid techniques.

1.5.1 Prior-based Technique

Prior-based methods are employed by utilizing the inherent characteristics of images, including pixel intensity, color distortion, structure, and sharpness. It is categorized as an image enhancement, transmission refinement, and image fusion techniques.

The image enhancement techniques utilize only one filter to improve atleast one of the characteristics of the image, i.e., edges, color, or texture. In the transmission refinement techniques, the estimated map is optimized to restore the degraded image. The image fusion techniques perform multiple iterations to optimize the dehazed results.

1.5.2 Deep Learning Technique

The multi-layer perceptron algorithm is only applicable to the 1D dataset. When the 2D dataset is used, machine learning algorithms are implemented to convert it to vector form for feature extraction. These are applicable to a small dataset. When the dataset is large, deep learning techniques are implemented to improve accuracy. The size of the dataset is increased using augmentation techniques. When the model is trained using deep learning techniques, it extracts the features automatically. Deep learning techniques such as convolutional neural network (CNN) and generative adversarial network (GAN) are commonly used for dehazing tasks [45]. The well-known deep learning techniques, as depicted in Figure 1.3, are discussed in subsequent subsections.

- **Convolution Neural Network**

A convolution neural network (CNN) is a feed-forward artificial neural network used for various computer vision tasks. It is a type of supervised learning. CNN architecture consists of a simple convolution layer, a pooling layer, and a fully connected layer. Various image filters are used to extract spatial features from the input image using the convolution layer. The dimension of the image is reduced using the pooling layer. It can be either max pooling or average pooling. Finally, classification is performed by the fully connected layer [46]–[50]. The goal is to estimate atmospheric light and transmission map of the image using dehazing algorithms. A CNN architecture is classified as convolution-based network, end-to-end network, or an autoencoder.

A convolution-based network comprises the permutation and combination of several layers to achieve the desired image processing tasks. The autoencoders are used to reduce computational time and decrease memory storage. The various stages of the working of autoencoders are downsampling, processing, and upsampling, followed by image recovery [38]. If the output obtained from a convolution network or an autoencoder does not require further processing, it is said to be an end-to-end network. Figure 1.4 represents a generalized CNN architecture.

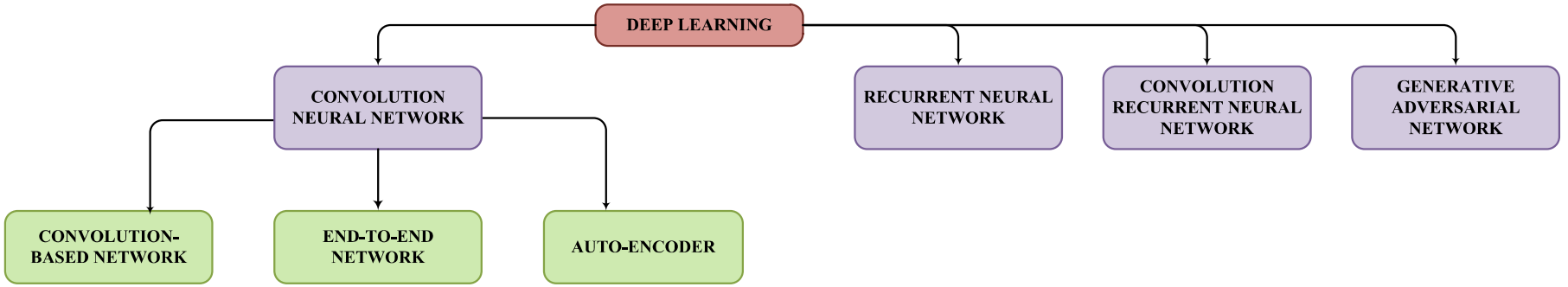


Fig. 1.3 Types of deep-learning architectures

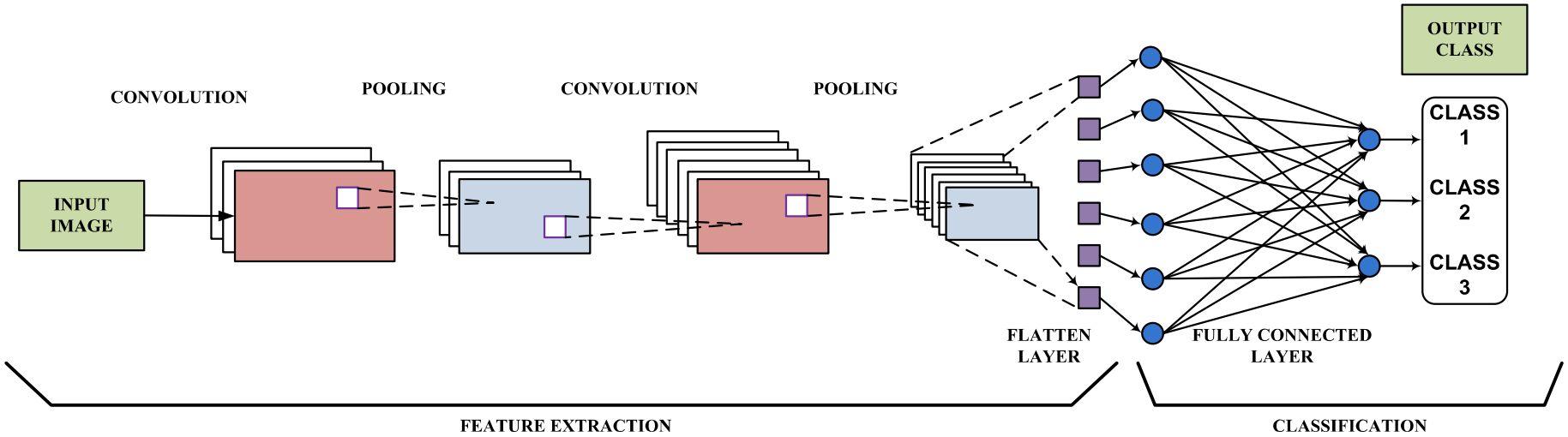


Fig. 1.4 Convolution neural network architecture

- **Recurrent Neural Network**

In a recurrent neural network (RNN), the classification system consists of long short-term memory (LSTM). This network uses backpropagation through time (BPTT) algorithm for learning. The spatial variant property did not allow CNN to use small size models and time-sequential data. Jiang et al. [51] proposed the real-time image defogging algorithm based on RNN. This network utilizes supervised learning to present a relationship between the extracted features and scene depth. The atmospheric scattering model is used to restore the fog image (see Figure 1.5).

- **Convolution Recurrent Neural Network**

The network comprising both CNN and RNN is known as convolution recurrent neural network (CRNN). In this network, future and past computations are dependent on each other. CRNN is trained using gradient backpropagation similar to a feedforward neural network. Oleg et al. [52] proposed CNN to remove non-uniform noise from thermal images, which is estimated using an infinite impulse response (IIR) filter and weighting parameters. Figure 1.6 represents the architecture of the CRNN network.

- **Generative Adversarial Network**

Generative adversarial network (GAN) is an unsupervised learning technique. It consists of a generator and discriminator couple, which are trained to check if the input

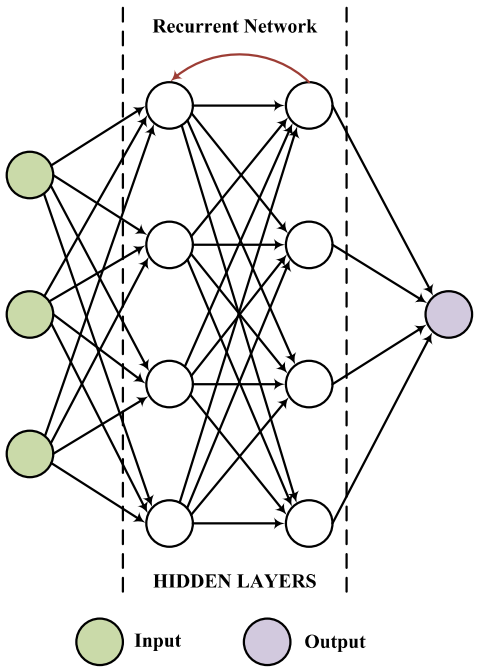


Fig. 1.5 Recurrent neural network architecture

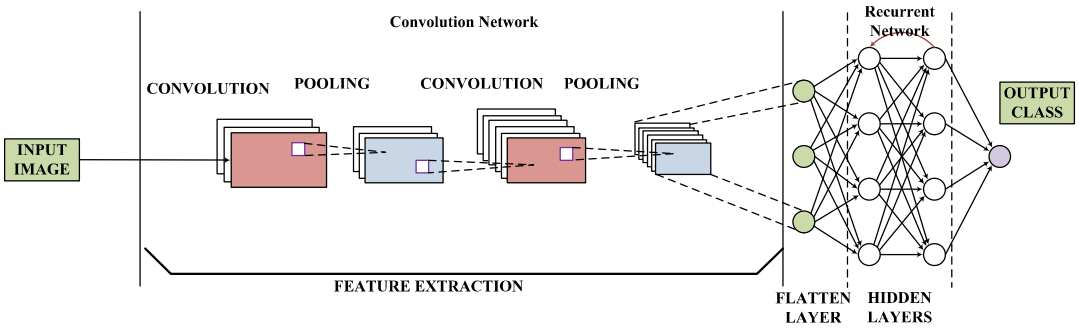


Fig. 1.6 Convolution recurrent neural network architecture

image is real or fake. Here, depth maps are not available, and the output image suffers from color distortion and artifacts. There has been recent growth in the implementation of GAN, such as neural style transfer techniques [53], [54]. GAN works on features such as color, texture, and structure of the objects present in the image to obtain better defogging, dehazing, and desmogging results. Figure 1.7 represents a block diagram of GAN network.

1.5.3 Hybrid Technique

The hybrid techniques consist of both prior-based and learning-based methodologies. They enhance the image characteristics and extract the features to estimate the feature map.

1.6 Applications of Defogging Techniques

Defogging algorithms are important aspects of computer vision applications. This section discusses various uses of visibility restoration-based techniques that utilize defogging algorithms as a pre-processing tool, as depicted in Figure 1.8 [55].

1.6.1 Remote Sensing

The hyperspectral images are of poor quality. As a result, they are not used in detection and recognition applications such as agriculture or mining [56], [57]. Hence, defogging algorithms are applied to high-quality remote sensing images [58], [59].

1.6.2 Intelligent Transportation System

Pollutants in the air, such as fog, mist, haze, smoke, and smog, make it difficult for humans to see far away. The level of fog particles has increased over time due to rising human needs. The most important aspects of defogging algorithms made to improve road safety involve finding vehicles and lanes [60]. Some applications of intelligent transportation systems are as follows:

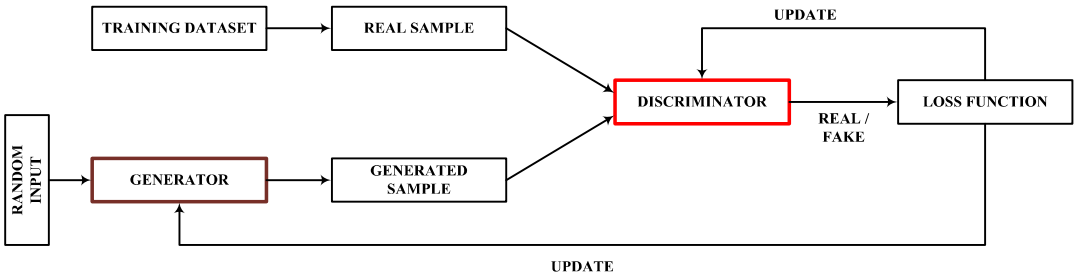


Fig. 1.7 Generative adversarial network architecture

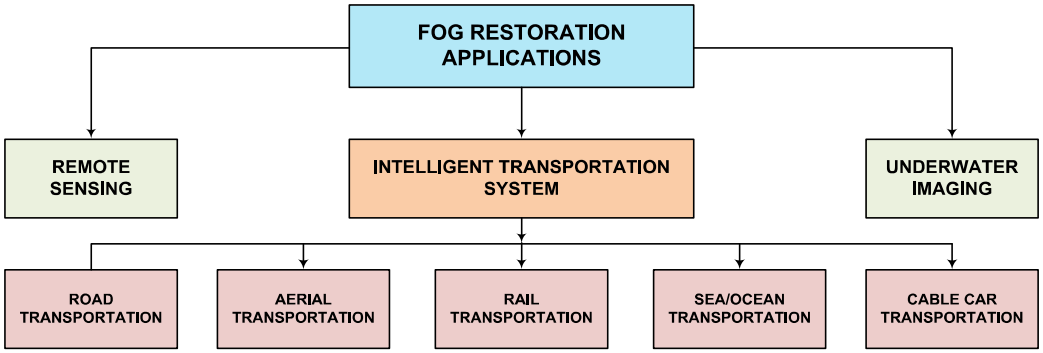


Fig. 1.8 Applications of defogging techniques

A real-time efficient and fast image defogging method is required to reduce the number of accidents on foggy days. The drivers need a high-speed, free live stream to make timely decisions [11]. An efficient defogging system can be designed. It can be used to prevent flight schedule changes during take-off and landing of airplanes. The delay or cancellation of trains due to severe pollution can be avoided. To address this issue, air pollution safety measures for locomotives are being tested [61]. However, their functionality is limited. As a result, locomotive pilots are kept up-to-date on signals and other approaching trains [62]. The defogging algorithms must be implemented to minimize train delays and ensure the safety of the passengers [63]. The smoke from industries near the sea shore rises above the sea or ocean water. The smoke particles interact chemically with the surface-condensed water vapors.

This results in the formation of smog, which reduces the helmsman's visibility within a kilometre. This affects sea traffic because severe pollution makes navigating difficult. Thus, the helmsman can obtain a clear scene of the sea/ocean fog by using defogging algorithms [64]. Cable car transportation is a means of transportation that

uses cables and motor-driven pulleys to move the cable car at a constant rate [65]. The cables are driven by mechanical energy via and are used to pull the trolley. It is most commonly used in hilly areas to reduce the distance between two distant intersections. Due to fog, it is either delayed or cancelled in order to avoid accidents. Thus, defogging techniques can be enforced to enhance the visibility of cable car drivers.

1.6.3 Underwater Imaging

Color casting and color sprinkling are issues encountered during underwater imaging. The spectrum of colors in sunlight has distinct wavelengths. Color casting refers to the appearance of a bluish color in the underwater environment. The light attenuates as it travels further into the ocean. Color sprinkling is a light scattering phenomenon produced by water particles that reduces visibility in underwater images. Defogging techniques can be used to mitigate the effects of color cast and color sprinkling [66].

1.7 Weather Degradation Datasets

Researchers have contributed a number of real and synthetic datasets for image defogging. It is observed from the literature that real datasets have very few images [67]. Foggy Road Image DAtabase (FRIDA) with 66 images is the first synthetic dataset in which traffic scenes were generated with the help of computer graphics [68]. Zhang et al. [69] established a dataset called HazeRD that consists of 14 pairs of clear images and their depth maps. D-HAZY was introduced by Ancuti et al. [70]. It is a synthetic dataset built on Middlebury [71] and New York University (NYU) datasets [72]. It consists of 1400+ pairs of hazy and clear images. Koschmieder's light propagation model [39] and image depth maps are used to generate hazy images from clear images. Ancuti et al. [73] proposed a real dataset called Dense-Haze for image defogging. O-Haze [74] dataset consists of light hazy images whereas the Dense-Haze is an updated version of the O-Haze dataset that has dense hazy scenes. Dense-Haze consists of 33 pairs of haze-free and real hazy images. Li et al. [75] presented a realistic single image dehazing (RESIDE) dataset that consists of three training subsets and two testing subsets. Ancuti et al. [76] introduced an image dataset that consists of non-homogenous haze. A haze generator was used to introduce the artificial haze in the background. Liu et al. [67] presented a real dataset for image defogging known as the Multiple Real Fog Image Dataset (MRFID). It consists of 200 natural and clear

outdoor scenes and their corresponding foggy scenes with different intensities. Table 1.2 depicts the comparative analysis of different datasets that exist in the literature. Apart from them, other frequently used datasets in fog restoration techniques are Berkeley Segmentation Dataset (BSDS300) [77], color lines [78], Image and Vision Computing [79], FRIDA2 [80], I-HAZE [81], structural static scenarios [82], LIVE [83], visual object classes [84], make 3D [85], ImageNet [86], Karlsruhe Institute of Technology and Toyota Technological Institute (KITTI) [87], NH-HAZE 2 [88], NH-HAZE 3 [89], Foggy Cityscapes [90], Foggy Driving Dataset [91], and REal-world VIdeo DEhazing (REVIDE) [92].

1.8 Performance Measures

This section presents various performance metrics for evaluating the performance of defogging techniques. Figure 1.9 depicts different performance metrics derived to assess the performance of defogging techniques [93]–[97]. Some of these metrics are calculated by comparing reconstructed image and ground truth images, whilst others do not. Furthermore, the calculation of computational complexity has become more important over time.

1.8.1 Comparison with Ground Truth Images

When reference or ground truth foggy images are available, the performance metrics such as the structure similarity index metric ($SSIM$), peak signal-to-noise ratio ($PSNR$), mean squared error (MSE), absolute mean error (AME), and Commission Internationale de l'Éclairage (CIE) for color deviation, may be computed.

- **Structure Similarity Index Metric**

Structure similarity index metric ($SSIM$) is calculated by comparing the edges, contrast, and luminance of the ground truth image and reconstructed image [98]. It is computed as [55]:

$$SSIM(g,p) = \left(\frac{2\,\mu_g\,\mu_p + c_1}{\mu_g^2 + \mu_p^2 + c_1}\right)\left(\frac{2\,\sigma_j + c_2}{\sigma_g^2 + \sigma_p^2 + c_2}\right) \tag{1.3}$$

where g refers to ground truth image. p represents the reconstructed image. μ_g and μ_p are sample means of g and p, respectively. σ_g^2 and σ_p^2 are sample variances of g and p, respectively. σ_j is sample cross co-variance between g and p. c_1 and c_2 are constants.

Table 1.2 Comparison of benchmark hazy datasets (Out*-Outdoor; In*-Indoor; Re*-Real; Syn*-Synthetic)

Specifications	Datasets							
	FRIDA	**HAZERD**	**D-HAZY**	**DENSE-HAZE**	**O-HAZE**	**RESIDE**	**NH-HAZE**	**MRFID**
Year	2010	2017	2016	2018	2018	2018	2020	2019
No. of Images	108	45	1400+	55	70	429,292	110	1000
Resolution	640×480	-	-	-	5456× 3632	620×460	5456×3632	2284×914
Software / Hardware Used	SiVIC™ Software	-	-	Sony A500	Sony RM-VPR1	MATLAB, Pycaffe	Sony RM-VPR1	-
Location of Images	Out*	In*, Out*	In*	In*, Out*	Out*	In*, Out*	Out*	Out*
Nature of Fog	Syn*	Re*	Syn*	Syn*	Re*	Syn*	Re*	Re*
Application Domain	Fewer traffic roads	Empty locations in wide areas	Few objects in a small area	Empty parks	Parks and playgrounds	Pedestrians, furniture	Parks and playgrounds	Fixed cameras on road-side
Advantages	Less computation time	Good contrast of images	Size of dataset is large	-	Good contrast of images	Good for object detection	Color restoration is good	Various images of same locations
Disadvantages	Poor illumination conditions	Objects with color similar to atmospheric light are not defogged	Less accuracy of computed images	Suffers from color distortion and structural distortion	Poor accuracy	Suffers from color distortion	No appropriate algorithm for dense fog	Suffers from halo effect and color distortion

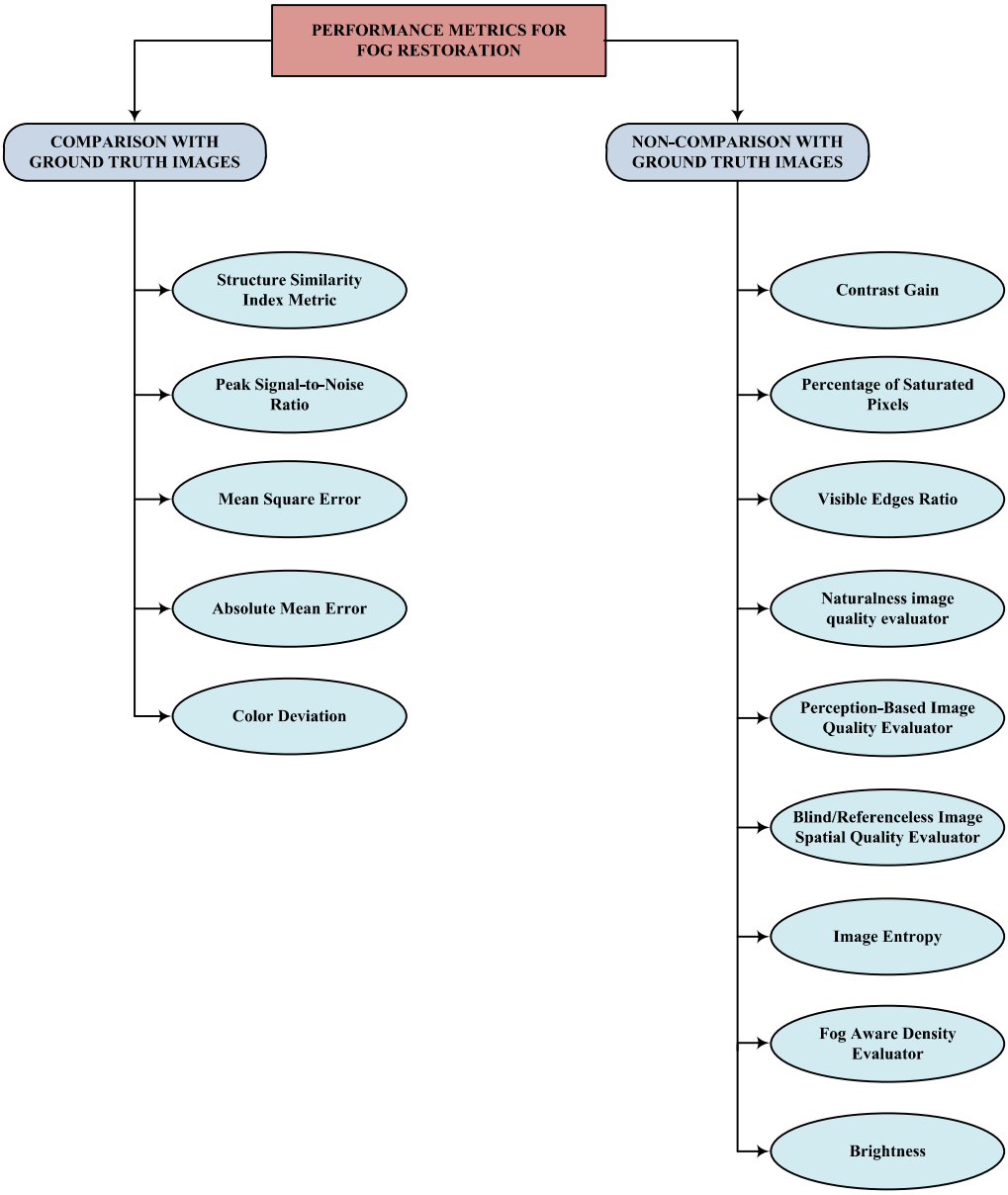


Fig. 1.9 Performance measures for defogging techniques

- **Peak Signal to Noise Ratio**

Peak signal-to-noise ratio ($PSNR$) is most common performance metric used to analyze the visibility restoration techniques. It computes the distorted image index against the mean square error (MSE) [55], [99]. A large value of $PSNR$ indicates less image distortion. $PSNR$ is computed as [99]:

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \tag{1.4}$$

- **Mean Square Error**

Mean square error (MSE) is the average of difference of squared intensity values of pixels of the target image ($J'(x,y)$) and reconstructed image ($J(x,y)$). It is calculated as [55], [99]:

$$MSE = \frac{1}{M \times N}\left(\sum_{i=1}^{M}\sum_{j=1}^{N}[J(x,y) - J'(x,y)]^2\right) \tag{1.5}$$

where M and N represent length and breadth of the image, respectively.

- **Absolute Mean Error**

Absolute Mean Error (AME) is calculated to estimate the absolute difference between the ground truth image and reconstructed image. For an 8-bit image, AME has a value between 0 and 255. A smaller value of AME is prefered. It is computed as [100], [101]:

$$AME = \frac{1}{M \times N}\left(\sum_{p=1}^{(M \times N)}|J(x,y) - J'(x,y)|\right) \tag{1.6}$$

- **Color Deviation**

When different defogging algorithms are used to process the images, the resulting images suffer from color deviation from the ground truth images. The color difference is estimated by implementing the CIEDE color-difference formula [102]. A small value of color deviation is preferred. The color deviation (ΔE_{00}) in the resultant images occurs due to large differences in pixels' values from the ground truth images based on lighting, chroma, and hue. It can be calculated using the CIE color difference formula [103].

$$\Delta E_{00} = \sqrt{\left(\frac{\Delta \mathrm{L}'}{k_L S_L}\right) + \left(\frac{\Delta \mathrm{C}'}{k_C S_C}\right) + \left(\frac{\Delta \mathrm{H}'}{k_H S_H}\right) + R_T\left(\frac{\Delta \mathrm{C}'}{k_C S_C}\right)\left(\frac{\Delta \mathrm{H}'}{k_H S_H}\right)} \tag{1.7}$$

where $\Delta \mathrm{L}'$, $\Delta \mathrm{C}'$, and $\Delta \mathrm{H}'$ represents the lighting, chroma, and hue components of the pair of reference and computed image, respectively. k_L, k_C, and k_H are parametric factors for lighting, chroma, and hue components, respectively. These are adjusted according to background and texture of the image. S_L, S_C, and S_H are weighting functions for lighting, chroma, and hue components, respectively. R_T presents rotation of different planes of color spaces.

1.8.2 Non-Comparison with Ground Truth Images

The ground truth images are not available for real-time applications. They are also known as blind measures. The performance metrics that can be calculated in the absence of ground truth images are contrast gain (C_G), percentage of saturated pixels ($\mathbb{P}$), visible edge ratio ($\mathcal{E}$), naturalness image quality evaluator ($NIQE$), perception-based image quality evaluator ($PIQE$), blind/referenceless image spatial quality evaluator ($BRISQUE$), image entropy (IE), brightness ($\beta 1$), and fog aware density evaluator ($FADE$).

- **Contrast Gain**

Contrast (C_G) is the ratio of standard deviation ($\mathcal{s}$) and mean (m) of all the pixels of the image. It is computed as [104]:

$$C_G = \frac{\mathcal{s}}{m} \tag{1.8}$$

where standard deviation of pixels is computed as follows:

$$\mathcal{s} = \sqrt{\left(\frac{1}{M \times N - 1}\left(\sum_{p=1}^{(M \times N)} (p_i(m_r, n_c) - \mathfrak{b})^2\right)\right)} \tag{1.9}$$

where $M \times N$ is the resolution of the image. p_i is normalized pixel intensity. m_r and n_c are number of rows and number of columns, respectively. $\mathfrak{b}$ is the brightness or the average intensity of all pixels of the image.

Contrast gain (C_{Diff}) is defined as the difference of contrast of foggy and defogged images.

$$C_{Diff} = C_{G_{fog-free}} - C_{G_{foggy}} \tag{1.10}$$

where $C_{G_{fog-free}}$ is the contrast of reconstructed image. $C_{G_{foggy}}$ is the contrast of foggy image. A larger value of contrast gain is desired.

- **Percentage of Saturated Pixels**

Saturated pixels have either a minimum pixel intensity of 0 or a maximum pixel intensity of 255. The percentage of saturated pixels ($\mathbb{P}$) is calculated by estimating the ratio of the number of saturated pixels to the total number of pixels [99], [105]. A small value of $\mathbb{P}$ means illumination conditions are uniform.

$$\mathbb{P} = \frac{P_S}{M \times N} \tag{1.11}$$

where P_S is the number of pixels with intensities ranging from 0 to 255.

- **Visible Edges Ratio**

This performance metric comprises the number of new visible edges ($\mathcal{E}$) and the ratio of the average gradient ($\overline{g}$). $\mathcal{E}$ denotes the new and improved visible edges after the implementation of defogging technique. A larger value of $\mathcal{E}$ is desired. It is computed as [106]:

$$\mathcal{E} = \frac{n_{defog} - n_{foggy}}{n_{foggy}} \tag{1.12}$$

where n_{defog} and n_{foggy} represents the cardinal number of visible edges, respectively.

The gradients of visible edges $\overline{g}$ denote the restoration degree and texture details of the reconstructed image. A larger value of $\overline{g}$ is desired. It is computed as [106]:

$$\overline{g} = \mathrm{e}^{\left[\frac{1}{n_{defog}} \sum_{m \in \phi_{defog}} \log(g_m)\right]} \tag{1.13}$$

Here, g_m can be represented as:

$$g_m = \frac{\Delta_{\mathrm{defog}}}{\Delta_{foggy}} \tag{1.14}$$

where g_m represents the visible edges of the reconstructed image. Δ_{defog} and Δ_{foggy} are gradients of the computed image and foggy image, respectively.

- **Naturalness Image Quality Evaluator**

The naturalness image quality evaluator ($NIQE$) is used to calculate the variation of the features determined from the defogged image as compared to the natural scene statistics [107]. A smaller value of $NIQE$ is preferred.

- **Perception-Based Image Quality Evaluator**

Perception-Based Image Quality Evaluator ($PIQE$) is used to calculate the block-wise distortion estimation [108]. A smaller value of $PIQE$ is desired for high-quality images with less distortion.

- **Blind/Referenceless Image Spatial Quality Evaluator**

$BRISQUE$ [109] does not consider distortion-related features such as block and blur when evaluating the image quality. It is estimated for real-time applications due to its lesser computational complexity as compared to $PIQE$ and $NIQE$.

- **Image Entropy**

Image entropy (IE) is defined as the estimation of randomness present in the constitution of the image. Shannon entropy of a random discrete variable (say Z) is defined as [110]:

$$IE(Z) = -\sum_{z \in Z} p(z) log p(z) \tag{1.15}$$

where $p(z) \epsilon [0,1]$ and $\sum_{z \in Z} p(z) = 1$. $-log p(z)$ represents the information associated with single occurrence of z.

- **Fog Aware Density Evaluator**

The effectiveness of RRNet on real-world images is calculated using the fog aware density evaluator ($FADE$) [83], [111]–[113]. It depends on different factors such as color saturation, contrast energy, image entropy, and sharpness of the image instead of geographical position of camera, depth information, and human-rated judgements. A low value of fog density is desired. If the computed value is zero, then the image is fog-free.

- **Brightness**

Brightness ($\beta 1$) is defined as the overall darkness or lightness present in the image. It is the mean value of intensities of all pixels of the image, i.e., ratio of sum of all elements to number of elements. It is computed as [114]:

$$\beta 1 = \frac{1}{H \times V} \sum_{h=1}^{H} \sum_{v=1}^{V} (x_i, y_i)(h, v) \tag{1.16}$$

The brightness of a dehazed image should be less than the brightness of the corresponding hazy image because hazy regions have higher intensity values. An image is considered to have good quality if the mean value of all pixels lies in between 115 and 140. H and V represent the width and height of the images, respectively.

1.9 Thesis Organization

The structure of the thesis is as follows:

Chapter 1 presents an introduction to the fog imaging problem and the motivation to solve it using computer vision technique. The chapter discusses in detail the different weather degradation techniques, along with the scope and applications of real-time dehazing. Also, the comparative analysis of several weather degradation

datasets is discussed. Furthermore, various reference-based and reference less performance metrics used to evaluate the performance of dehazing techniques are discussed.

Chapter 2 presents an enhanced survey and analysis of the previous state-of-the-art methods for image and video dehazing. This chapter discusses different prior-based, deep learning-based, and hybrid techniques for different weather conditions proposed by various researchers over the years. It is followed by a detailed discussion of different real-time visibility restoration frameworks that exist in the literature.

Chapter 3 discusses the challenges present in existing datasets, followed by detailed description of the developed dataset, namely Hazy Unpaired Dataset for Road Safety (HUDRS). No external haze machine is used to generate haze. Real-world foggy images are taken under natural environmental conditions.

Chapter 4 presents the design and working of grey wolf optimization technique. Most of the existing filter-based techniques suffer from edge and texture distortion. In the proposed approach, the edges and texture of a computed image are improved by evaluating a fitness function using iterative method. This chapter also presents the design and working of a novel wavelet and inverted wavelet-based multi-scale convolution neural network for single image dehazing. The importance of retinex and wavelet filters for image processing are discussed. Also, it highlights the importance of gamma correction.

Chapter 5 presents the design and working of a novel deep learning image and video dehazing architecture called Aethra-Net. The existing video dehazing algorithms suffer from color distortion due to the continuous processing of frames. Thus, they are not suitable for videos with dense haze. This drawback is resolved in Aethra-Net architecture.

Chapter 6 presents designing and working of a bounding function for gray-world kernel prior algorithm (BFGKP). In this model, the computational complexity is reduced and the airlight is updated for each frame. This algorithm is used in design and development of novel static and dynamic dehazing frameworks for road safety using readily available electronic equipment and devices. Further, this model is used for the design of a web-based application for real-time dehazing.

Chapter 7 presents the design and working of a novel deep learning architecture called residual regression network for single image desmogging. The performance evaluation of the proposed and existing smoggy image restoration techniques is performed over several datasets to predict the most suitable model for this task.

Chapter 8 concludes the research contributions and findings of the work. The chapter also sheds light on few limitations of the proposed approaches to be addressed in future.

CHAPTER-2
LITERATURE REVIEW

The visibility restoration algorithms for fog, haze, and smog for still images and real-time framework presented by various researchers are implemented using prior-based techniques, learning-based models, or their hybrid. The detailed comparative analysis is presented in the following sections.

2.1 Single Image Visibility Restoration

This section discusses different single image visibility restoration techniques for fog, haze, and smog proposed by various researchers over the years. The visibility restoration techniques are classified as prior-based, deep learning-based, and hybrid techniques. The prior-based techniques are computed using the intrinsic properties of images such as pixel intensity, color distortion, texture, and contrast. The learning-based dehazing techniques, in contrast to prior-based techniques, learn models rather than manually extracting features. The hybrid architectures utilize the advantages of both prior-based and deep learning-based techniques.

2.1.1 Prior-based Techniques

Kratz and Nishino [115] discussed the importance of the scene depth and ratio of reflected-to-incident light in image processing by using the Markov Random Field architecture. It factorized the hazy input by estimating the maximum a posteriori (MAP). The MAP was predicted by using an expectation maximization algorithm. This technique had high saturation and low accuracy during the night-time.

He et al. [116] presented an algorithm based on dark channel prior (DCP) for dehazing images based on the assumption that atleast one of the three color-channels of the clear images namely, Red (R), Green (G), and Blue (B), had low intensity. This was achieved by estimating a dark channel. But the brighter areas, mainly sky, did not follow the DCP and resulted in color distortion. The pixel values of the dark channel in clear images were assumed to tend towards zero. This assumption was proven false when the contrast of scene objects was similar to the atmospheric light.

Zhu et al. [117] proposed a color attenuation prior (CAP) to estimate the transmission map. It established a linear relationship between the brightness of images

and scene depth. In this method, supervised learning was used to learn the parameters of the given model, such as scene depth and scene radiance. This helped in estimating the transmission map. In this method, the value of scattering coefficient was supposed to be constant but it decreased as the depth increased.

In [118], the smoggy images were restored using a reconciled color algorithm to estimate depth map information and scattering light coefficient. It utilized traditional statistical methods, i.e., the calculation of mean and standard deviation, to estimate the transmission map and airlight. However, this technique was not effective in smoky regions.

The Weiner filter had been shown to be effective for noise reduction and deblurring in grayscale smoky images [119]. The transmission map was refined using the traditional DCP technique. Although a soft matting layer was implemented for edge smoothening, the computation time was increased.

Berman et al. [77] observed that the hundreds of colours forming a tight cluster in a haze-free image change to haze-lines in the RGB domain when synthetic haze was computed in images after passing through airlight coordinates. When the scene was dimmer than the airlight, the haze lines were difficult to detect because pixels pointed in the same direction that led to the failure of this method.

Zhang et al. [120] proposed a method called maximum reluctance prior to estimate ambient illumination and the transmission map to find the source of local intensity in hazy images during the daytime for each individual channel {R, G, B}. The estimated ambient illumination was used to recover the nighttime haze-free image. Although the results generated by using prior-knowledge based techniques were remarkable, the dehazing quality varied with the properties of the actual image.

In [121], the transmission values were computed by using ellipsoid geometry. It resulted in a contrast gain for the dehazed pixels. The simultaneous execution of fuzzy segmentation process was performed to refine the transmission map. Although the computation time was less and gradient artifacts were minimized, the resultant images had distorted edges.

Chen et al. [122] discussed a multi-channel smoke removal technique to overcome this limitation. It utilized a random forest regression. The transmission map and

atmospheric light were estimated for different color channels using feature extraction. However, individual processing of color channels increased the computation time.

In [123], a gain coefficient-based trilateral filter was developed that produced a fog-free image with little or no artifact. It overcame the limitation of DCP to defog the objects that had same color as atmospheric light. But, halo effects were introduced in the computed images.

A dehazing technique implying a combination of dark channel prior (DCP) and bright channel prior (BCP) was developed by Singh and Kumar [124]. In this technique, DCP computes the non-sky region and BCP computes the sky region. The resultant images had poor contrast and low brightness.

Singh and Kumar [125] discussed the defogging process using a gain gradient image filter. This model focused on the reduction of gradient reversal effects and halo impacts in the fog-free image. The atmospheric veil was improved by using the gain coefficient-based filter. In this method, the edges of the objects were preserved in a more significant manner compared to previous techniques. However, the desired results were not obtained for dense foggy images.

Ngo et al. [126] proposed an enhanced colour attenuation prior (ICAP) comprised of a low-pass filter (LPF) and an improved median filter (mHMF) for determining depth and refinement. Also, the image's background distortion was minimized using LPF. A quadratic decomposition algorithm was used to estimate atmospheric light. Also, an adaptive tone mapping was used to post-process the reconstructed hazy image. In terms of performance metrics, it was ineffective compared to previously utilised strategies.

In [127], the transmission map was improved by utilizing guided anisotropic diffusion and an iterative learning filter (GADILF). This reduced pixel saturation, but color was distorted. A window-based integrated means filter (WIMF) was developed to refine the transmission map in [128]. It calculated the gradient magnitude of a hazy image and divided it into low and high-frequency regions.

Bala and Lakhwani [129] used an illumination channel prior-based technique to estimate the transmission map and the atmospheric light. It was further refined using a gradient magnitude-based filter. Although it preserved the color texture of the computed images, it over-exposed them to light.

Li et al. [130] proposed a two-stage technique for desmogging and dehazing. This method utilized a homomorphic filter (HF) and a modified DCP algorithm. HF was utilized to evenly disperse the smog thickness throughout the image. The conventional DCP algorithm was introduced to smoothen the edges. But it introduced color distortion and gradient artifacts into the resultant images as it was sensitive to noise. To circumvent this limitation, a sphere-shaped window was used instead of a rectangular window to analyze the approximated cluster of pixels with higher efficiency. The resultant images, however, were darker than the original smog-free images. The model became overfit and was rectified using an L0 filter in [131].

In [132], an iterative optimization technique for natural color restoration was presented. The haze-related features were estimated using contrast energy, image entropy, standard deviation, and normal dispersion. But the computation cost was high.

In [133], a non-linear model was proposed for the estimation of transmission function for quality enhancement. The density of haze presented in the image was controlled by tuning the quality control parameter. However, the desired results were not obtained for dense foggy images.

In [134], the transmission map was estimated by integrating the oblique gradient prior map and luminance map. The refined map was generated by using a variational regularized model. The hyperparameters tuning was performed by implementing non-dominated sorting genetic algorithm (NSGA) to preserve the edges and color of the computed image. However, the computed images were still affected by color distortion.

In [135], the transmission map was estimated by using a gradient channel prior. Further, an information gain-based bilateral filter was also used to refine the transmission map. This filter changed the intensity of the pixel if all the sibling pixels had significant information gain. However, the hyperparameter tuning was not considered.

Multi-exposure fusion technique was used for a thorough investigation of structural patch decomposition (SPD) [136]. It was an efficient technique for high-resolution dynamic scenarios. It minimized edge distortion and halo artefacts in the resultant image.

Dhara et. al. [137] performed color correction on hazy images by implementing non-linear color balancing and adaptive airlight refinement. Their method reduced the

halo effect in dehazed images.

A channel difference prior (CDP) method was proposed for dense hazy images in [138]. The blurred image was divided into patches using a quick block-based method to estimate the transmission map. The guided filter and weighted interpolation technique smoothed the edges. Although the resultant image had high visual quality, it suffered from high pixel intensity. This limitation was minimized for image dehazing and exposure (IDE) using Gray-world assumption (GWA) algorithm [139].

Hu et al. [140] proposed a sky segmentation and optimal transmission map-based dehazing algorithm. It produced images with high color fidelity and clear pixels. But it was computationally inefficient for video dehazing.

Table 2.1 presents the comparative analysis of the discussed prior-based dehazing techniques.

2.1.2 Deep Learning-based Techniques

Hussain and Jeong [141] proposed a learning-based deep neural network (DNN) to convert a 2D vector into 1D vector by rasterization during image restoration. It demonstrated resiliency for unseen image data. The dehazed image had low pixel intensity and a software-generated appearance.

Two activation functions were used in the same architecture for the dehazing approach suggested by Cai et al. [142]. The first function "maxout," was used for non-linear mapping, while the second function, "bilateral rectified linear unit (BReLu)," enhanced the convergence and decreased the search space. A hazy input image could be restored to a clear image using DehazeNet. the medium transmission map was estimated. This approach was not found effective for densely foggy images.

A multi-scale convolutional neural network (MSCNN) based dehazing architecture was developed by Ren et al. [143] This architecture was comprised of two layers, i.e., a coarse layer to estimate the holistic transmission map of the scene and a fine layer to modify it using local information. Images taken at night did not give the expected results.

All-in-one dehazing architecture was designed to eliminate color distortion, edge distortion, and halo effects [144]. However, this algorithm was not suitable for images with dense haze. Also, it had gradient artifacts.

Table 2.1 Comparative Analysis of Prior-based Defogging Techniques (Ref – Reference, CD – Color Distortion, ED – Edge Distortion, HE – Halo Effect, LHG – Large Haze Gradients, BA – Block Artifacts, EL – Excessive Lighting)

Algorithm	Ref	Year	Dataset(s)	Type of Algorithm	Prior Knowledge	Remarks / Knowledge Gap	Quality Metrics					
							CD	ED	HE	LHG	BA	EL
Scene Albedo (SA)	[115]	2009	Internet Browsing	Image Enhancement	Gradient orientation in chromaticity image is invariant to illumination.	• No comparative analysis available.	✓	✓	✗	✓	✓	✗
Dark Channel Prior (DCP)	[116]	2011	Internet Browsing	Transmission Refinement	Pixel intensity is very small in a local patch in atleast one of the color channels.	• No comparative analysis available.	✓	✗	✓	✓	✓	✗
Color Attenuation Prior (CAP)	[117]	2015	Middlebury	Transmission Refinement	Linear relationship of brightness and depth map minimizes the color distortion in sky region.	• Resultant images suffer from partial color distortion. • High computation time as compared to other algorithms	✓	✗	✗	✓	✗	✓
Reconciling Color (RC)	[118]	2015	Internet Browsing	Image Enhancement	Reconciling color enhances the image's brightness	• Inadequate results.	✓	✓	✓	✗	✗	✗
Weiner Filter (WF)	[119]	2016	Internet Browsing	Image Enhancement	Classic DCP enhances image contrast	• Inadequate results.	✗	✓	✗	✓	✓	✗
Non-local Dehazing (NLD)	[77]	2016	Berkeley Segmentation Dataset (BSDS300)	Image Enhancement	Images comprise of more than hundred different colours.	• Computed images appear artificial due to high contrast.	✓	✗	✗	✗	✗	✓
Maximum Reflectance Prior (MRP)	[120]	2017	Middlebury	Transmission Refinement	Images are visible in day time due to reflected light.	• Invalid for images with proper illumination or distinct colors.	✓	✗	✓	✗	✗	✓

Table 2.1 (continue) Comparative Analysis of Prior-based Defogging Techniques (Ref – Reference, CD – Color Distortion, ED – Edge Distortion, HE – Halo Effect, LHG – Large Haze Gradients, BA – Block Artifacts, EL – Excessive Lighting)

Algorithm	Ref	Year	Dataset(s)	Type of Algorithm	Prior Knowledge	Remarks / Knowledge Gap	Quality Metrics					
							CD	ED	HE	LHG	BA	EL
Color Ellipsoid Prior (CEP)	[121]	2018	Color Lines	Image Fusion	Vector with minimum color component on ellipsoidal surface produces pixels with high contrast.	• -	✓	✓	✗	✗	✗	✗
Multi-color channel enhancement with machine learning	[122]	2018	Internet Browsing	Image Enhancement	Intensity distribution of a pixel is different for different colours.	• Inadequate results.	✗	✓	✓	✓	✗	✗
Gain Gradient Filter (GGF)	[125]	2019	-	Image Enhancement	All objects are highly illuminated in atmospheric light.	• Inadequate results.	✗	✗	✓	✓	✗	✗
Improved Color Attenuation Prior (ICAP)	[126]	2019	IVC; FRIDA2; D-HAZY; O-HAZE; I-HAZE	Transmission Refinement	Linear relationship of brightness and depth map minimizes the color distortion in sky region.	• Poor performance in terms of quantitative analysis.	✗	✗	✗	✓	✗	✓
Gradient Channel Prior (GCP)	[127]	2019	RESIDE; FRIDA; FRIDA2; IVC; D-HAZY	Transmission Refinement	Depth estimation plays a key role to process complex background.	• Ineffective for dense hazy images.	✓	✗	✓	✓	✗	✗
Window-based Integrated Means Filter (WIMF)	[128]	2019	FRIDA; FRIDA2; RESIDE; IVC; D-HAZY	Transmission Refinement	Gradients can decompose image into different frequencies.	• High noise. • Distorted transmission map.	✗	✗	✓	✓	✓	✗
Illumination Channel Prior (ICP)	[129]	2019	Grayscale Images	Transmission Refinement	-	• Inadequate results.	✓	✓	✓	✓	✓	✓

Table 2.1 (continue) Comparative Analysis of Prior-based Defogging Techniques (Ref – Reference, CD – Color Distortion, ED – Edge Distortion, HE – Halo Effect, LHG – Large Haze Gradients, BA – Block Artifacts, EL – Excessive Lighting)

Algorithm	Ref	Year	Dataset(s)	Type of Algorithm	Prior Knowledge	Remarks / Knowledge Gap	Quality Metrics					
							CD	ED	HE	LHG	BA	EL
Improved Dark Channel Prior (IDCP)	[130]	2019	Satellite Images	Transmission Refinement	-	• Inadequate results.	✓	✗	✓	✓	✓	✓
Guided L_0 Filter (GLF)	[131]	2020	RESIDE; FRIDA; FRIDA2; IVC; D-HAZY	Transmission Refinement	-	• No naturalness in resultant images.	✗	✗	✓	✗	✓	✗
Optimal Transmission Map and Adaptive Atmospheric Light (OTM-AAL)	[132]	2020		Optimization	-	• Not applicable for real-time systems due to lack of optimization	✗	✗	✗	✓	✓	✗
Non-Linear Bounding Function (NLBF)	[133]	2020	Internet Browsing	Transmission Refinement	Bounding technique minimizes reconstruction error	• Smoothing coefficient is inefficient for dense hazy images.	✗	✗	✗	✓	✓	✗
Variational Minimization (VM)	[134]	2020	Internet Browsing	Image Fusion	Non-dominated Sorting Genetic Algorithm tunes the hyperparameters.	-	✓	✗	✗	✗	✗	✗
Information Gain-based Bilateral Filter (IGBF)	[135]	2020	Internet browsing	Transmission Refinement	Large smog gradients hinder texture information.	• Inadequate results.	✓	✗	✓	✗	✓	✗

Table 2.1 (continue) Comparative Analysis of Prior-based Defogging Techniques (Ref – Reference, CD – Color Distortion, ED – Edge Distortion, HE – Halo Effect, LHG – Large Haze Gradients, BA – Block Artifacts, EL – Excessive Lighting)

Algorithm	Ref	Year	Dataset(s)	Type of Algorithm	Prior Knowledge	Remarks / Knowledge Gap	Quality Metrics					
							CD	ED	HE	LHG	BA	EL
Structural Patch Decomposition (SPD)	[136]	2020	Static scenarios	Image Fusion	Three characteristics of a signal, namely strength, structure, and intensity, must be separately enhanced.	• Not entirely devoted to defogging application.	✓	✗	✗	✗	✓	✓
Color Cast Dependent Image (CCDI)	[137]	2021	LIVE; RESIDE; O-HAZE	Transmission Refinement	Airlight estimation is equally important as transmission estimation.	• Computed images have dark pixels, therefore it is not applicable for night time images	✗	✗	✗	✓	✓	✓
Channel Difference Prior (CDP)	[138]	2021	RESIDE	Image Enhancement	Transmission map degrades as restoration factor increases.	• No naturalness in resultant images.	✓	✓	✗	✗	✗	✓
Image Dehazing and Exposure (IDE)	[139]	2021	RESIDE	Image Enhancement	Scene depth can be estimation without trained network.	-	✓	✗	✗	✗	✗	✗
Sky Segmentation (SS)	[140]	2023	RESIDE	Transmission Refinement	Adaptive estimation of airlight based on sky color.	• Ineffective for video dehazing due to high computation time.	✓	✗	✓	✓	✗	✓

A deep fully convolutional regression network (DFCRN) was presented in Zhao et al. [145] to obtain a transmission map. The upsampling unit was used to avoid the problem of overfitting.

Li et al. [146] proposed a cascaded CNN in which the input hazy image was initially processed using shared hidden layers. Its output was connected to two different sub-networks for global atmospheric light and medium transmission map estimation. It was suitable for images with synthetic fog only.

Liu et al. [147] proposed an unsupervised learning model called GAN for training the database. The relationship between the transmission map and features, such as extracted multi-scale structural features was dependent on sparse coding and haze-related color features. However, the learning-based techniques failed when the condition became unsatisfying.

Ren et al. [148] proposed an autoencoder-based gating fusion network, such that three layers of transmission map namely, contrast enhancement, white balance, and gamma correction were generated. The white balance minimized the chromatic casts occurring due to atmospheric light, contrast enhancement increased the luminance, and gamma correction was used to limit the darkness in the haze-free output. This model was not applicable for dense foggy situations.

Kim et al. [149] proposed patch classification-based CNN. A method called quadtree decomposition was used to decompose hazy images into patches. The smaller patches preserved more details. This model was claimed to be the first architecture to use a deep learning algorithm for dehazing tasks using patch classification. In this method, only one-fourth of the patch was considered on each iteration, which reduced the processing time.

Wang et al. [150] proposed a novel method called scene classification region-based CNN (SC-R-CNN) which used remote sensing and object detection network to locate the ship using the defogging algorithm. This network simultaneously focused on object detection and classification making the identification process dual-stream. The accuracy of this model was higher than conventional R-CNN architecture.

Dudhane and Murala [151] proposed cardinal colour fusion multi-scale CNN (C^2MSNet) in two phases. The first phase was used to estimate multi-depth channels using color fusion. The second phase introduced multi-channel, multi-scale CNN for

refinement of the feature maps. The model was robust, and the peak signal-to-noise ratio had been significantly enhanced. But it was color-centric and ignored other image parameters.

Li et al. [152] proposed the conditional GAN (cGAN) technique by utilizing a regularized gradient and a visual geometry group (VGG) network. The generator used a pair of encoder and decoder to process hazy images. The discriminator was used to detect if the image was fake or real. This method did not have heterogeneity in the training datasets. Thus, it was not suitable for night-time and less hazy images.

A gated context aggregation network was presented in [153]. It was implemented for defogging as well as dehazing tasks. In this network, features were fused by using sub-networks. For instance, semantic features were fused using the feature pyramid. Also, dilated convolution was performed to reduce the effect of gridding artifacts. But the dehazed results suffered from color distortion.

An unsupervised deep learning method for image dehazing was discussed in [154]. It utilized CycleGAN architecture with two generators. The first generator covered the haze-free image with a haze mask synthesized to recover an improved dehazed image. The second generator repeated the process of covering and removing the haze mask to train the image. The initial image was recovered and discriminator detected if the image was real or fake. The implementation of multiple generator networks increased the processing time.

Wang et al. [155] utilized prior knowledge of YCrCb luminance channel called atmospheric illumination prior (AIP). It was designed to extract the features of hazy regions and perform image dehazing using multiscale restoration. The color was preserved in the processed image. These combination techniques had a high success rate but needed significant optimization because computed images have block artifacts and excessive lighting.

An enhanced version of CycleGAN was proposed in [67]. In this method, two different architectures were used. One is to obtain the fog-free images from the foggy images and another is to obtain the foggy images from the fog-free images to train the dataset. However, it was not suitable for images with dense fog. Also, the computed images had color and edge distortion.

Feature fusion attention network (FFA-Net) [156] consisted of three components. The first component, i.e., feature attention (FA), was a combination of channel attention (CA) and pixel attention (PA) to provide flexibility in CNN. The second component had an additional component with FA called local residual learning. This additional component focused only on high-level features and did not process the less important information. The third component, feature fusion preserved the information from the imbecile layers and sent it to the deep layers. Although it preserved the texture and the color of the restored image, it was not suitable for images with dense fog.

Singh et al. [157] proposed a back-projected pyramid network (BPPNet). It consisted of iterative U-Net blocks for complexity learning and multi-scale spatial learning was used for pyramid convolution. This method preserved the edges, but there were multi-coloured patches observed throughout the image.

Bennur et al. [158] proposed a convolution encoder that had a lighter architecture compared to CNN-based dehazing architectures such as FFA-Net and BPPNet. This feature made it compatible with the systems of lower specifications, but it was not reliable for images with dense haze.

Dong et. al. [159] proposed a multi-scale boosted dehazing network that composed of dense feature fusion modules. These modules were used to preserve the spatial information from the high-resolution features except those at non-adjacent levels. The spatial information of high-resolution features was preserved by using feature fusing modules. Also, the non-adjacent features were exploited.

The dehazing approach presented in Ren et al. [160] was similar to the conventional MSCNN [143]. In addition to the coarse scale and fine scale network, a holistic edge guided network was designed to smooth the edges of the complete objects depicted in the transmission map. The map became discontinuous at the scene depth.

Wu et al. [161] suggested a contrastive regularization-based autoencoder to restore naturalness in the images. In this network, contrastive regularization was applied to the restored image, such that the restored image was pushed towards a clear image and away from the hazy image. As a result, this algorithm performed well in both high and low-frequency regions of the images. The only drawback of this algorithm was the large memory requirement, as it was tested on a dataset with 10 haze-free and hazy images.

Fu et al. [162] proposed a wavelet transformation-based general adversarial network in which an autoencoder and skip connections were used to preserve the texture of the image. It was best suited for non-homogeneous haze. However, it was found to be ineffective for real-time roadside images.

In [163], the transmission map was estimated by using CNN. The image restoration was performed by implementing an oblique gradient channel prior. Jain and Kumar [164] proposed an oblique gradient profile prior to estimate depth map. The presence of smog gradients was predicted using transfer learning based GoogleNet Inception V3. It used energy minimization function to refine the transmission map. However, hyperparameter tuning was not considered in this architecture.

Zhang et al. [165] estimated the clear scene radiance by utilizing visible light polarization technique to detect smoggy regions. Further, an autoencoder generated a smog-free image. The process was repeated until the clear image was obtained. It was found effective near the source of smog pollution.

A multi-scale architecture was presented in [166] for image dehazing. It consisted of an adaptive feature selection module for the extraction of shallow features at the beginning of each scale, a self-calibration attention module to combine the features between different scales, and an adaptive feature enhancer module to produce the output image of desired resolution. This architecture was not suited for dense haze.

In [167], dilated densely connected blocks were used to obtain feature maps. It utilized multiple skip connections and residual connections to enhance the feature maps. However, the texture and edge visibility of the computed images are poor.

Tu et al. [168] proposed a multi-axis, multi-layer perceptron model for dehazing, deraining, denoising, deblurring, and light enhancement. The computed results had no color distortion, edge distortion, or halo effects. However, this model was not suitable for images with dense haze. Also, it could not perform two operations at the same time. For instance, it was not possible to restore the rain and the fog visibility in a single image.

Some of the latest deep learning-based dehazing techniques designed by researchers are presented in [169]–[171]. Tables 2.2 and 2.3 present the comparative analysis of the existing deep learning-based dehazing techniques.

2.1.3 Hybrid Techniques

Song et al. [172] proposed a novel method called ranking CNN in which features were extracted from hazy images using a random forest regression algorithm to produce a transmission map. In this technique, the structural and statistical features were arranged in ascending order, which reduced the computational efficiency of the network. But it was not proven effective for large haze gradients. Li et al. [173] proposed residual CNN to estimate atmospheric light. The sky region was optimally enhanced, and the image was natural and clear. But this technique suffered from large haze gradients.

Table 2.4 presents the comparative analysis of ranking CNN and residual CNN.

2.2 Real-time Visibility Restoration Frameworks

This section presents different real-time visibility restoration frameworks that exist in the literature.

Kim et al. [174] discussed an optimized contrast enhancement dehazing algorithm to reduce flickering artifacts in video sequences. In this algorithm, the atmospheric light was estimated using a quadtree-based subdivision, and the transmission map was estimated by extracting patches of different constraints calculated from the hazy image. However, it was unsuitable for videos with dense haze.

Ren et al. [175] suggested an autoencoder-based semantic segmentation (AESS) technique for video dehazing. However, this technique did not estimate a refined transmission map for real-world scenes with non-homogenous haze.

Mandal et al. [176] focused on enhancing the driver's visibility in real-time while also reducing the time complexity. The computation time required to produce dehazed frames was related to frame resolution. A look-up table was designed for the pixel's intensity and visibility restoration factor, known as gamma. But the computing time was significant for lower quality frames with lesser resolution as well.

Soma and Jatoth [177] suggested a video dehazing framework based on the Raspberry Pi 3 and Jetson Nano to minimize the color distortion. The hue, saturation, and light (HSL) color spaces were used in conjunction with the RGB color spaces for dehazing video frames. The halo effect decreased while the edges were preserved. Researchers have recently focused on a deep learning-based video dehazing approach. However, the computed frames suffered from color distortion.

Kumar et al. [178] suggested a VLSI-based embedded system to reduce the computing time and predict the dehazed output of desired brightness. It utilized fewer resources and power than the previous architectures. The adaptive airlight was updated to eliminate flickering artifacts. But the throughput of this system was approximately three times that of previous architectures.

Peng et al. [179] described a real-time video dehazing technique. It gave a frame-by-frame estimation of a spatial-temporal coherent transmission map by utilizing the dark channel prior. A guided joint bilateral filter was used to suppress the visual artifacts. The pixel intensities and depth of the frame, on the other hand, were ambiguous. To overcome this ambiguity, different color spaces such as RGB, HSV, and XYZ were combined to extract image features in an onboard camera-based dehazing architecture. This algorithm detected traffic signs in clear, cloudy, and foggy weather conditions. But the computed images suffer from color distortion and halo effect.

Dark channel prior is a traditional approach to perform image dehazing. Adidela et al. [180] proposed a DCP-based dehazing algorithm for single image and video dehazing. The transmission map was estimated for each frame. However, the computed images had color distortion and low brightness.

Zhang et al. [181] suggested a REal-world VIdeo DEhazing (REVIDE) dataset comprised of hazy videos. It was tested on the Confidence Guided and Improved Deformable Network (CG-IDN). The computed results were improved for videos with non-homogeneous haze. But it was still inefficient for large haze gradients.

Table 2.5 presents the comparative analysis of the discussed real-time video dehazing frameworks.

2.3 Research Gaps

After the extensive literature survey, the following research gaps were identified:

1) The majority of existing deep learning-based algorithms are applicable for synthetic images with less fog but are partially effective for images with real dense fog. Most of the existing datasets have synthetic foggy images. There is a need for a new dataset with roadside real-world images.

2) The existing algorithms proven effective for single images have some drawback such as color distortion, edge distortion, and halo effects.

3) The computational time is high for the existing real-time defogging frameworks. Also, the airlight effect is not considered for each individual frame.

4) There is no real-time web-based smartphone application for the drivers to rely on while driving through foggy areas to prevent accidents.

2.4 Research Objectives

Based on the state-of-the-art discussions and research gaps identification, following research objectives have been defined:

1) To study the existing deep learning-based defogging techniques and to collect the real-time dataset from the roadside.

2) To design a novel modified filter to enhance the coarse estimated atmospheric veil for roadside foggy images.

3) To develop a novel deep learning-based visibility restoration framework to quickly and efficiently remove the fog from road images to prevent accidents.

4) To develop a stand-alone and/or web-based application which helps the researchers and research community to remove fog from roadside still images.

2.5 Summary

This chapter presents a detailed comparative analysis of different defogging techniques, namely prior-based, deep learning-based, hybrid-based, and video defogging techniques, in terms of datasets used, knowledge gap, quality metrics, hyperparameters, prior knowledge, modules in architectures, and software or hardware system specifications. The research gaps are found in this analysis. They are followed by the research objectives of the study.

Table 2.2 Comparative Analysis of Deep Learning-based Defogging Techniques (Ref – Reference, CD – Color Distortion, ED – Edge Distortion, HE – Halo Effect, LHG – Large Haze Gradients, BA – Block Artifacts, EL – Excessive Lighting)

Algorithm	Ref	Year	Dataset(s)	Type of Architecture	Modules in Architecture	Remarks / Knowledge Gap	Quality Metrics					
							CD	ED	HE	LHG	BA	EL
Perceptron-based Neural Network	[141]	2016	FRIDA	DNN	• Normalization • Rasterization • Backpropagation • Weight Optimization	• No comparative analysis available.	-	✓	-	✓	✓	-
DehazeNet	[142]	2016	Internet Browsing	CNN	• Feature Extraction • Multi-scale Mapping • Local Extremum • Non-linear Regression	• Ineffective for dense hazy images.	✗	✗	✗	✓	✗	✓
Multi-scale CNN	[143]	2016	NYU	CNN	• Coarse-scale Network • Fine-scale Network	• Ineffective for nighttime images.	✗	✓	✗	✓	✗	✗
All-in-one dehazing network	[144]	2017	Middlebury; VOC	CNN	• K-Estimation • Clean Image Generation	• Parameters are not optimized for individual weather degrading variant.	✗	✗	✗	✓	✓	✗
Deep Fully Convolutional Regression Network	[145]	2017	NYU; Make 3D	CNN	• Downsampling • Upsampling	• Ineffective for nighttime images.	✗	✗	✓	✓	✓	✗
Cascaded Network	[146]	2018	Internet Browsing; NYU-V2	CNN	• Shared Hidden Layers • Global Atmospheric Light Estimation Subnetwork • Medium Transmission Estimation Subnetwork	• Although color distortion is reduced, the image noise is not suppressed.	✗	✗	✗	✓	✓	✓
Unsupervised Neural Network	[147]	2018	NYU	GAN	• Sparse Coding and Structure Feature Map • RGB, DCP, and Haze-lines Feature Map	• Ineffective for dense hazy images.	✗	✗	✗	✓	✗	✓

Table 2.2 (continue) Comparative Analysis of Deep Learning-based Defogging Techniques (Ref – Reference, CD – Color Distortion, ED – Edge Distortion, HE – Halo Effect, LHG – Large Haze Gradients, BA – Block Artifacts, EL – Excessive Lighting)

Algorithm	[Ref]	Year	Dataset(s)	Type of Architecture	Modules in Architecture	Remarks / Knowledge Gap	Quality Metrics					
							CD	ED	HE	LHG	BA	EL
Gated Fusion Network	[148]	2018	NYU	CNN	• White Balance • Contrast Enhancement • Gamma Correction	• Ineffective for dense hazy images.	✗	✓	✗	✗	✓	✓
Quad-tree Decomposition	[149]	2018	Internet Browsing	CNN	• Quad-tree Decomposition • Patch Classification • Initial Transmission • Refinement	• Depth is assumed constant due to which the color contrast in resultant images is irregular.	✓	✗	✓	✗	✗	✓
Remote Sensing Image	[150]	2018	VOC	R-CNN	• Classification Network • Object Detection Network	• Inadequate experimental details.	✓	✓	✗	✗	✓	✗
Cardinal Color Fusion	[151]	2018	ImageNet; D-Hazy	CNN	• Multi-channel Network • Multi-scale Network for each channel	• Computed images have poor texture and high noise.	✗	✗	✗	✓	✓	✓
Conditional GAN	[152]	2018	NYU; Make 3D	GAN	• Generator with Autoencoder • Discriminator with Classification	-	✗	✗	✓	✓	✗	✓
Gated Context Aggregation Network (GCANet)	[153]	2019	RESIDE	CNN	• Encoder • Smooth Dilated Residual Block • Gated Fusion • Decoder	• Parameters are not optimized for individual weather degrading variant.	✓	✗	✓	✗	✗	✓
Cycle GAN	[154]	2019	RESIDE	GAN	• Mask Computation using Generator 1 • Mask Computation using Generator 2	• Inadequate results.	✓	✗	✓	✗	✓	✓

Table 2.2 (continue) Comparative Analysis of Deep Learning-based Defogging Techniques (Ref – Reference, CD – Color Distortion, ED – Edge Distortion, HE – Halo Effect, LHG – Large Haze Gradients, BA – Block Artifacts, EL – Excessive Lighting)

Algorithm	Ref	Year	Dataset(s)	Type of Architecture	Modules in Architecture	Remarks / Knowledge Gap	Quality Metrics					
							CD	ED	HE	LHG	BA	EL
Atmospheric Illumination Prior Network	[155]	2019	IMAGENET ILSVRC2012	CNN	• Feature Extraction • Multiscale Restoration • Fusing Network • Combination Network	• Model is not optimized in YCrCb colorspace.	✗	✗	✗	✓	✓	✓
Feature Fusion Aggregation Network (FFA-Net)	[156]	2020	RESIDE	CNN	• Local Residual Learning • Feature Attention → Channel Attention + Pixel Attention	• Ineffective for dense hazy images.	✗	✓	✗	✓	✓	✗
Back Projected Pyramid Network	[157]	2020	I-HAZE; O-HAZE; NH-HAZE; Dense-Haze	CNN	• Multi-level Complexity Learning using Iterative U-Net Block • Multi-scale Spatial Learning using Pyramid Convolution Block	• Dehazed results suffer from color cast.	✓	✗	✗	✓	✗	✗
Light Convolution Autoencoder	[158]	2020	RESIDE	CNN	• Encoder • Decoder	• Only purpose of using an autoencoder is automatic feature extraction. • Inadequate Results.	-	-	-	✓	-	✓

Table 2.2 (continue) Comparative Analysis of Deep Learning-based Defogging Techniques (Ref – Reference, CD – Color Distortion, ED – Edge Distortion, HE – Halo Effect, LHG – Large Haze Gradients, BA – Block Artifacts, EL – Excessive Lighting)

Algorithm	Ref	Year	Dataset(s)	Type of Architecture	Modules in Architecture	Remarks / Knowledge Gap	Quality Metrics					
							CD	ED	HE	LHG	BA	EL
Multi-scale Boosted Dehazing Network (MSBDN)	[159]	2020	RESIDE; HAZE-RD, I-HAZE; O-HAZE	CNN	• Strengthen-Operate-Subtract Boosted Module • Residual Group • Dense Feature Fusion Module	-	✗	✗	-	✓	-	✓
Holistic Edges Multi-Scale Network	[160]	2020	Make 3D; KITTI	CNN	• Course Scale Network • Fine Scale Network • Holistic Edge Guided Network	• Enhanced version of Multi-scale CNN. Only improvement was edge Refinement.	✗	✗	✗	✓	✗	✓
Autoencoder-based Contrastive Learning (AECL)	[161]	2021	RESIDE	CNN	• Autoencoder-like Dehazing Network → Adaptive Mix-up for Feature Preserving + Dynamic Feature Enhancement • Contrastive Regularization	-	✓	✗	✗	✓	✓	✗
Discrete Wavelet Network	[162]	2021	RESIDE; NH-HAZE; NH-HAZE2	GAN	• Discrete Wavelet Transform Branch • Knowledge Adaptation Branch	• Ineffective for real-life hazy scenarios.	✗	✗	✗	✗	✓	✗
Energy Minimization	[163]	2021	Internet browsing	CNN	• Classification Network • Smog Gradient Prediction Model	• Spatial information is not considered.	-	-	-	-	-	-

Table 2.2 (continue) Comparative Analysis of Deep Learning-based Defogging Techniques (Ref – Reference, CD – Color Distortion, ED – Edge Distortion, HE – Halo Effect, LHG – Large Haze Gradients, BA – Block Artifacts, EL – Excessive Lighting)

Algorithm	Ref	Year	Dataset(s)	Type of Architecture	Modules in Architecture	Remarks / Knowledge Gap	Quality Metrics					
							CD	ED	HE	LHG	BA	EL
Smoggy Channel Prior (SCP)	[164]	2021	Internet browsing	CNN	• Smog Prediction Model • Oblique Gradient Profile Prior • Energy Minimization	• No novelty in smog removal technique.	✗	✗	✗	✓	✓	✗
Polarization via Cycle Neural Network	[165]	2021	Large polarization images	CNN	• Polarization degree image computation • Clear region recovery • Coarse clear image generation	• Ineffective for atmospheric smog. • Only relevant around source of smog pollution.	-	-	-	-	-	-
Attention-based Adaptive Network	[166]	2022	RESIDE; NYU2; NH-HAZE; O-HAZE; I-HAZE	CNN	• Multi-stage Architecture • Feature Extraction Module • Self-Calibration Attention Module • Adaptive Feature Enhancer Module	• Although texture details and semantic information are preserved, architecture is mathematically complex.	✗	✗	✗	✓	✓	✗
Dense Residual and Dilated Network	[167]	2022	NYU	CNN	• Dense Residual Network • Dilated Densely Connected Block	• Multiple skip connections make the model overfit for real-life scenarios.	✗	✓	✓	✗	✓	✗
Multi Axis Multi-layer Perceptron (MAMLP)	[168]	2022	RESIDE	CNN	• Main Backbone • Encoder / Decoder / Bottleneck • Cross Gating Block	• Inadequate results.	✗	✗	✗	-	-	-

Table 2.2 (continue) Comparative Analysis of Deep Learning-based Defogging Techniques (Ref – Reference, CD – Color Distortion, ED – Edge Distortion, HE – Halo Effect, LHG – Large Haze Gradients, BA – Block Artifacts, EL – Excessive Lighting)

End-to-end Dehaze and Detection	[169]	2023	ImageNet	CNN	• Encoder-Decoder Module • Detector Module → Convolution Down sampling Block + Spatial Pyramid Pooling Block	-	✗	✗	✗	✓	✓	✓
Progressive Encoding-Decoding network	[170]	2023	RESIDE; D-HAZY; I-HAZE; O-HAZE	CNN	• Feature Memory Module • Encoding-Decoding Module	-	✗	✗	✗	✓	✗	✓
TransConv Ensemble Reconstruction	[171]	2023	RESIDE; NH-HAZE; NH-HAZE 2; NH-HAZE 3	CNN	• TransConv Fusion Dehaze → Feature Attention + Gate Convolution + Selective Kernel Fusion • Lightweight Ensemble Reconstruction • Teacher Reconstruction Network	-	✗	✗	✗	✓	✓	✗

Table 2.3 Experimental Setup of Deep Learning-based Defogging Techniques (Ref – Reference, MSE – Mean Square Error, MAE – Mean Absolute Error, SGD – Stochastic Gradient Descent, RMSProp – Root Mean Square Propagation)

Ref	Programming Language	Processor	RAM (GB)	Hyperparameters				
				Optimizer	Initial Learning Rate	Epochs	Loss Function	Batch Size
[141]	-	-	-	-	-	-	MSE	-
[142]	MATLAB	Nvidia GeForce GTX 780 GPU	-		0.005	-	MSE	128
[143]	MATLAB	(Intel CPU 3.40 GHz	16	SGD	0.001	70	MSE	100
[144]	Pycaffe	Intel(R) Core (TM) i7-6700 CPU @3.40GHz	16	-	-	10	MSE	-
[145]	Caffe	Nvidia Tesla K80 GPU	-	SGD	0.0009	290	-	8
[146]	Python (TensorFlow)	a Intel(R) i7-6700 CPU @3.40GHz and a Nvidia GTX 1080 Ti GPU	-	Adam	0.001	-	MSE	32
[147]	MATLAB	Intel Core i7 CPU	16	-	-	-	-	-
[148]	-	Nvidia K80 GPU	-	Adam	0.0001	-	MSE	10
[149]	-	NVIDIA GeForce GTX TITAN XP GPU	-	Adam	0.00001	10	Cross Entropy	64
[150]	-	-	-	-	-	-	-	-
[151]	MATLAB	4.20 GHz Intel Core i7 processor + NVIDIA GTX 1080	8	SGD	0.002	18	MSE	64
[152]	Python (Torch7)	Nvidia Titan-XGPU + Intel(R) Core (TM) i7-6700 CPU @3.40GHz	-	Adam	0.002	500	-	-
[153]	-	4 GPU	-	Adam	0.01	100	MSE	12
[154]	-	-	-	-	-	-	MAE	-
[155]	Caffe	NVIDIA GeForce GTX 960 GPU + Intel 2.4 GHz CPU	-	RMSProp	0.001	54	MSE + MAE	4

Table 2.3 (continue) Experimental Setup of Deep Learning-based Defogging Techniques (Ref – Reference, MSE – Mean Square Error, MAE – Mean Absolute Error, SGD – Stochastic Gradient Descent, RMSProp – Root Mean Square Propagation)

Ref	Programming Language	Processor	RAM (GB)	Hyperparameters				
				Optimizer	Initial Learning Rate	Epochs	Loss Function	Batch Size
[156]	Pytorch	RTX 2080Ti GPU	-	Adam	0.0001	-	MSE	-
[157]	-	GPU	-	Adam	0.001	-	MSE + Adversarial Loss + Content Loss + Similarity Loss	-
[158]	-	-	-	Adam	0.001	28	MSE	-
[159]	-	NVIDIA 2080Ti GPU	-	Adam	0.0001	100	MSE	16
[160]	MATLAB	Intel CPU @3.40GHz	16	SGD	0.001	70	MSE	10
[161]	PyTorch 1.2.0 + MindSpore	NVIDIA TITAN RTX GPU	-	Adam	0.0002	100	MAE	16
[162]	-	Two NVIDIA 1080 Ti GPUs		Adam	0.0001	8000	Smooth Loss + Perceptual Loss + Similarity Loss + Adversarial Loss	16
[163]	MATLAB	Intel (R) Core (TM)i5	8	-	-	-	Laplacian Loss	-
[164]	MATLAB	Intel (R) Core (TM)i7	16	-	-	-	Laplacian Loss	-
[165]	Python	-	-	Adam	0.0001	300	-	32
[166]	PyTorch	NVIDIA RTX2080TI GPU.	-	Adam	0.0002	150	Charbonnier Loss	16
[167]	Tensorflow	Nvidia GTX1080Ti GPU	-	Adam	0.0001	30	L1 Loss + Gradient Loss + Similarity Loss	1
[168]	-	-	-	Adam	0.0002	-	Charbonnier Loss	-
[169]	Tensorflow2.0	NVIDIA GTX 1080Ti GPU	11	Adam	0.0001	40	Content Loss + Similarity Loss + Object Detection Loss	18
[170]	PyTorch	GeForce RTX 3090	-	Adam	0.0001	50	Similarity Loss	4
[171]	Pytorch	Two NVIDIA Titan X GPUs	12	Adam	0.0004	8000	Reconstruction Loss + Perceptual Loss + Similarity Loss + Standard Loss + Knowledge Distillation Loss	16

Table 2.4 Comparative Analysis of Hybrid Defogging Techniques (Ref – Reference, CD – Color Distortion, ED – Edge Distortion, HE – Halo Effect, LHG – Large Haze Gradients, BA – Block Artifacts, EL – Excessive Lighting, MSE – Mean Square Error)

Parameters	**Algorithms**											
	Ranking CNN						Residual-based Deep CNN					
Ref	[172]						[173]					
Year	2018						2018					
Dataset(s)	Self-generated						NYU-V2					
Prior Knowledge	Ranking layer facilitates the statistical attribute extraction						Airlight is different in different atmospheric conditions					
Modules in Architecture	**Map Estimation:** CNN; Random Forest regression **Map Refinement:** Guided Filter						**Map Estimation:** Classic convolution network **Residual Image:** Residual Network Element					
Remarks / Knowledge Gap	High computation time because transmission is estimated for every pixel of a local patch.						Not effective for real-world images					
Quality Metrics	**CD**	**ED**	**HE**	**LHG**	**BA**	**EL**	**CD**	**ED**	**HE**	**LHG**	**BA**	**EL**
	-	-	-	✓	✗	-	✗	✗	✗	✓	✗	✗
Programming Language	MATLAB						Caffe					
Processor	NVIDIA Geforce GTX980 GPU						-					
Hyperparameters	**Optimizer**	**Initial Learning Rate**	**Epochs**	**Loss Function**	**Batch Size**		**Optimizer**	**Initial Learning Rate**	**Epochs**	**Loss Function**	**Batch Size**	
	Stochastic Gradient Descent	0.01	100	Softmax	64		Stochastic Gradient Descent	0.001	-	MSE	-	

Table 2.5 Comparative Analysis of Real-time Defogging Techniques (Ref – Reference, CD – Color Distortion, ED – Edge Distortion, HE – Halo Effect, LHG – Large Haze Gradients, BA – Block Artifacts, EL – Excessive Lighting)

Software / Hardware	Ref	Year	Dataset(s)	Remarks / Knowledge Gap	Programming Language	Processor	RAM	Quality Metrics					
								CD	ED	HE	LHG	BA	EL
Optimized Contrast Enhancement	[174]	2013	Internet Browsing (https://www.flickr.com/)	• Computed frames suffer from flickering artifacts.	Single Instruction and Multiple Data Stream	Intel Core i5	4GB	✓	✗	✓	✓	✓	✗
Autoencoder-based Semantic Segmentation	[175]	2018	NYU-V2	• Only indoor dataset is used for training	-	NVidia K80 GPU	-	✓	✗	✗	✗	✗	✗
Minimum Filter	[176]	2020	Foggy Cityscapes; The Foggy Driving Dataset	• Pixel mapping with lookup table may increase computation time	-	Intel Core i5; Intel Core i7	-	✗	✗	✗	✓	✓	✗
Raspberry Pi 3	[177]	2021	Middlebury	• On chip processing unit can be damaged even due to small fault	MATLAB	Jetson Nano	-	✓	✗	✗	✗	✗	✓

Table 2.5 (continue) Comparative Analysis of Real-time Defogging Techniques (Ref – Reference, CD – Color Distortion, ED – Edge Distortion, HE – Halo Effect, LHG – Large Haze Gradients, BA – Block Artifacts, EL – Excessive Lighting)

Software / Hardware	Ref	Year	Dataset(s)	Remarks / Knowledge Gap	Programming Language	Processor	RAM	Quality Metrics					
								CD	ED	HE	LHG	BA	EL
VLSI-based architecture	[178]	2021	RESIDE; NYU-V2; Middlebury	• On chip processing unit can be damaged even due to small fault	-	-	-	✗	✗	✗	✓	✗	✗
Spatial-Temporal Coherent Regularization	[179]	2021	-	• It cannot estimate the transmission map in case of ambiguity between scene pixels and depth	C++	Intel Core i5	8GB	✓	✗	✓	✗	✗	✗
Dark Channel Prior	[180]	2021	-	• Computational complexity may be high	-	-	-	✓	✗	✓	✓	✗	✗
Confidence Guided - Improved Deformable Network	[181]	2021	REal-world VIdeo DEhazing (REVIDE)	-	-	NVIDIA TITAN RTX GPU	-	✓	✗	✗	✓	✗	✗

CHAPTER-3
HAZY UNPAIRED DATASET FOR ROAD SAFETY

This chapter discusses the challenges of existing datasets, followed by the qualitative and quantitative analysis of the different images computed using various visibility restoration models.

3.1 Challenges of Existing Datasets

The real-time dataset collection suffers from various challenges. These challenges are divided into three categories: environment-based factors, article-based, and algorithm-based challenges. The issues arise as a result of poor illumination conditions, the resemblance of weather conditions such as fog, haze, and smog, degraded image quality due to halo artifacts, gradient artifacts, edge degradation, and color distortion, detection of the objects present in dense fog, and detection of the objects hidden by bigger objects or not visible in a group of objects. Apart from the physical and environmental limitations, the existing algorithms have some technical limitations as well. To train a large dataset, hyperparameter tuning is a major concern in achieving higher accuracy. Thus, each color channel is processed separately in some dehazing techniques to improve the accuracy.

3.1.1 Environment-based Challenges

The environment-based challenges faced during the implementation of dehazing algorithms are as follows:

- **Illumination Conditions**

The distorted images are captured in different illumination conditions by using various image capture equipment such as stereo camera, visible light polarization camera, and single shot detector. Images captured with different cameras produce different results even if same algorithm is implemented. Some specifications of the camera such as focus and viewing geometry, are adjusted whenever a new site is selected. Due to poor illumination conditions, the dataset collection involves travelling to different locations.

- **Resemblance in Weather Conditions**

It is difficult to analyze an image captured in these distorted weather conditions. Sometimes there is an unavoidable resemblance in the weather conditions such as fog,

haze, and smog, which makes it difficult to determine the air pollutant with the human eye.

3.1.2 Article-based Challenges

The article-based challenges faced during the implementation of dehazing algorithms are as follows.

- **Degraded Image Quality**

The quality of hazy images gets worse because of halo artifacts (HA), gradient reversal artifacts (GRA), block artifacts (BA), edge distortion (ED), color distortion (CD), and large haze gradients (LSG). Thus, the significance of reconstructing the images by implementing dehazing algorithms is to restore the details of the distorted images. The $gamma$ parameter is used to adjust the degree of haze in the reconstructed image without distorting the quality of features.

- **Object Detection**

In poor weather conditions, the atmospheric light that reaches the surface of Earth is mainly scattered by vapours in the atmosphere. So, it is difficult to distinguish between the sky and non-sky regions. Also, the objects are mainly covered in white light, which makes it difficult to detect them. This drawback can be overcome by using object detection-based camera, such as LiDar and Radar.

- **Cluttering**

Cluttering refers to the collection of indiscriminately lying articles. Some salient articles get ignored during object segmentation due to cluttering. For instance, a two-wheeler vehicle is partially visible on the road when it is hidden behind bigger vehicles. To circumvent this limitation, multiple images or frames of the same articles are recorded at different angles.

- **Occlusion**

Occlusion refers to the accumulation of multiple articles such that a less useful articles present in the foreground hides some articles of interest lying in the background. Various computing techniques, such as object detection followed by its recognition and analysis are implemented to track the hidden objects.

3.1.3 Algorithm-based Challenges

Some algorithm-based challenges faced during the implementation of dehazing algorithms are as follows:

- **Training Datasets**

The dehazing techniques have increasingly used deep-learning based algorithms throughout the years. Real-time hazy image datasets are rarely introduced. The network is trained on synthetic hazy images as well as the smoggy images obtained through the internet. Thus, the existing deep-learning-based dehazing algorithms do not produce efficient results for real-world densely hazy images.

- **Multi-Channel Dehazing**

The default color format of an image is RGB, which means it comprises of red, green, and blue color channels. Sometimes the quality of the image is degraded due to variations in pixel values when image parameters such as brightness, contrast, texture and saturation are adjusted. Other color formats available are Hue Saturation Value (HSV), and YCbCr.

Table 3.1 summarizes the various challenges addressed during the implementation of dehazing techniques.

3.2 Proposed Dataset

The first step towards the creation of a dataset is to identify the problem. It helps to collect the valuable data. The algorithms implemented on the image dataset are categorized as classification and regression-based algorithms. The next step is to choose an appropriate time and place to capture the images. Hazy Unpaired Dataset for Road Safety (HUDRS) consists of real-world roadside images captured under foggy/hazy and clear weather. Real-time dense foggy or hazy images have been captured during the early morning or late night in the winter season. Also, clear images have been acquired during daylight for training the deep learning models.

Figure 3.1 depicts an overview of the process of development of HUDRS and its performance evaluation. The proposed dataset has been created in three stages, such as pre-capturing, capturing, and post-capturing [182]. Various steps for the establishment of HUDRS are presented in Table 3.2.

HUDRS database consists of 522 clear images acquired during the daylight when fog is clear. 1050 foggy images have been captured during the morning and evening. Fog-free images have been computed with different fog intensities. Images of road scenes have been captured on National Highways NH-1, NH-44, and NH-7 in Punjab, India. Data acquisition has been done during the winter season. Unlike existing datasets such as FRIDA and RESIDE, no external haze machine has been used to generate haze. Real-world foggy images have been taken under natural environmental conditions.

Table 3.1 Various challenges associated with visibility restoration of hazy images

Challenges	Causes	Solutions
Environment-based Challenges		
Illumination Conditions	Attenuation and scattering of atmospheric light affect different cameras in different manner	Adjustments in specifications of cameras such as brightness, contrast, and view geometry
Resemblance in Weather Conditions	Similar composition with common components such as water vapors and dust particles	Visibility range and density or thickness of the air pollutant distinguish them
Article-based Challenges		
Degraded Image Quality	Halo artifacts, gradient reversal artifacts, blocking artifacts, edge detection, color distortion, large haze gradients	Adjustment of the gamma parameter without distorting the features
Object Detection	Scattering of atmospheric light	Lidar and Radar cameras are used
Cluttering	Sky and non-sky regions are not distinguished accurately	Capture multiple images of same object at different angles
Occlusion	More useful articles are hidden among less useful articles	Analysis of all the objects to detect, and identify the object of interest
Algorithm-based Challenges		
Training Datasets	Majority of the algorithms are validated upon images with artificial haze	Creation of a new dataset with unpaired real and artificial haze which can be used to train the network
Multi-channel Dehazing	Different values of specification of the images such as brightness, contrast, and texture for different channels.	Use YCrCb-based color format in images as it reduces the trainable variables

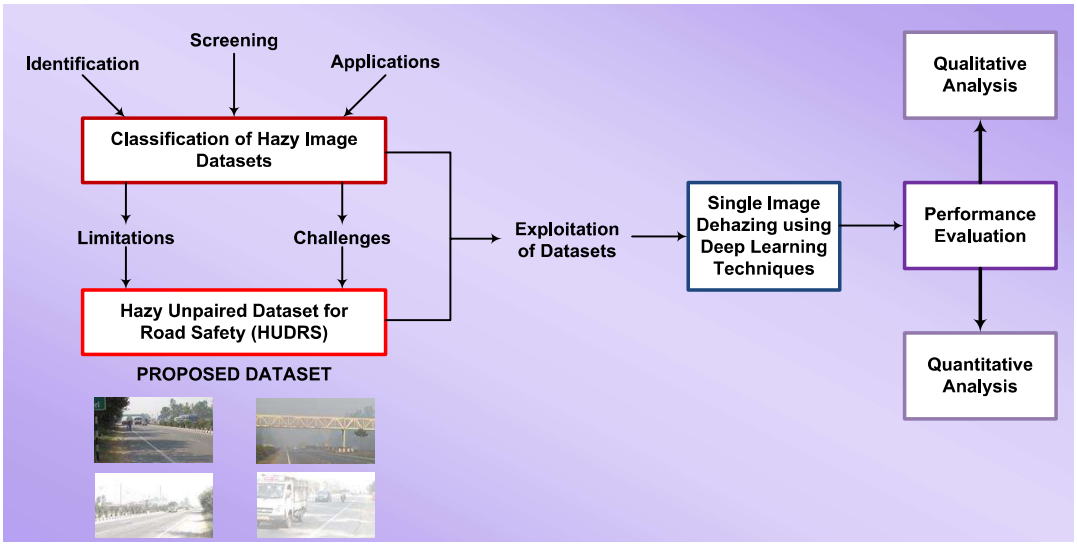


Fig. 3.1 Overview of development of HUDRS

Table 3.2 Steps for establishment of HUDRS

ALGORITHM: ESTABLISHMENT OF HUDRS	
Input: Identification of foggy/hazy region	
Output: Labelled dataset	
Step-1	Selection of an appropriate time and place
Step-2	Capture the images from the specified regions
Step-3	Check the quality of captured images
Step-4	Check the resolution of images captured at different spatial locations
Step-5	Remove the redundant images
Step-6	Resize the captured images
Step-7	Perform augmentation on the captured images
Step-8	Perform annotation on the images

Table 3.3 Technical specifications of HUDRS

Number of Pixels	16.0 MP
Color Filter Type	Primary Color
Image Processor	DIGIC 4+ with iSAPS
Focal Length	4-720 mm
Image Size	4608 × 3456
Still Image Type	JPEG compression
Color Matrix	RGB
White Balance	Through The Lens (TTL)
Approximate Pixel Pitch	1.34 microns

In terms of hardware, a Canon Power Shot SX400 has been used for image acquisition. The resolution of JPG images is 4608×3456. The technical details of image capture are shown in Table 3.3. Each image has been captured with manual adjustment of camera from roadside. The main aim of this chapter is to propose HUDRS for analyzing the dehazing algorithms for road safety. After capturing the images, their quality has been examined. The image quality can be degraded due to human error as well as technical problems. The captured images are considered as degraded due to the human error if they are blurred or have improper illumination conditions due to the improper placement of camera. The improper settings of the camera may lead to white balance, shutter speed, auto-focus, and exposure compensation. These images have not been considered in the dataset.

The remaining data has been converted into file formats for storing in a system. This helps to view and process multiple images at a time. The redundant images have been deleted as it causes overfitting or underfitting during the training of models in later stages. The data has been rescaled to minimize the computation time of processing the training data. It has been augmented to increase the size of dataset for better accuracy of the restoration model. The training dataset has been prepared by inserting synthetic fog of different intensities in clear ground truth images. Four different versions with synthetic fog have been produced by processing each clear image. Finally, the data has been annotated or labelled to train the model [183]. This is also called pre-processing of images.

3.3 Performance Evaluation of Existing Datasets

Some hazy images of HUDRS have been analyzed using the existing dehazing algorithms namely FFA-Net, CEP, GCANet, ICAP, NLBF, MSBDN, CCDI, and AECL. Figure 3.2 depicts the results obtained from different dehazing algorithms over IMSET-1 of HUDRS (i.e., IMGN1 to IMGN6). Sample images from various datasets, namely DENSE-HAZE (i.e., IMGN7), O-HAZE (i.e., IMGN8), RESIDE (i.e., IMGN9), NH-HAZE (i.e., IMGN10), and HUDRS (Synthetic) (i.e., IMGN11) have been evaluated using state-of-the-art single image deep learning-based techniques such as FFA-Net, GCANet, MSBDN, and AECL as depicted in Figure 3.3. Also, sample images from various datasets, namely FRIDA (i.e., IMGN12), HAZERD (i.e., IMGN13), D-

HAZY (i.e., IMGN14), MRFID (i.e., IMGN15), and HUDRS (i.e., IMGN16) have been evaluated by implementing prior-based dehazing techniques such as CEP, ICAP, NLBF, and CCDI as depicted in Figure 3.4.

3.3.1 Qualitative Analysis

As observed in Figure 3.3, the images of IMSET-2A are partially defogged when FFA-Net is implemented. Also, this technique is not applicable for night images as in IMGN8. The defogged images of IMGN7, IMGN9, IMGN10, and IMGN11 obtained from GCANet suffer from color distortion as atmospheric light has not been estimated in this algorithm. When IMGN9 is processed, the white-balance is affected because the computed image has some traces of fog.

When MSBDN is implemented to obtain the dehazed images, the traces of fog particles are visible in the IMGN7 and IMGN8 images. The results obtained from AECL are much closer to the ground-truth images as compared to FFA-Net, GCANet, and MSBDN. Figure 3.4 depicts the visual representation of results computed for IMSET-2B. It may be noted that real-world images IMGN15 and IMGN16 are evaluated for MRFID and HUDRS, respectively, so ground truth images are not available. The results obtained from CEP have no halo artifact. However, defogged images have random color patches in the computed images that results in the loss of information.

ICAP reduces the color distortion and edge distortion in defogged images. However, it is not applicable for images with heavy fog, as in IMGN13, IMGN15, and IMGN16. IMGN13 and IMGN15 suffer from the color distortion and edge distortion, respectively, when NLBF is implemented. CCDI distributes the effect of blurring near edges, but it increases the noise in the image. Hence, the evaluated algorithms suffer from the color shifting and structural artifacts.

3.3.2 Quantitative Analysis

The quantitative evaluation of state-of-the-art techniques based on $BRISQUE$, β, and $\mathbb{P}$ is shown in Tables 3.4 and 3.5. In Table 3.4, AECL performs better than other algorithms when $BRISQUE$ is calculated for IMGN7. Similarly, MSBDN gives better result for IMGN8 and GCANet gives better result for IMGN11. When images of IMGN9 and IMGN10 are evaluated, FFA-Net performs better than other algorithms. In Table 3.5, ICAP performs better than other algorithms for images of IMGN12,

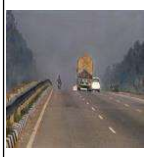
Dehazing Algorithm	IMGN1	IMGN2	IMGN3	IMGN4	IMGN5	IMGN6
Haze Image						
FFA-Net						
CEP						
GCANet						
ICAP						
NLBF						
MSBDN						
CCDI						
AECL						

Fig. 3.2 Image restoration of IMSET-1 using different dehazing algorithms over HUDRS dataset

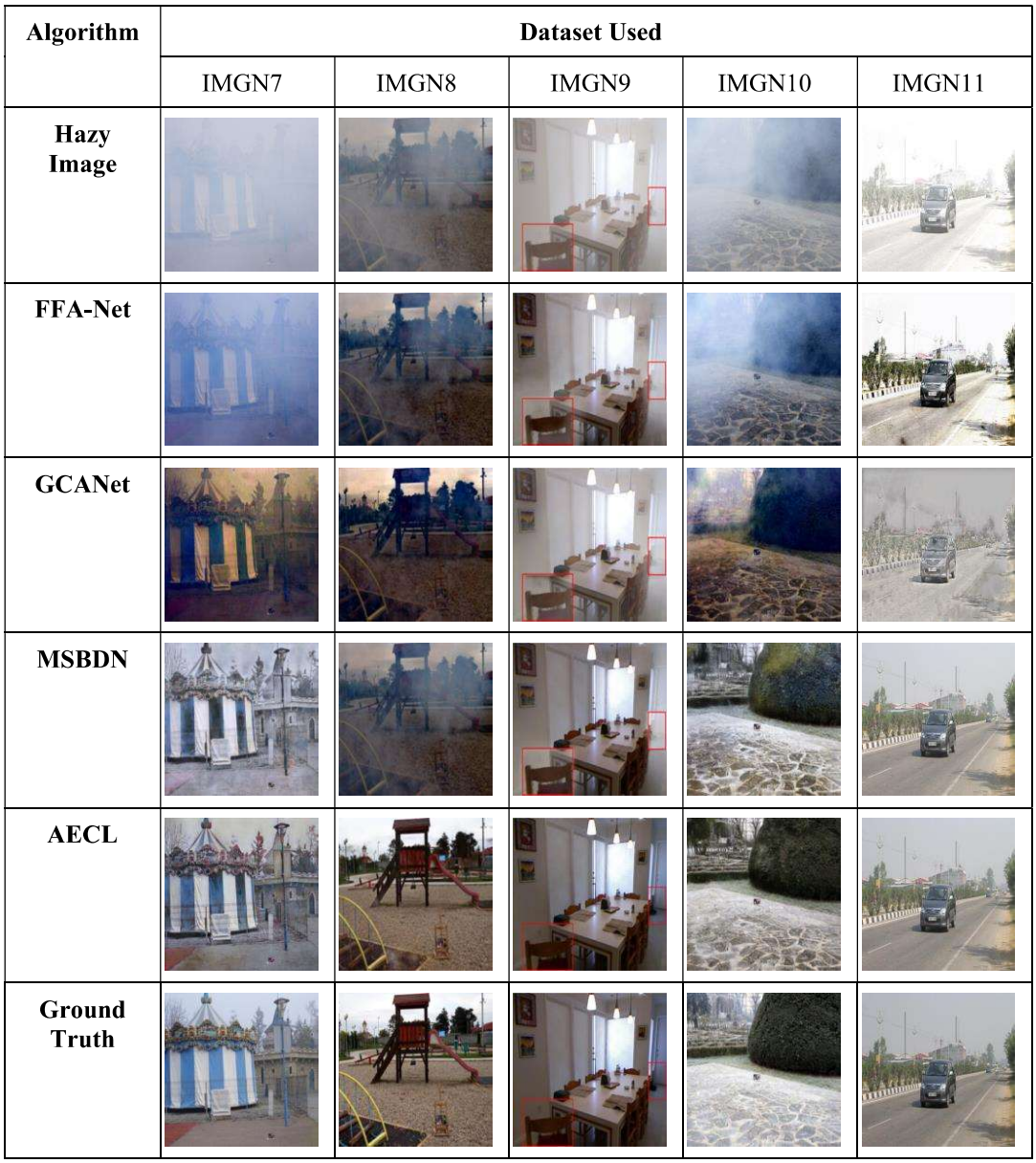


Fig. 3.3 Qualitative analysis of sample images from IMSET-2A using different deep learning-based dehazing techniques

IMGN13, IMGN14, and IMGN16 when $BRISQUE$ is calculated. For IMGN15, CEP performs better than other algorithms.

Figures 3.5 – 3.12 depict the graphical representation of $SSIM$, $PIQE$, $PSNR$, and IE of various images. It has been observed that AECL is better than other algorithms whenthe performance measures such as $SSIM$ and $PSNR$ are calculated. $PIQE$ is small when synthetic images are evaluated by using CEP. ICAP performs better than other algorithms when images with real haze are evaluated. Also, none of the techniques perform better than others when IE is calculated.

Algorithm	Dataset Used				
	IMGN12	IMGN13	IMGN14	IMGN15	IMGN16
Hazy Image					
CEP					
ICAP					
NLBF					
CCDI					
Ground Truth				-	-

Fig. 3.4 Qualitative analysis of sample images from IMSET-2B using different prior-based dehazing techniques

3.4 Summary

In this chapter, various phases of creating Hazy Unpaired Dataset for Road Safety (HUDRS) have been discussed. The foggy images have been captured during morning and evening, while fog-free images have been captured during day-time on NH-1, NH-7, and NH-44, Punjab, India. The resolution of the captured images is 4608×3456. Various pre-processing operations, such as the elimination of duplicate images and resizing of images have been applied to the dataset.

Table 3.4 Quantitative evaluation of sample images from IMSET-2A

Dataset	Performance Metric	Dehazing Algorithms			
		FFA-Net	**GCANet**	**MSBDN**	**AECL**
IMGN7	$BRISQUE$	40.6418	42.9772	43.8292	**39.5162**
	$\beta 1$	145.4131	78.3549	151.0388	**127.2563**
	$\mathbb{P}$	0	1.3298	0	0
IMGN8	$BRISQUE$	47.5166	47.4255	**46.5967**	48.8980
	$\beta 1$	76.5246	65.2379	83.0868	**105.6646**
	$\mathbb{P}$	0.1068	8.2710	0	0.0183
IMGN9	$BRISQUE$	**36.8644**	37.3202	39.1106	42.0460
	$\beta 1$	**137.4352**	152.4871	118.3355	99.1131
	$\mathbb{P}$	0.0645	2.7943	0.5791	0.0023
IMGN10	$BRISQUE$	**39.0290**	44.9085	45.2065	44.4088
	$\beta 1$	**138.5185**	80.9585	121.2033	111.1342
	$\mathbb{P}$	0.0320	10.6281	0	0
IMGN11	$BRISQUE$	32.8307	**23.9184**	36.2444	36.3585
	$\beta 1$	199.1086	173.7866	170.8651	**160.8732**
	$\mathbb{P}$	0.5688	0.0076	0.2529	0.0771

Table 3.5 Quantitative evaluation of sample images from IMSET-2B

Dataset	Performance Metric	Dehazing Algorithm			
		CEP	**ICAP**	**NLBF**	**CCDI**
IMGN12	$BRISQUE$	36.1425	**20.4358**	41.1473	35.7833
	$\beta 1$	106.3229	**119.7097**	99.7314	114.8621
	$\mathbb{P}$	22.1771	0.0065	37.1449	22.4232
IMGN13	$BRISQUE$	39.2929	**37.3809**	40.2685	41.7534
	$\beta 1$	132.0477	**134.4472**	129.9938	132.9884
	$\mathbb{P}$	0.0656	5.4195	1.7944	1.1341
IMGN14	$BRISQUE$	30.9514	**22.4359**	29.5633	32.4068
	$\beta 1$	144.1214	**138.1102**	129.3620	152.1807
	$\mathbb{P}$	13.7440	0	0.4990	1.4526
IMGN15	$BRISQUE$	**43.4801**	50.1187	44.3742	57.5795
	$\beta 1$	61.1443	**119.5205**	73.4737	75.6214
	$\mathbb{P}$	0.1587	0	4.5765	2.8263
IMGN16	$BRISQUE$	38.4893	**27.3713**	38.2332	39.7492
	$\beta 1$	102.5987	**133.7707**	82.1732	122.6038
	$\mathbb{P}$	0.0877	0	5.5809	0.3052

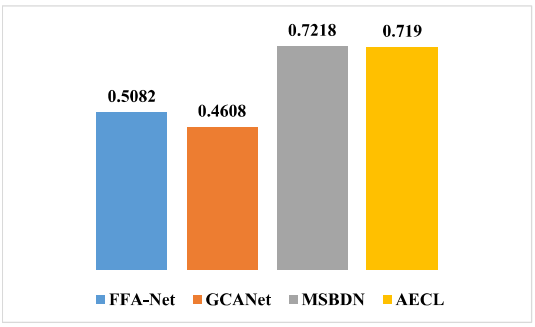
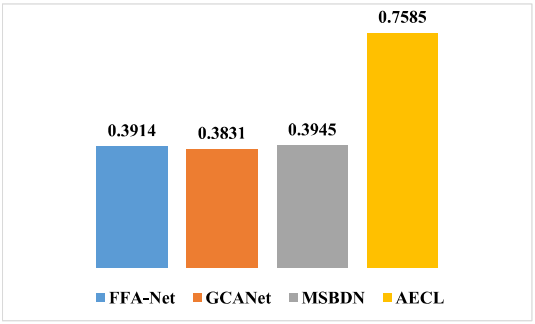
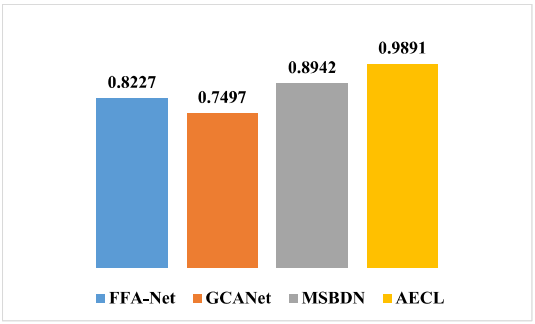
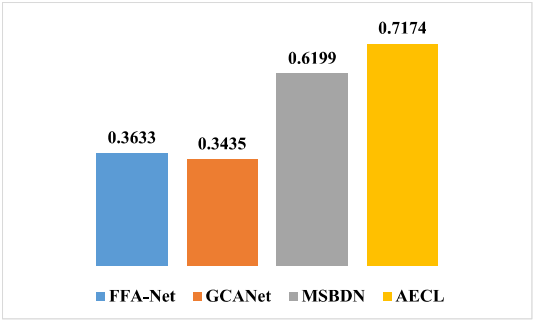
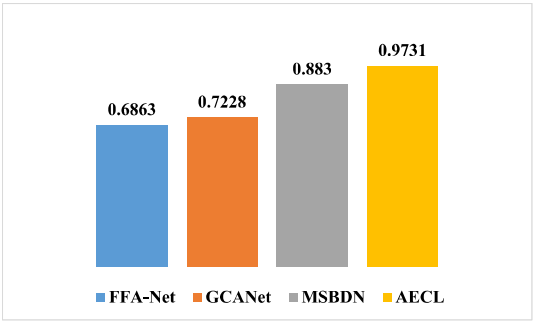
IMGN7	
IMGN8	
IMGN9	
IMGN10	
IMGN11	

Fig. 3.5 Quantitative analysis of sample images from various datasets of IMSET-2A using different dehazing techniques by evaluating *SSIM* of images

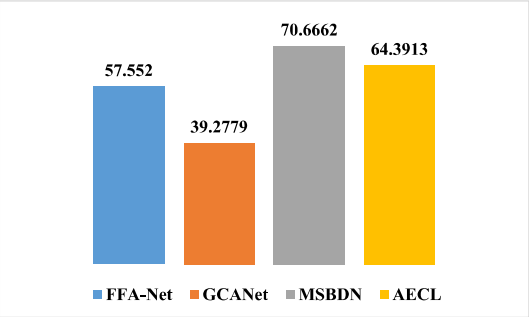
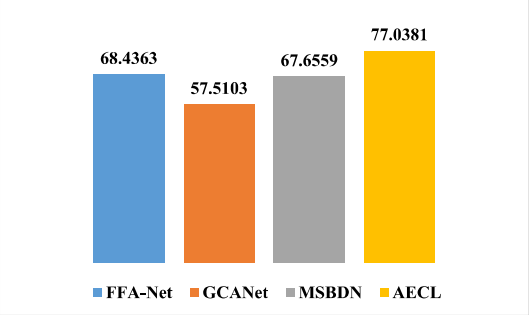
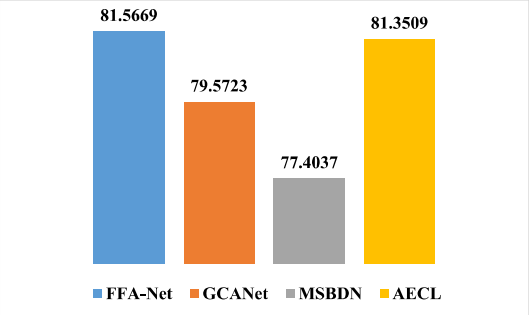
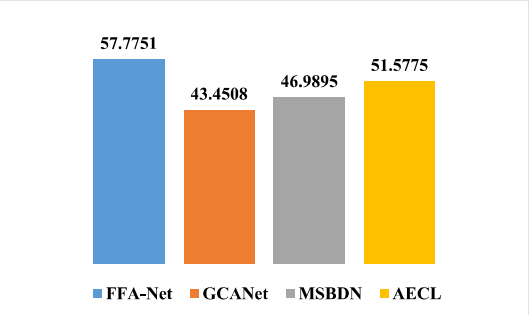
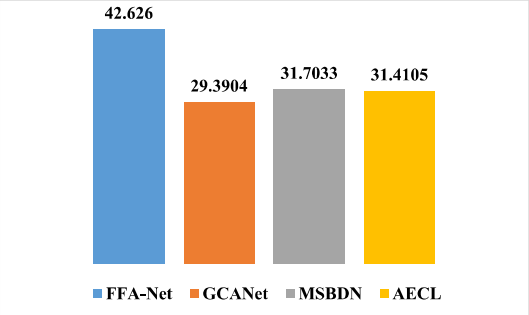
IMGN7	
IMGN8	
IMGN9	
IMGN10	
IMGN11	

Fig. 3.6 Quantitative analysis of sample images from various datasets of IMSET-2A using different dehazing techniques by evaluating *PIQE* of images

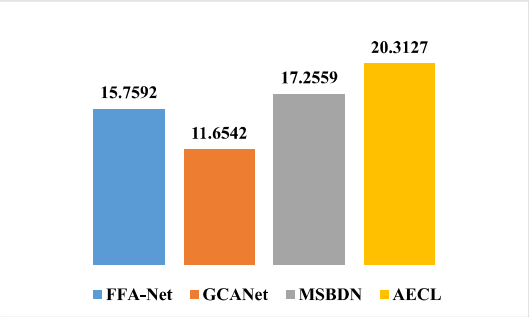
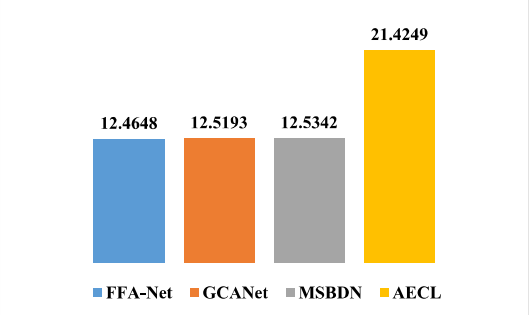
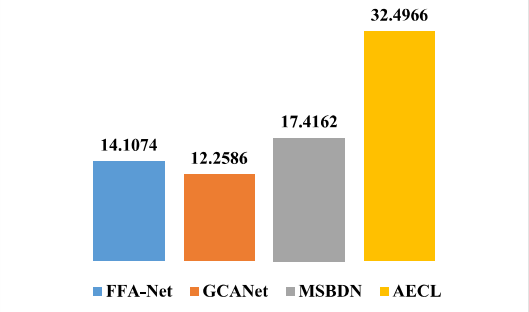
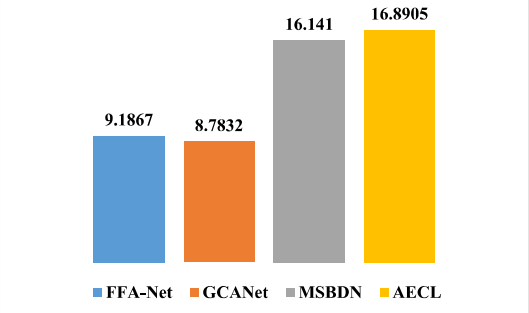
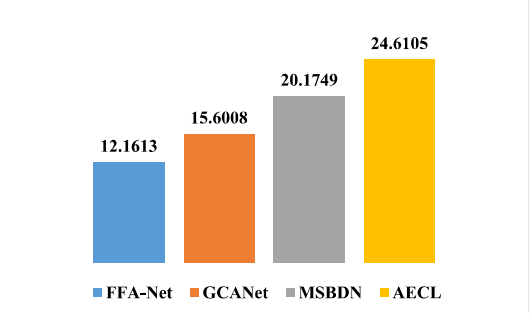
IMGN7	
IMGN8	
IMGN9	
IMGN10	
IMGN11	

Fig. 3.7 Quantitative analysis of sample images from various datasets of IMSET-2A using different dehazing techniques by evaluating *PSNR* of images

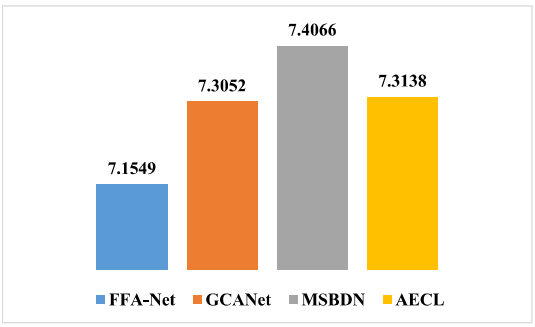
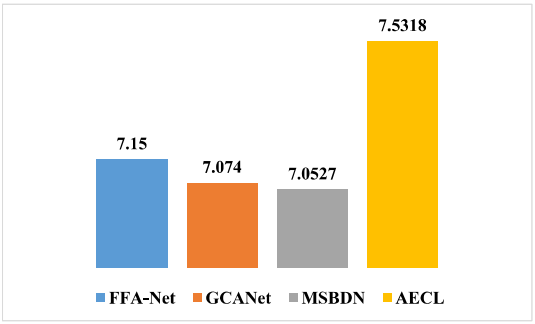
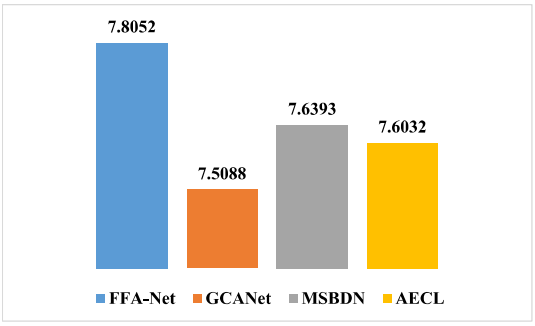
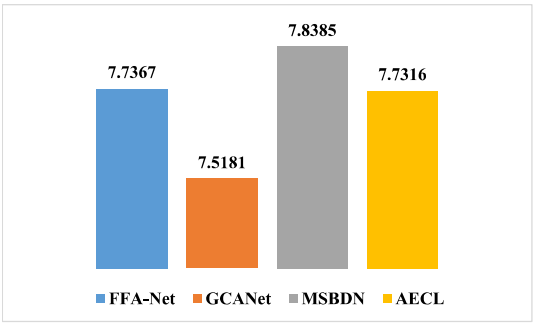
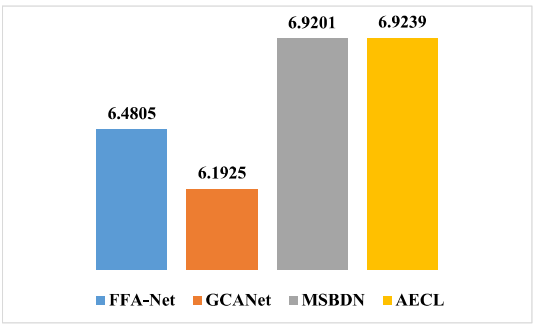
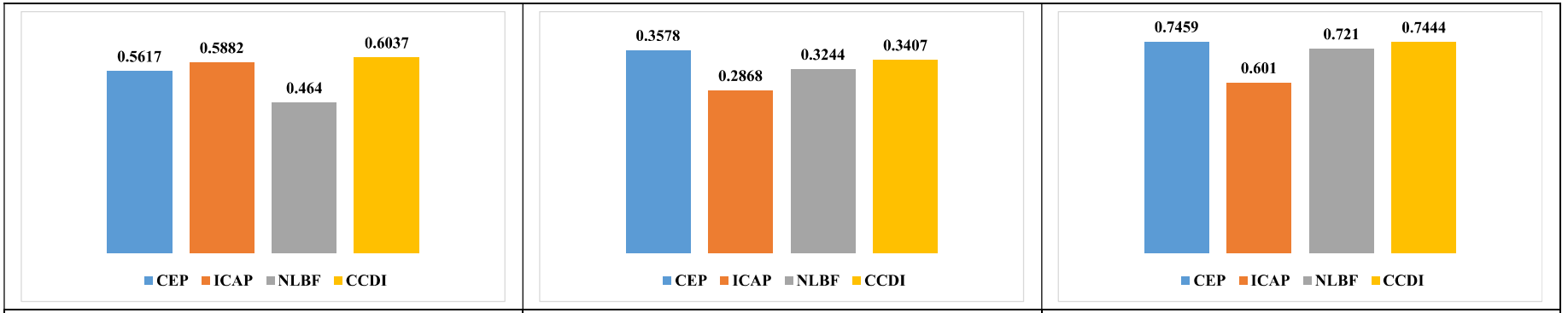
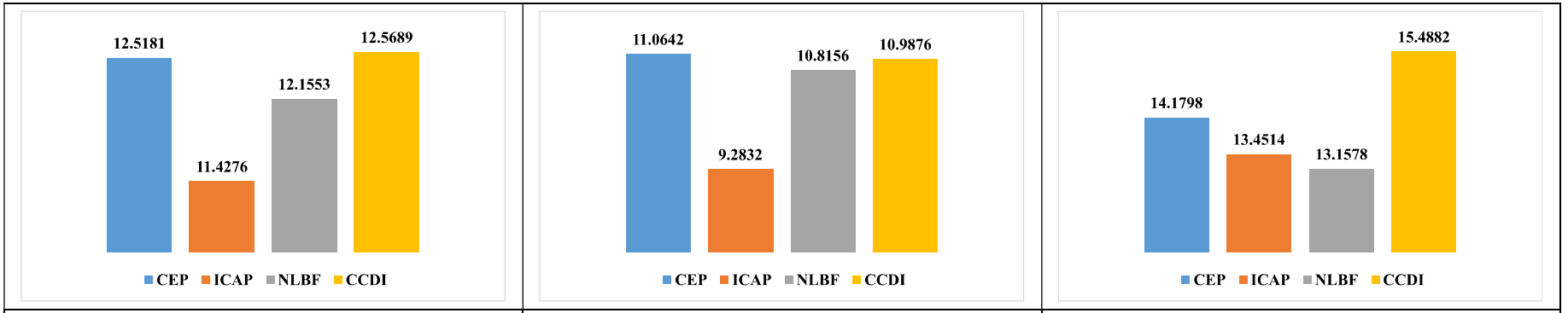
IMGN7	
IMGN8	
IMGN9	
IMGN10	
IMGN11	

Fig. 3.8 Quantitative analysis of sample images from various datasets of IMSET-2A using different dehazing techniques by evaluating *IE* of images



IMGN12	IMGN13	IMGN14

Fig. 3.9 Quantitative analysis of sample images from various datasets of IMSET-2B using different dehazing techniques by evaluating *SSIM* of images



IMGN12	IMGN13	IMGN14

Fig. 3.10 Quantitative analysis of sample images from various datasets of IMSET-2B using different dehazing techniques by evaluating *PSNR* of images

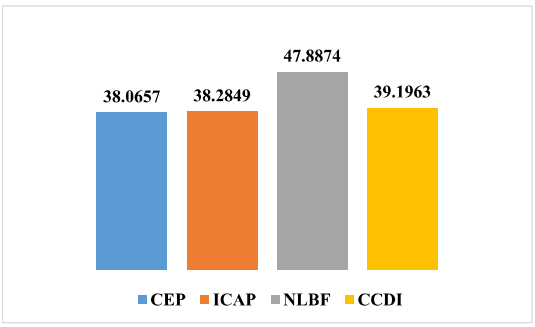
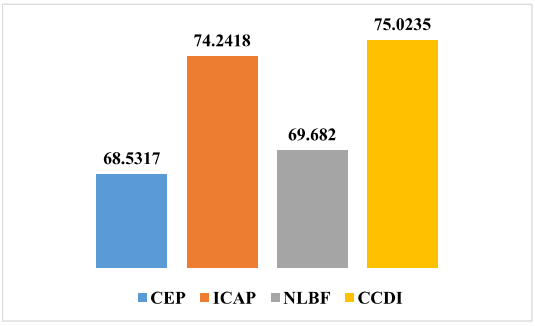
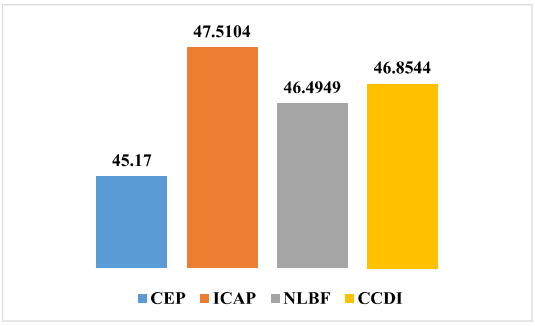
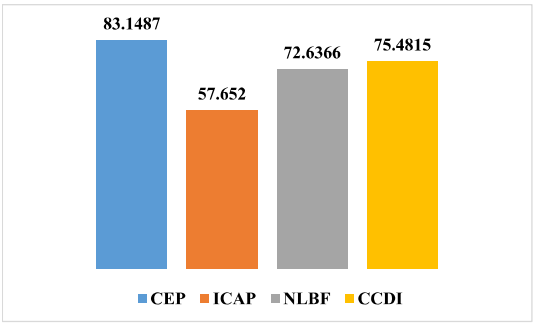
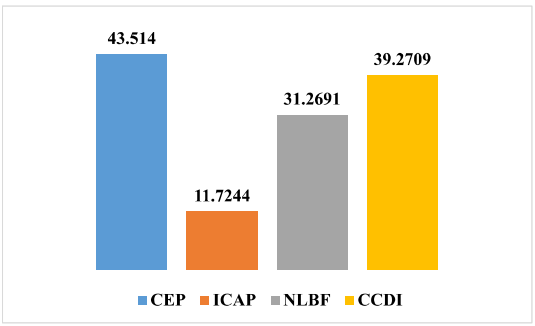
IMGN12	
IMGN13	
IMGN14	
IMGN15	
IMGN16	

Fig. 3.11 Quantitative analysis of sample images from various datasets of IMSET-2B using different dehazing techniques by evaluating *PIQE* of images

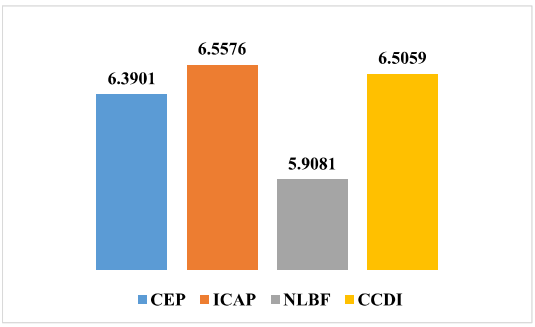
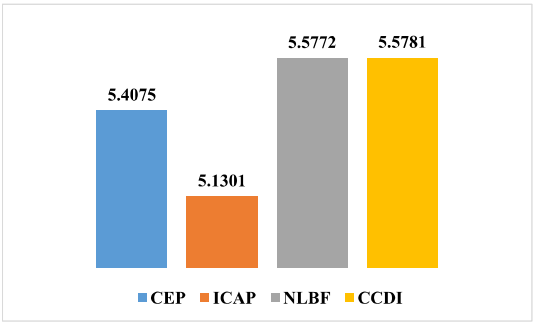
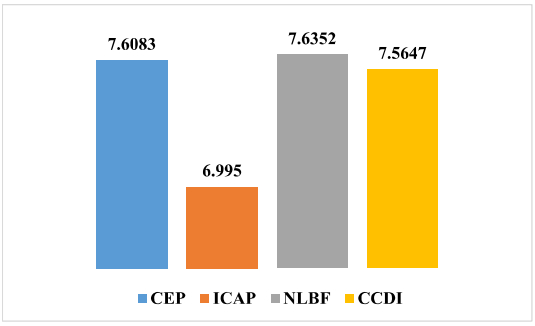
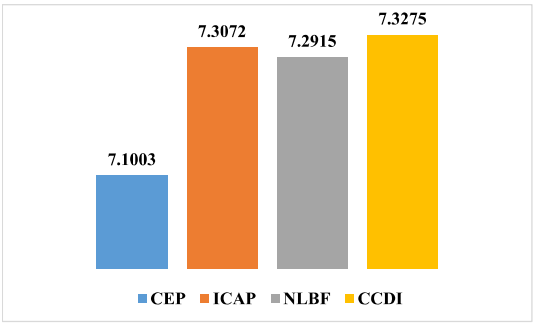
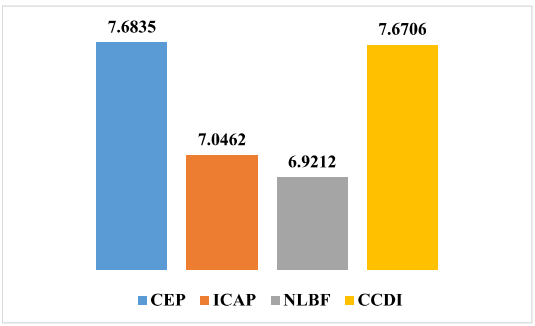
IMGN12	
IMGN13	
IMGN14	
IMGN15	
IMGN16	

Fig. 3.12 Quantitative analysis of sample images from various datasets of IMSET-2B using different dehazing techniques by evaluating *IE* of images

The size of dataset has been increased by introducing haze of four different densities in the fog-free images. Data augmentation operations, such as cropping and rotation also increase the size of dataset. The hazy images have been captured without utilizing any external haze-producing machines. The complete dataset has been annotated by using four different versions of hazy images. The comparative analysis of dehazing algorithms has been performed on sample images from the existing hazy datasets, namely FRIDA, HAZERD, D-HAZY, DENSE-HAZE, O-HAZE, RESIDE, NH-HAZE, MRFID, and the proposed HUDRS dataset.

Also, the images of FRIDA, HAZERD, D-HAZY, MRFID, and real-world hazy images of HUDRS have been evaluated by implementing prior-based dehazing algorithms. Furthermore, the images of DENSE-HAZE, O-HAZE, RESIDE, NH-HAZE, and synthetic hazy images of HUDRS have been evaluated by implementing deep learning-based dehazing algorithms. Experimental results reveal that the majority of existing dehazing techniques exploited the synthetic datasets. Overall, none of the techniques performed better than the overall set of images.

CHAPTER-4
SINGLE IMAGE DEHAZING

In this chapter, a metaheuristic algorithm and a deep learning-based technique for single image dehazing are discussed.

4.1 Grey Wolf Optimization

Metaheuristic techniques are used to find near-optimal solutions with ease of implementation. This chapter presents a metaheuristic technique called grey wolf optimization to perform single image dehazing.

4.1.1 Motivation

Grey wolves are characterized as elite predators, inhabiting the highest trophic level in the hierarchy. They prefer to reside in social teams known as packs. The average number of members ranges from five to 12. Furthermore, they exhibit an immensely rigid social dominance hierarchy.

The leaders of the group responsible for decision-making related to hunting, selection of sleeping locations, and wake-up times are called alphas. Their decisions are enforced upon the pack. Sometimes, an alpha wolf exhibits a tendency to follow other members of the pack. It is responsible for issuing commands that must be obeyed by the pack [184]. Although, it is not the strongest member, it excels at managing the pack. This highlights the significance of organization and discipline in a pack, which surpasses the importance of its strength.

The beta rank represents the second tier in the hierarchy. The beta wolves assist the alpha in group activities. This wolf, regardless of gender, is a strong candidate for accepting the alpha position in the event of the alpha wolf's death or advanced age. It exhibits deference towards the alpha while also exerting authority over subordinate wolves. It assumes the roles of both instructor of the team and advisor to the alpha. It provides support to the alpha by reinforcing their commands within the pack and offering feedback.

The omega members are the lowest-ranking individuals in a grey wolf pack. They fulfil the role of a scapegoat. They consistently exhibit subordinate behavior towards the higher-tier wolves. The omega, despite being insignificant, plays a crucial role in

maintaining pack cohesion. Research has shown that the absence of an omega can lead to internal conflicts and disruptions within the pack. The venting of violence and disappointment by the omega(s) leads to this outcome. This behavior helps to establish and maintain the hierarchical structure within the group.

In wolf packs, delta wolves exhibit submission to alphas and betas, while asserting dominance over the omega. They are either elders or weaklings, and are assigned the jobs of observers, guardians, and hunters. They monitor the territory and alert the fellow wolves to potential threats. The hunters are the defenders of the pack. They assist the dominant and subordinate members of the pack by participating in prey hunting and food provisioning activities. The elder wolves have previously held the superior position. The caretakers attend the debilitated, ailing, and wounded wolves within the pack. Grey wolves demonstrate collective hunting as a prominent social behavior, associated with their social hierarchy. Muro et al. [185] identified the following fundamental phases involved in the hunting behavior of grey wolves:

- The act of chasing and striving for prey is a fundamental behavior observed in predatory species.
- The predator engages in a tactical approach of encirclement and persistent harassment of the prey until it discontinues its movement.
- Predatory assault.

4.1.2 Proposed Methodology

In this section, the grey wolf algorithm has been used to estimate the airlight and transmission map for visibility restoration.

- **Estimation of Airlight**

In the proposed algorithm, an order statistic filter has been employed for airlight estimation. The hazy image has been represented as a two-dimensional matrix. The color channel with the least number of non-zero pixels has been selected. Its pixels have been arranged in ascending order. Each pixel has been substituted with the corresponding information from the sequential index in the patch.

- **Grey Wolf Optimization Map**

The transmission map has been produced by adjusting the pixels' intensity to the relevant positions of the wolves namely, alpha wolf, beta wolf, and delta wolf.

Firstly, the control parameters namely, $alpha\_factor$ and $beta\_factor$ have been initialized as 0.5 and 0.8, respectively, along with the dimension and threshold boundary of the feature map to be estimated. In GWO, $alpha\_factor$ and $beta\_factor$ are the distance of prey from alpha wolf and beta wolf, respectively. This distance has been considered in the range [0,1], such that 0 means closest to the prey, and 1 means farthest from the prey. So, $alpha\_factor$ has been considered as "0.5." The value of $beta\_factor$ has been varied from 0.6 to 0.9, but the optimal value of 0.8 had been obtained. Secondly, the population size and position of the wolves have been initialized. The positions of alpha wolf and beta wolf have been directed towards the $alpha\_factor$ and $beta\_factor$ because they are considered as best and second-best solutions, respectively.

The initial positions of the wolves are defined as follows:

$$pos = lb + (ub - lb) \times pop\_size \tag{4.1}$$

where pos is the position vector of the wolves. lb and ub are lower and upper boundary of the threshold. pop_size is the population size, i.e., number of wolves considered for the algorithm.

The elements of pos are considered as $alpha\_pos$, $beta\_pos$, $delta\_pos$, and so on. The $alpha\_pos$ has been updated using fitness function and has been directed towards $alpha\_factor$. Similarly, the $beta\_pos$ has been updated using fitness function and has been directed towards $beta\_factor$. The different edge filters such as Sobel, Prewitt, and fibermetric filters, had been tested, but the performance of GWO had not been found better in terms of qualitative or quantitative analysis. In this chapter, the "canny" edge filter has been used as fitness function [186]. The fitness value has been calculated as sum of pixels of the edge filtered map.

Another vector 'a' has been used to introduce stochastic behavior during the position update. It is calculated as follows [187]:

$$a = 2 \times alpha\_pos * (rand\_vector)_{1 \times dim} - alpha\_pos \tag{4.2}$$

where $rand\_vector$ is a random vector of dimensions $(1 \times dim)$.

The randomness and variation have been introduced in the current position of alpha and beta wolves, namely cur_pos_alpha and cur_pos_beta, respectively.

It is mathematically represented as:

$$cur\_pos\_alpha = (2 \times alpha\_pos \times (rand\_vector)_{1\times dim}) - alpha\_pos \quad (4.3)$$

$$cur\_pos\_beta = (2 \times beta\_pos \times (rand\_vector)_{1\times dim}) - beta\_pos \quad (4.4)$$

These positions are further updated using initial and current positions as follows.

$$ud\_pos\_alpha = |(cur\_alpha\_pos \times alpha\_pos) - pos\{wolf\ number\}| \quad (4.5)$$

$$ud\_pos\_beta = |(cur\_beta\_pos \times beta\_pos) - pos\{wolf\ number\}| \quad (4.6)$$

$$ud\_pos\_delta = |alpha\_pos - pos\{wolf\ number\}| \quad (4.7)$$

where ud_pos_alpha, ud_pos_beta, and ud_pos_delta are updated positions of alpha, beta, and delta wolves, respectively. $pos\{wolf\ number\}$ is the initial position from the population of wolves.

The final position of wolves (fin_pos) has been calculated using the updated positions of the wolves as follows:

$$fin\_pos = (alpha\_pos - (a \times ud\_pos\_alpha)) + (beta\_pos - (a \times ud\_pos\_beta)) + (delta\_pos - (a \times ud\_pos\_delta)) \quad (4.8)$$

The final map has been estimated by comparing the pixel intensity of the normalized grayscale hazy image and fin_pos. The generated transmission map (t) includes the pixel intensities greater than the elements of fin_pos.

- **Visibility Restoration**

The hazy image has been restored by constraining the illumination, referred to as gamma, after estimating the airlight and transmission map. Assuming t_0 as minimum desired intensity of light, the restored image is produced using Eq. (4.9):

$$J(x,y) = \frac{I(x,y) - A}{[max(t(x,y), t_0(x,y))]^{gamma}} + A \quad (4.9)$$

Table 4.1 presents the pseudo code of grey wolf algorithm-based dehazing algorithm.

4.1.3 Qualitative Analysis

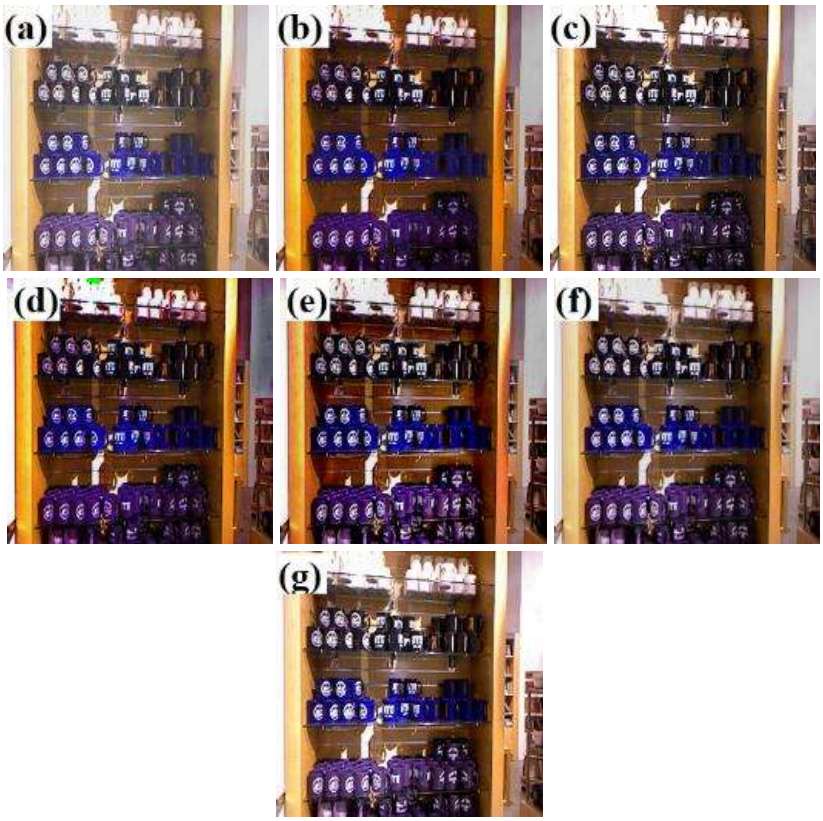
This section presents the performance evaluation of grey wolf optimization

algorithm and other existing techniques on IMSET-3 using qualitative and quantitative analysis. Figure 4.1 depicts the results of three synthetic foggy images namely, IMGN17, IMGN18, and IMGN19, obtained from RESIDE dataset. Also, four real foggy images namely, IMGN20, IMGN21, IMGN22, and IMGN23 obtained from HUDRS dataset have been exploited. The images IMGN17 through IMGN23 have been computed using the proposed technique and other existing dehazing techniques namely, CAP, ICAP, CEP, and SPD. Color distortion is a prevalent issue observed in the outcomes obtained by the utilization of CEP.

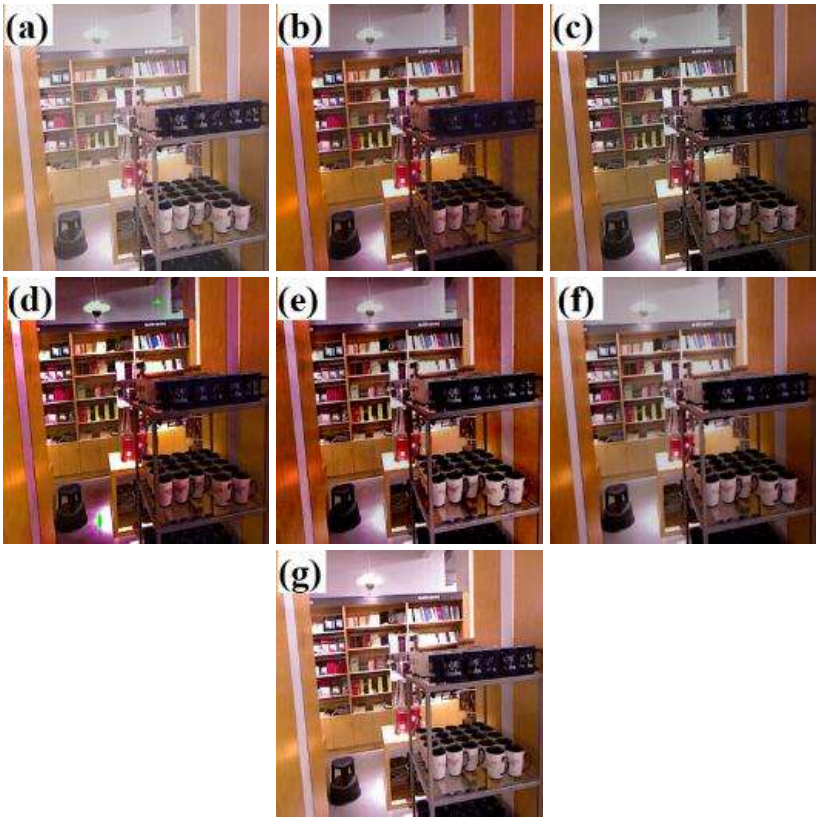
Also, the edges and texture of the results produced using SPD are distorted as compared to other techniques. The pixels' intensity tends towards zero when images are computed using CAP and ICAP. The images are dehazed up to the desired range of

Table 4.1 Pseudo code grey wolf algorithm-based single image dehazing algorithm

ALGORITHM: GREY WOLF OPTIMIZATION ALGORITHM	
Input: Hazy image	
Output: Transmission map of hazy image	
Step-1	Initialize population of wolves (alpha, beta, and delta) with upper and lower thresholds
Step-2	Initialize the dimension, $alpha\_factor$, and $beta\_factor$ for transmission map estimation
Step-3	for iteration = 1 to $maximum\_iterations$
Step-4	for i = 1 to $population\_size$ Evaluate fitness of current wolf using canny edge filter
Step-5	if iteration == 1 or fitness > best fitness Update best fitness and best wolf position end
Step-6	Calculate distances to alpha, beta, and delta wolves based on fitness Update position of current wolf using alpha, beta, and delta positions Ensure position stays within map boundary
Step-7	end % End of inner loop for population
Step-8	end % End of outer loop for iterations
Step-9	Obtain the best transmission map and its corresponding fitness value

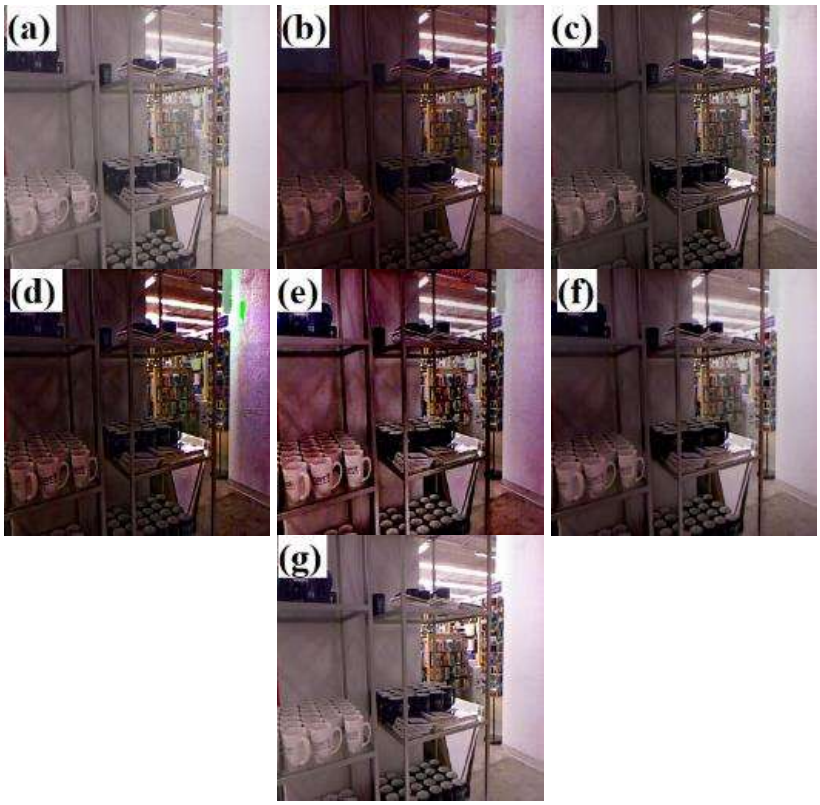


(i) IMGN17



(ii) IMGN18

Fig. 4.1 Qualitative analysis of images obtained from IMSET-3 (a) hazy image (b) CAP (c) ICAP (d) CEP (e) SPD (f) grey wolf (proposed) (g) ground truth



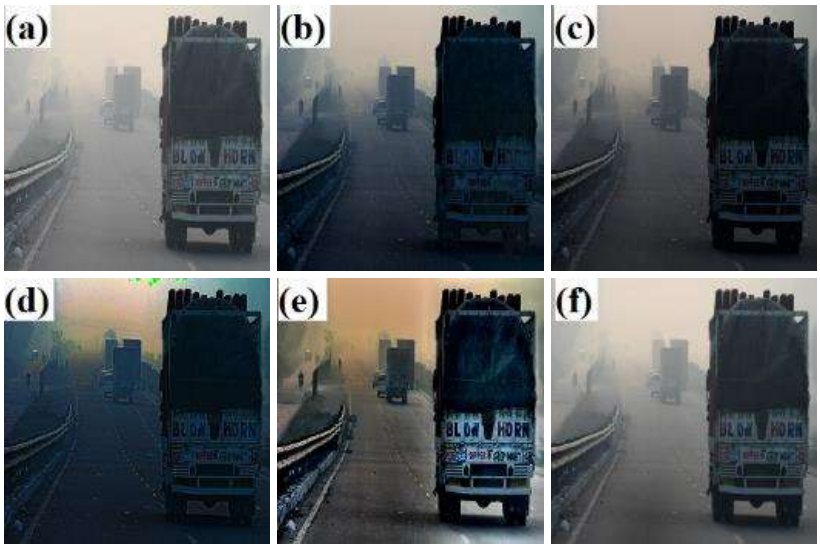
(iii) IMGN19

Fig. 4.1 (continue) Qualitative analysis of images obtained from IMSET-3 (a) hazy image (b) CAP (c) ICAP (d) CEP (e) SPD (f) grey wolf (proposed) (g) ground truth

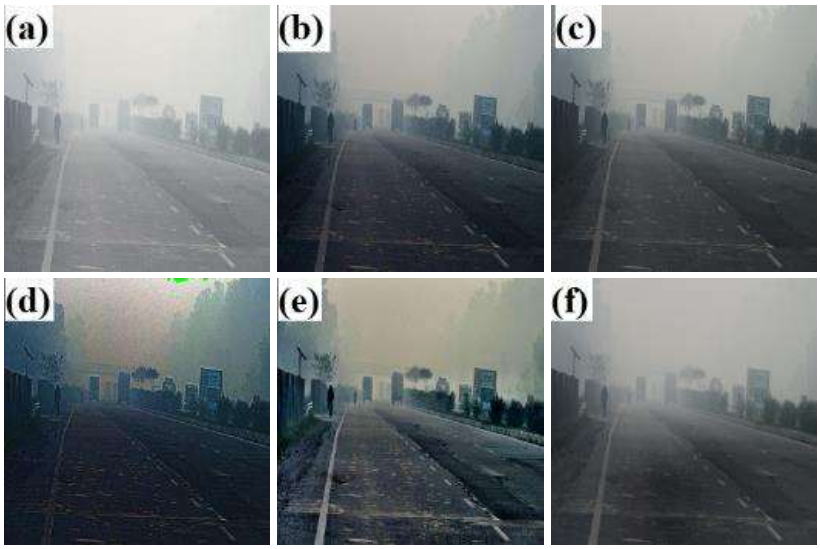
vision when GWO has been implemented. They do not suffer from color, edge, or texture distortion.

4.1.4 Quantitative Analysis

The quantitative analysis of the proposed and other algorithms is depicted in Figures 4.2 and 4.3, in terms of $SSIM$, and $PSNR$, respectively. In Figure 4.2, GWO produces highest value of $SSIM$ for IMGN17 and IMGN18 as 0.95 and 0.87, respectively. The values of $SSIM$ calculated for IMGN17 by using ICAP, CAP, SPD, and CEP are 0.94, 0.86, 0.83, and 0.80, respectively. Similarly, the values of $SSIM$ calculated for IMGN18 by using CAP, ICAP, SPD, and CEP are 0.85, 0.84, 0.81, and 0.79, respectively. The highest values of $SSIM$ for IMGN19 have been obtained as 0.85 and 0.82 when ICAP and GWO are implemented, respectively. It is followed by CAP, SPD, and CEP with $SSIM$ as 0.68, 0.66, and 0.58, respectively.



(iv) IMGN20



(v) IMGN21

Fig. 4.1 (continue) Qualitative analysis of images obtained from IMSET-3 (a) hazy image (b) CAP (c) ICAP (d) CEP (e) SPD (f) grey wolf (proposed)

In Figure 4.3, GWO produces highest value of $PSNR$ for IMGN17 and IMGN18 as 22.08 and 17.53, respectively. The values of $PSNR$ calculated for IMGN17 by using ICAP, CAP, SPD, and CEP are 21.47, 16.53, 15.76, and 13.41, respectively.

Similarly, the values of $PSNR$ calculated for IMGN18 by using ICAP, CAP, SPD, and CEP are 16.30, 15.71, 14.29, and 13.30, respectively. The highest values of $PSNR$ for IMGN19 have been obtained as 16.23 and 16.03 when ICAP and GWO are implemented, respectively. It is followed by SPD, CAP, and CEP with $PSNR$ as 13.82, 12.83, and 11.09, respectively.

(vi) IMGN22

(vii) IMGN23

Fig. 4.1 (continue) Qualitative analysis of images obtained from IMSET-3 (a) hazy image (b) CAP (c) ICAP (d) CEP (e) SPD (f) grey wolf (proposed)

Table 4.2 discusses the $NIQE$ of the computed images (symbols '*' and '#' indicates least and second-least values). The $NIQE$ has been obtained for IMGN17 through IMGN23 when GWO and other dehazing algorithms are used. It has been observed that the best value of $NIQE$ for IMGN18 is obtained by using GWO is 4.61. However, the least value of $NIQE$ for IMGN21 and IMGN23 are obtained by using SPD as 3.74 and 3.96, respectively. It is followed by CAP with the second-least values of $NIQE$ for IMGN21 and IMGN23 as 4.11 and 4.51, respectively. Furthermore, the best results for IMGN17 and IMGN20 are yielded by using CEP (4.88) and SPD (5.03),

respectively. It is followed by GWO with the values of $NIQE$ for IMGN17 and IMGN20 as 5.06 and 5.24, respectively. Also, the best values of $NIQE$ evaluated for IMGN19 and IMGN22 yielded by using CAP are 4.36 and 3.79, respectively. GWO produces second best values of $NIQE$, i.e., 4.68 and 4 for IMGN19 and IMGN22, respectively.

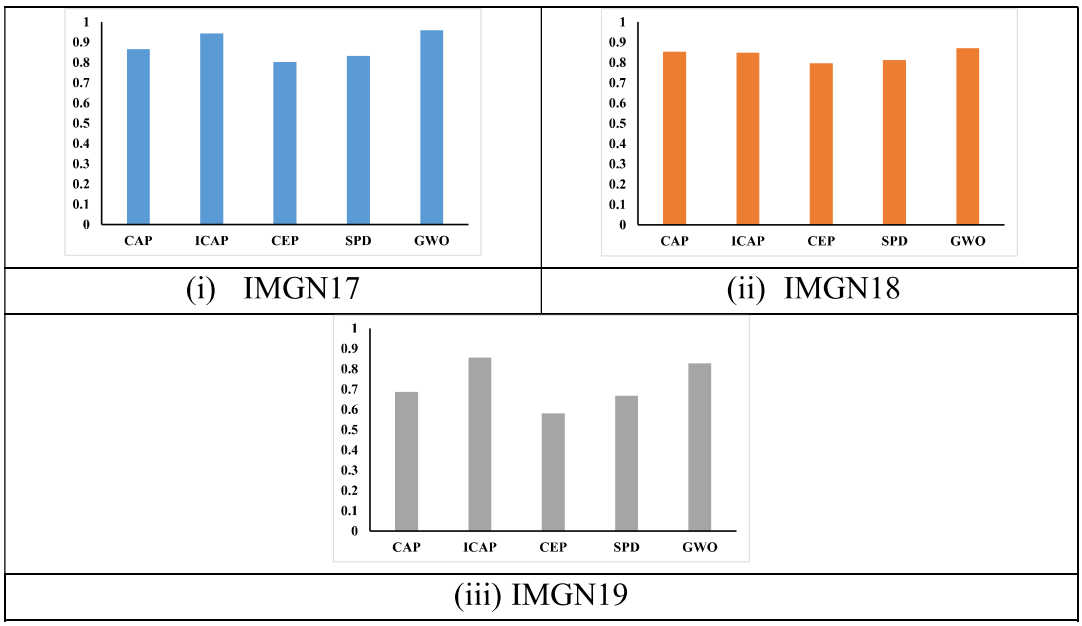


Fig. 4.2 Performance evaluation of grey wolf algorithm and existing dehazing techniques using $SSIM$

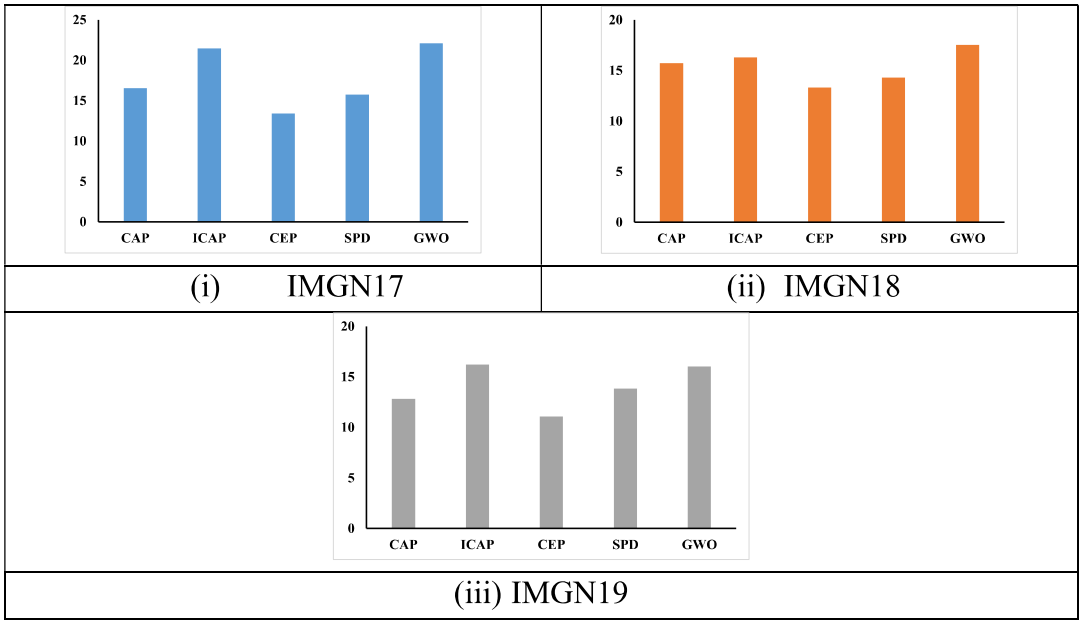


Fig. 4.3 Performance evaluation of grey wolf algorithm and existing dehazing techniques using $PSNR$

Table 4.2 Performance evaluation of grey wolf algorithm and existing dehazing techniques using $NIQE$

Images	CAP	ICAP	CEP	SPD	GWO
IMGN17	5.32	5.77	***4.88**	5.76	**#5.06**
IMGN18	4.93	4.86	5.2	**#4.73**	***4.61**
IMGN19	***4.36**	5.07	5.18	5.06	**#4.68**
IMGN20	5.7	5.34	5.85	***5.03**	**#5.24**
IMGN21	**#4.51**	5.08	5.04	***3.74**	4.9
IMGN22	***3.79**	4.08	5.09	4.03	**#4**
IMGN23	**#4.11**	4.17	5.87	***3.96**	4.13
Avg	4.67	4.91	5.30	***4.61**	**#4.66**

4.2 Hybrid Convolution Neural Network

Due to the dispersion of atmospheric particles, the quality of the image becomes inferior. The quality of the frames captured in poor weather conditions such as fog, smog, and haze declines as a result. These images are difficult to process using conventional visibility restoration algorithms. It is an intriguing task to estimate the concentration of haze at an unknown depth. In this chapter, a novel single image dehazing algorithm called hybrid CNN has been proposed that consists of a wavelet and inverted wavelet-based multi-scale convolution neural network (WIW-MSCNN).

4.2.1 Motivation

The word 'Retinex' is derived from the combination of words retina (a part of human eye) and cortex (a part of human mind) [188]. In Retinex theory, an image is represented as a product of incident light and reflected light. The aforementioned theory explains that the features of the incident light depend upon the source, whereas the features of the reflected light depend upon the object. The incident light is estimated by using prior-based knowledge of reflected light, and vice-versa. Therefore, assumptions are made about either illuminance, reflectance, or both. The foremost assumption is that the edges remain the same for hazy and transformed images. Also, the retinex filter is utilized to smoothen the feature map. Thus, illuminance is assumed to change gradually.

Xizhi [189] stated that different details can be extracted from an image, such as horizontal, vertical, diagonal, and approximate details. They can be enhanced or deteriorated as per the need for the computer vision problem. This has been achieved

with the help of multiple low-pass and high-pass filters. Thus, the image quality can be improved with a wavelet filter.

Ren et al. [143] proposed a MSCNN-based dehazing network, which consisted of a linearly connected coarse-scale and fine-scale layer. The coarse-scale network and the fine-scale network estimate the feature map from the inverted wavelet transformed image and the wavelet transformed image, respectively. One network generates a map for sky region, while other network generates a map for non-sky region. These maps have been combined to produce an estimated transmission map. The proposed model has the advantages of both prior-based techniques and deep learning-based techniques.

4.2.2 Modified Multi-scale Dehazing Architecture

In this section, retinex and wavelet-based filter has been implemented for image dehazing. It is followed by refining of the generated feature maps using a modified multi-scale CNN.

- **Features Enhancement**

Initially, the retinex algorithm [190], [191] was used to estimate the pixels' illuminance. The pixels' illuminance of the hazy image has been calculated by estimating through different paths and directions. This has been achieved by increasing the effect of reflected light on the image and reducing the effect of incident light on it.

$$R_c(x,y) = \log I_c(x,y) - \log[F(x,y) * I_c(x,y)] \tag{4.10}$$

where c denotes the channel number (1 for red, 2 for green, and 3 for blue in an RGB image) for point (x,y). $I_c(x,y)$ is the input image and $R_c(x,y)$ is the retinex processed output image. The symbol '$*$' represents the convolution operation. $F(x,y)$ is the centre represented by Gauss function.

$$F(x,y) = CONST \times \mathrm{e}^{\left(-(x^2+y^2)/d_{control}^2\right)} \tag{4.11}$$

where $CONST$ is proportionality constant usually taken as 1, and $d_{control}$ controls the function's range to decide the processed region.

The details of the retinex processed image have been enhanced using wavelet transformation [190], [192]. Initially, the discrete 2D wavelet transform operation was performed to extract the image details. The horizontal, vertical, and diagonal coefficients of the image have been modified. An inverse discrete 2D wavelet transform

has been performed by using the Haar wavelet. The reciprocal of the latter processed image has been performed to obtain an inverted wavelet transform to minimize the difference between the bright and dark pixels' intensities in the final output.

It helps to increase the visibility of drivers for both close and far objects to reduce the risk of accidents. Thus, wavelet and inverted wavelet transforms enhance the features of sky and non-sky regions, respectively. The filtered images have been passed through a multi-scale CNN to estimate and refine the transmission map.

- **Estimation of Transmission Map Using Multi-Scale Network**

The multi-scale neural architecture consists of two deep learning networks namely, coarse-scale and fine-scale networks. In these networks, filter banks of convolution layer have been utilized. The feature maps have been extracted using wavelet and inverted wavelet transforms and convolved with filter banks. The initial layer extracts only shallow features. The filter size has been reduced in each consecutive layer to extract deep features. The convolution operation is performed using Eq. (4.12).

$$conv(row, col) = \sum_{x,y}(f_x^l * k_{x,y}^{l+1}) + b_y^{l+1} \tag{4.12}$$

where $conv(row, col)$ is the convolved output. f_x^l represents the features of l^{th} layer. row and col are the numbers of rows and columns in the convolved image. k is the convolutional kernel that represents a mapping from the feature map of current layer to the feature map of next layer. b_y^{l+1} represents biasing used to shift the kernel.

The convoluted output has been passed through an activation function known as ReLU [48]. ReLU is an activation function that allows only positive values to pass through it.

$$f_y^{l+1} = \mathbb{R}(conv(row, col)) = \begin{cases} f_x^l; & (row, col) > 0 \\ 0; & otherwise \end{cases} \tag{4.13}$$

where $\mathbb{R}$ represents ReLU function that acts as an activation function.

The output of ReLU layer has been connected to maximum pooling layer. The size of feature maps at the input and output of the deep learning architecture must be same. An up-sampling layer has been connected to achieve the task. The output of up-sampling layer is given as:

$$f_y^{l+1}\Big(\big((2x-1):2x\big), \big((2y-1):2y\big)\Big) = f_y^l(x, y) \tag{4.14}$$

The combination of max pooling and upsampling has been used to introduce non-

linearity in the model to make it effective for real-world images. It has been followed by normalization layer to reduce the computation time and complexity [50]. The penultimate layer in the coarse-scale and the fine-scale network is another activation function known as the sigmoid function. It has been used to combine the features from preceding upsampling layer. The final output of either sub-network is given as:

$$t(x,y) = sigm\left(\sum_n \omega_y f_y^p + b_y^{l+1}\right) \tag{4.15}$$

where ω is the weight on y^{th} indexed feature map. $sigm$ denotes the sigmoidal function as given below:

$$sigm(x,y) = \frac{1}{1+\exp(-(x,y))} \tag{4.16}$$

Figure 4.4 depicts the neural network of the proposed approach to recover $J(x,y)$ from $I(x,y)$. The input to the coarse-scale network is the inverted wavelet transformed image. The different filter sizes in three convolution layers are 11×11, 9×9, and 7×7. The number of filters used in first, second, and third convolution layers are 5, 5, and 10, respectively. Similarly, the wavelet transformed image is the input to fine-scale network. In this sub-network, the filter sizes used in the convolution layers are 7×7, 5×5, and 3×3. The number of filters in convolution layers of a fine-scale network is 4, 5, and 10, respectively. The sigmoid functions have been used at the end of course and fine-scale network to generate separate feature maps. These maps have been combined to generate an enhanced transmission map.

- **Estimation of Airlight**

After the estimation of transmission map, the next task is to predict the airlight. As the scene depth increases, the transmitted light decreases.

$$I(x,y) = A, \quad if\ d(x,y) \to \infty, \qquad and\ t(x,y) \to 0 \tag{4.17}$$

Thus, the atmospheric light has been estimated using DCP [116]. In this technique, a group of pixels called a dark patch has been selected as a non-sky region by using the minimum filter in the hazy image. In this dark region, the top 0.1% of pixels with the highest luminance have been selected to estimate the airlight. The pixel with highest intensity in the dark patch has been estimated by 2D order statistic filter given as [193]:

$$WOS_o(n) = WOS_{i_{(w_o)}} = w_o'^{\,th}\ (largest[WOS_i(n) \times w]) \tag{4.18}$$

where $WOS_o(n)$ and $WOS_i(n)$ are the output and input of weighted order statistic (WOS) filter, respectively. $WOS_{i_{(w_o)}}$ is the order statistics for $w_o = 1, 2, \ldots, W_N$, and w is weight or strength of connection between two layers.

After refining the transmission map and estimating atmospheric airlight, the transmission map has been optimized by utilizing the input image, atmospheric airlight, and the refined transmission map. The refined transmission map's pixels have been scaled down to 0.1 times their original values relative to the square of the difference between the grayscale input image and the estimated airlight. Finally, the edges have been preserved and smoothed using a guided filter.

- **Visibility Restoration**

The clear image has been obtained by rearranging Eq. (1.1) and is given by Eq. (4.19). The exposure of light in the output image can be controlled by a parameter called "gamma." In the proposed scheme, $t(x, y)$ has been assumed to have a least value of 0.0001, which is referred as t_0.

$$J(x,y) = \frac{I(x,y) - A}{[max(t(x,y), t_0(x,y))]^{gamma}} + A \tag{4.19}$$

4.2.3 Result and Discussion

The performance of hybrid CNN has been evaluated using qualitative and quantitative analysis. The results have been compared with existing dehazing techniques namely, FFA-Net, CEP, GCANet, ICAP, NLBF, MSBDN, CCDI, and DW-GAN.

- **Experimental Set-up**

The proposed architecture has been simulated on an Intel (R) Core (TM) i7-6700 CPU @3.40GHz with 8GB RAM and 64-bit operating system on MATLAB R2021a. The architecture has been validated on hazy images from various datasets such as RESIDE, O-HAZE, DENSE-HAZE, NH-HAZE, FRIDA, FRIDA2, and HUDRS (synthetic and real-world). The training images have been resized to 598×598 pixels. The training and testing datasets consist of 2500 and 500 images, respectively.

The testing images have been divided into three test sets, such that each set represents images from different datasets, i.e., IMSET-4 consists of RESIDE and O-HAZE; IMSET-5 consists of DENSE-HAZE and NH-HAZE; IMSET-6 consists of FRIDA, FRIDA2, and HUDRS; and IMSET-7 consists of HUDRS (real-world) images.

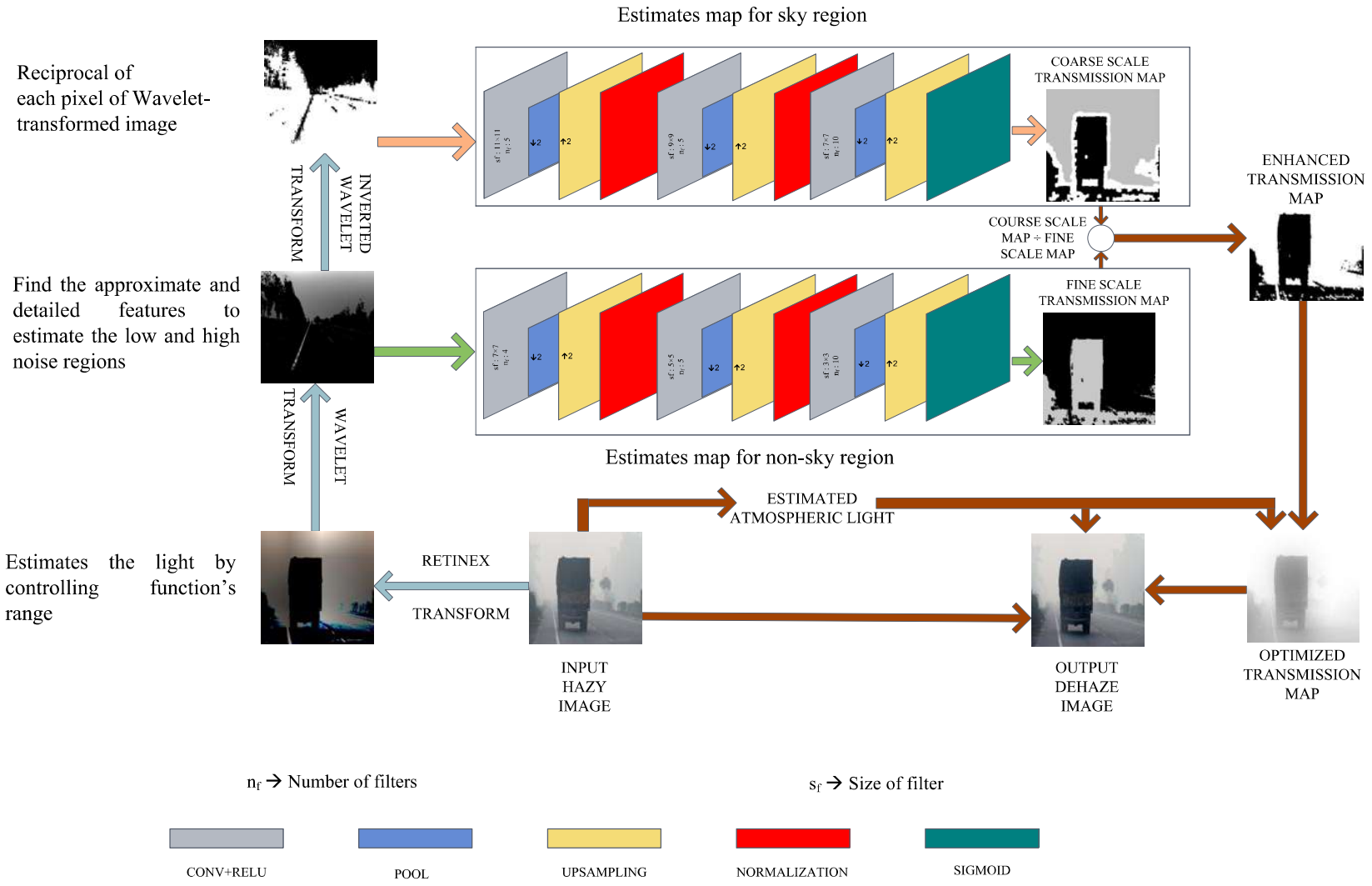


Fig. 4.4 Detailed architecture of hybrid CNN

The network has been trained using Adam optimizer with a decay factor of 0.9. It accelerates the convergence of gradient decent algorithm and smooths the training curve. The initial learning rate and batch size are 0.001 and 10, respectively. As the value of learning rate becomes closer to 1, the training curve becomes oscillatory in nature. On the other hand, if it becomes closer to 0, the computation time increases. So, an optimal learning rate has been selected with a decay factor of 0.9 for rapid convergence. The MatConvNet toolbox [194], [195] have been used to implement the proposed method. L2 loss function has been optimized to minimize the overfitting issue. It is calculated as:

$$L_2\ loss = \sum_{i=1}^{n} \left(S_{0_{true}} - S_{0_{predicted}} \right)^2 \tag{4.20}$$

where $S_{0_{true}}$ represents the ground truth image. $S_{0_{predicted}}$ denotes the predicted image.

- **Qualitative Analysis**

Figures 4.5 – 4.7 depict the visual comparison of images from IMSET-4 to IMSET-5, respectively. Various dehazing algorithms have been implemented on hazy images and results have been compared with the dehazed output generated by hybrid CNN. In Figure 4.5, the dehazed outcomes of IMSET-4 (i.e., IMGN24, IMGN25, IMGN26, IMGN27, IMGN28, and IMGN29) are represented. It is observed from this figure that RESIDE dataset has fewer haze images. Also, the images in the O-HAZE dataset comprise medium haze images. Figure 4.6 depicts non-homogeneous images in IMSET-5 (i.e., IMGN30, IMGN31, IMGN32, IMGN33, IMGN34, and IMGN35). The clearest results for IMSET-5 are generated using DW-GAN as they were designed for non-homogenous haze only. The dehazed images computed using the proposed approach are most suitable for FRIDA, FRIDA2, and HUDRS datasets. In Figure 4.7, images of IMSET-6 (i.e., IMGN36, IMGN37, IMGN38, IMGN39, IMGN40, and IMGN41) computed using hybrid CNN have no halo effects and color distortion due to the combination of a retinex filter, a wavelet filter, and an inverted wavelet filter. The hybrid CNN has been found to be effective in restoring visibility up to medium non-homogenous hazy regions.

$FADE$ has been used to calculate the effectiveness of several dehazing techniques on IMGN24, IMGN25, IMGN27, IMGN28, IMGN30, IMGN31, IMGN32, IMGN33, IMGN34, IMGN35, IMGN36, IMGN37, IMGN39, IMGN40, and IMGN41.

Algorithm	Images					
	IMGN24	IMGN25	IMGN26	IMGN27	IMGN28	IMGN29
Hazy Image						
Ground Truth						
FFA-Net						
CEP						
GCANet						
ICAP						
NLBF						
MSBDN						
CCDI						
DW-GAN						
Hybrid CNN (Proposed Approach)						

Fig. 4.5 Qualitative analysis of IMSET-4 from RESIDE and O-HAZE

Algorithm	Images					
	IMGN30	IMGN31	IMGN32	IMGN33	IMGN34	IMGN35
Hazy Image						
Ground Truth						
FFA-Net						
CEP						
GCANet						
ICAP						
NLBF						
MSBDN						
CCDI						
DW-GAN						
Hybrid CNN (Proposed Approach)						

Fig. 4.6 Qualitative analysis of IMSET-5 from DENSE-HAZE and NH-HAZE

Algorithm	Images					
	IMGN36	IMGN37	IMGN38	IMGN39	IMGN40	IMGN41
Hazy Image						
Ground Truth						
FFA-Net						
CEP						
GCANet						
ICAP						
NLBF						
MSBDN						
CCDI						
DW-GAN						
Hybrid CNN (Proposed Approach)						

Fig. 4.7 Qualitative analysis of IMSET-6 from FRIDA, FRIDA-2, and HUDRS

Table 4.3 Comparative analysis of hybrid CNN with existing dehazing techniques based on $FADE$

Images	FFA-Net	GCANet	ICAP	NLBF	MSBDN	DW-GAN	Hybrid CNN
IMGN24	0.4	0.41	1.28	0.57	**#0.39**	1.03	***0.25**
IMGN25	0.93	**#0.76**	2.11	***0.75**	0.97	3.35	0.9
IMGN27	2.3	***0.43**	1.43	**#0.51**	2.08	1.08	0.53
IMGN28	1.63	***0.5**	1.15	**#0.64**	1.26	1.25	1.79
IMGN30	0.79	***0.4**	2.1	1.74	1.48	0.63	**#0.42**
IMGN31	1.91	***1.15**	4.39	2.33	2.37	**#1.48**	1.66
IMGN32	2.28	***1.01**	8.89	1.34	5.39	1.86	**#1.08**
IMGN33	**#0.75**	***0.34**	1.61	1.23	1.03	0.88	0.8
IMGN34	**#0.45**	***0.3**	0.66	0.43	0.64	0.76	0.48
IMGN35	0.82	***0.42**	1.012	**#0.54**	1.06	0.91	0.8
IMGN36	1.04	**#0.77**	1.81	***0.75**	2.01	4.49	1.56
IMGN37	1.04	0.73	1.38	***0.49**	2.27	4.01	**#0.53**
IMGN39	1.35	1.21	1.52	**#0.56**	1.43	2.74	***0.4**
IMGN40	0.93	**#0.54**	0.78	***0.27**	1	1.54	0.62
IMGN41	1.34	0.75	1.08	***0.34**	1.14	2.22	**#0.45**
Avg	1.197333	***0.648**	2.080133	0.832667	1.634667	1.882	**#0.818**

Various factors namely, contrast energy, sharpness, image entropy, and color saturation of the image, are considered to evaluate the fog images' density. Human and camera-rated judgements such as depth information and camera position are not considered while calculating haze density. Thus, thin fog images can have higher $FADE$ values than dense fog, because the latter may preserve texture. A smaller value of $FADE$ is considered suitable.

Table 4.3 depicts the fog density in the images computed using various dehazing techniques (best and second best results are indicated by '*' and '#', respectively). It has been observed that the best values obtained for IMGN24 and IMGN39 are 0.25 and 0.4, respectively, using hybrid CNN. Also, it produces the second best $FADE$ values for IMGN30, IMGN32, IMGN37, and IMGN41 as 0.42, 1.08, 0.53, and 0.45, respectively.

Furthermore, GCANet and hybrid CNN has the best (i.e.,0.648) and second best (i.e., 0.818) $FADE$ values, respectively. The $FADE$ values of FFA-Net, ICAP, NLBF, MSBDN, and DW-GAN are 1.19, 2.08, 0.83, 1.63, and 1.82, respectively. Although the DW-GAN has better visual results, it has higher FADE value as compared to hybrid

CNN. It is because the results generated by GCANet suffer from color distortion. Also, the dehazed images obtained by implementing DW-GAN have low contrast energy and lack sharpness.

The effectiveness of the proposed approach has also been tested on real-world images (i.e., IMGN42, IMGN43, IMGN44, IMGN45, and IMGN46) obtained from HUDRS (IMSET-7). The intermediate steps of the proposed architecture, i.e., retinex map, wavelet map, inverted wavelet map, coarse scale transmission map, fine scale transmission map, enhanced transmission map, optimized transmission map, and restored image, are represented in Figure 4.8. Furthermore, the transmission maps of IMGN42 through IMGN46 produced using FFA-Net, CEP, ICAP, NLBF, and hybrid CNN are depicted in Figure 4.9. It has been observed that the pixels' intensities are low in the maps obtained using FFA-Net. The maps obtained using CEP, ICAP, and NLBF have smooth and blurred edges. The objects in the feature maps produced using hybrid CNN have been distinctively visible.

The image quality is affected by two parameters, i.e., haze density in the foreground and shade of light in the background. The haze density in the computed image varies with gamma, as shown in Eq. (4.19). The gamma value is high for dense foggy images and low for thin foggy images, irrespective of the duration of the day. For instance, gamma for NH-HAZE is more than 5. Similarly, the gamma range for FRIDA is 1 to 5, and that for RESIDE is 0 to 1. Figure 4.10 depicts the dehazed results of IMSET-8 (i.e., IMGN47, IMGN48, IMGN49, and IMGN50) to study the effect of "gamma." The value of gamma is less than 1 for all the images due to thin fog. These images have different shades of light in the background. For $\gamma = 0$, the computed image is same as the original image. When the value of γ increases, the sharpness increases and the haze decreases in the output image. The sky region is yellow in IMGN47 as it was captured in the evening.

The most suitable value of gamma for visibility restoration in IMGN47 is 0.07. The visibility in IMGN48 captured at noon has been restored at a gamma value of 0.03. Images such as IMGN49 and IMGN50 are captured in the early morning and night, respectively. Although the features are visible when gamma is set to 0.03, they start getting distorted at $\gamma = 0.07$. However, the latter value has been preferred in images with heavy traffic.

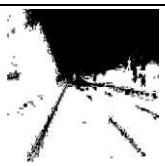
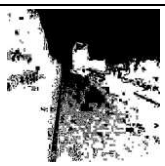
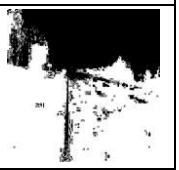
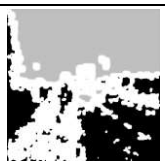
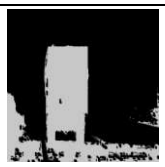
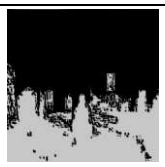
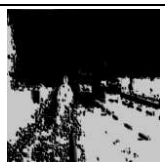
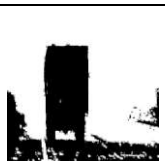
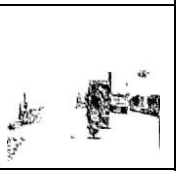
Proposed Approach	Images				
	IMGN42	IMGN43	IMGN44	IMGN45	IMGN46
Hazy Image					
Retinex Filter					
Wavelet Filter					
Inverted Wavelet Filter					
Coarse Scale Transmission Map					
Fine Scale Transmission Map					
Enhanced Transmission Map					
Optimized Transmission Map					
Dehazed Image					

Fig. 4.8 Qualitative analysis of IMSET-7 from HUDRS (real-world)

- **Quantitative Analysis**

Images from different datasets have been analyzed using hybrid CNN for various performance metrics such as IE, $SSIM$, $PIQE$, $PSNR$, and $NIQE$. The performances of

existing dehazing techniques namely, FFA-Net, CEP, GCANet, ICAP, NLBF, MSBDN, CCDI, and DW-GAN have been compared with the proposed dehazing technique.

Table 4.4 presents the comparative analysis of $SSIM$ of IMGN24, IMGN25, IMGN26, IMGN28, IMGN29, IMGN30, IMGN31. IMGN32, IMGN34, IMGN35, IMGN37, IMGN38, IMGN40, and IMGN41 calculated using FFA-Net, CEP, ICAP, NLBF, CCDI, and hybrid CNN (best and second best results are indicated by '*' and '#', respectively). It has been observed that the hybrid CNN produced the best values for IMGN25, IMGN28, IMGN32, and IMGN35 as 0.94, 0.52, 0.24, and 0.14, respectively. Also, second-best values have been obtained for IMGN30 and IMGN40 as0.07 and 0.8, respectively. The best values for IMGN24, IMGN26, IMGN38, IMGN40, and IMGN41 have been obtained as 0.93, 0.92, 0.75, 0.82, and 0.81, respectively, when FFA-Net is implemented. Similarly, ICAP produces best results for IMGN29, IMGN31, IMGN37, and IMGN40 as 0.1, 0.52, 0.54, and 0.82, respectively. Also, NLBF produces best results for IMGN32 (i.e., 0.24) and IMGN34 (i.e., 0.25). CCDI produces best results

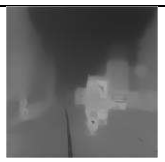
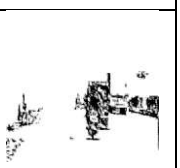
Algorithm	FFA-Net	CEP	ICAP	NLBF	Hybrid CNN
IMGN42					
IMGN43					
IMGN44					
IMGN45					
IMGN46					

Fig. 4.9 Comparative analysis of transmission maps generated by hybrid CNN and existing dehazing techniques

for IMGN29 (i.e., 0.1) and IMGN30 (i.e., 0.09). Overall, the best and second-best $SSIM$ values have been obtained using FFA-Net and hybrid CNN as 0.4962 and 0.4635, respectively. It is followed by ICAP, NLBF, CCDI, and ICAP with $SSIM$ values of 0.4631, 0.4414, 0.4235, and 0.3907, respectively.

Table 4.5 presents the comparative analysis of $PSNR$ for IMGN24, IMGN25, IMGN26, IMGN27, IMGN28, IMGN29, IMGN30, IMGN32, IMGN34, IMGN35, IMGN36, IMGN37, IMGN38, IMGN39, and IMGN40 calculated using FFA-Net, CEP, ICAP, NLBF, CCDI, DW-GAN, and hybrid CNN (best and second best results are indicated by '*' and '#', respectively). It has been observed that the hybrid CNN produced the best and second best values for IMGN25 (i.e., 23.8) and IMGN28 (i.e., 20.2), respectively. The best values for IMGN24, IMGN26, and IMGN38 have been obtained as 20.41, 19.18, and 12.74, respectively, when FFA-Net is implemented.

Similarly, ICAP produces best results for IMGN39 (i.e., 0.24) and IMGN40 (i.e., 0.25). Also, NLBF produces best results for IMGN36 (i.e., 14.02) and IMGN37 (i.e., 13.8), and DW-GAN produces best results for IMGN27 (i.e., 15.4), IMGN28 (i.e., 21.56), IMGN29 (i.e., 20.18), IMGN30 (i.e., 17.93), IMGN32 (i.e., 15.3), IMGN34 (i.e., 18.74), and IMGN35 (i.e., 16.83), respectively. Overall, the best and second-best $PSNR$ values have been obtained using NLBF and FFA-Net as 13.974 and 13.947, respectively.

It is followed by CCDI, DW-GAN, hybrid CNN, CEP, and ICAP with $PSNR$ values of 13.81, 13.62, 13.29, 13.24, and 13.16, respectively. Figure 4.11 presents the comparative analysis of average $NIQE$ for IMSET-4 through IMSET-6 calculated using several dehazing techniques namely, CEP, GCANet, NLBF, MSBDN, CCDI, DW-GAN, and hybrid CNN. It has been observed that GCANet and hybrid CNN produce the best and second best results as 4.31 and 4.47, respectively. They are followed by NLBF, CEP, CCDI, MSBDN, and DW-GAN as 4.51, 4.52, 4.55, 4.59, and 5.25, respectively.

Figure 4.12 presents the average $PIQE$ values of IMGN24, IMGN25, IMGN27, IMGN28, IMGN29, IMGN30, IMGN31, IMGN32, IMGN35, IMGN36, IMGN37, IMGN40, and IMGN41 obtained from several dehazing techniques namely, CEP, GCANet, ICAP, NLBF, MSBDN, CCDI, DW-GAN, and hybrid CNN.

Image Name	Output Images			
	γ = 0	γ = 0.03	γ = 0.07	γ = 0.1
IMGN47				
IMGN48				
IMGN49				
IMGN50				

Fig. 4.10 Analysis of hazy images of IMSET-8 from HUDRS (real-world) for different values of gamma

Table 4.4 Comparative analysis of hybrid CNN with existing dehazing techniques based on *SSIM*

Images	FFA-Net	GCANet	ICAP	NLBF	CCDI	Hybrid CNN
IMGN24	***0.93**	**#0.86**	0.6	0.78	0.77	0.76
IMGN25	**#0.93**	0.87	0.75	0.82	0.91	***0.94**
IMGN26	***0.92**	0.73	0.74	**#0.86**	0.77	0.7
IMGN28	0.48	0.25	**#0.49**	0.39	0.41	***0.52**
IMGN29	0.007	**#0.08**	***0.1**	**#0.08**	***0.1**	0.02
IMGN30	0.05	0.06	0.004	0.01	***0.09**	**#0.07**
IMGN31	0.35	0.29	***0.52**	**#0.47**	0.25	0.41
IMGN32	0.15	0.12	0.11	***0.24**	**#0.23**	***0.24**
IMGN33	0.14	**#0.23**	0.22	***0.25**	0.16	0.18
IMGN34	0.12	0.03	**#0.13**	0.11	0.04	***0.14**
IMGN37	0.49	0.48	***0.54**	**#0.53**	0.5	0.35
IMGN38	***0.75**	0.42	**#0.66**	0.51	0.48	0.58
IMGN40	***0.82**	0.54	***0.82**	0.6	0.64	**#0.8**
IMGN41	***0.81**	0.51	**#0.8**	0.53	0.58	0.78
Avg	***0.496214**	0.390714	0.463143	0.441429	0.423571	**#0.463571**

Table 4.5 Comparative analysis of hybrid CNN with existing dehazing techniques based on *PSNR*

Images	FFA-Net	CEP	ICAP	NLBF	CCDI	DW-GAN	Hybrid CNN
IMGN24	***20.41**	**#20.21**	12.41	16.71	19.19	9.12	14.6
IMGN25	**#22.41**	18.12	15.74	17.64	20.41	8.14	***23.8**
IMGN26	***19.18**	16.74	16.35	**#18.87**	18.53	9.59	12.2
IMGN27	10.64	13.55	14.27	**#14.56**	14.51	***15.4**	12.4
IMGN28	19.75	11.17	17.77	16.11	13.26	***21.56**	**#20.2**
IMGN29	11.79	13.97	14.87	**#16.67**	14.68	***20.18**	12.3
IMGN30	10.85	**#12.54**	8.78	8.66	9.98	***17.93**	12.3
IMGN32	12.82	12.75	8.31	12.69	**#13.37**	***15.3**	12.5
IMGN33	12.42	12.87	**#13.44**	13.11	13.18	***18.74**	11.76
IMGN35	10.48	10.86	**#11.66**	11.64	11.27	***16.83**	11
IMGN36	11.08	**#13.21**	11.18	***14.02**	13.11	9.37	9.43
IMGN37	10.29	**#12.57**	10.8	***13.8**	12.52	9.56	9.47
IMGN38	***12.74**	7.59	**#12.06**	8.86	7.51	11.52	11.96
IMGN39	**#10.64**	9	***11.8**	10.33	9.58	9.67	9.94
IMGN40	13.71	13.54	***18.05**	15.94	**#16.09**	11.43	15.6
Avg	**#13.94733**	13.246	13.166	***13.974**	13.81267	13.62267	13.29733

.

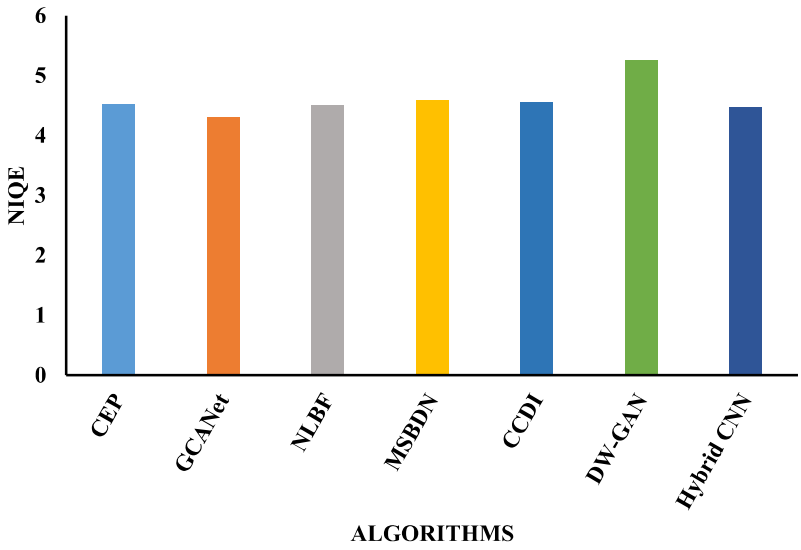


Fig. 4.11 Comparative analysis of hybrid CNN with existing dehazing techniques based on $NIQE$

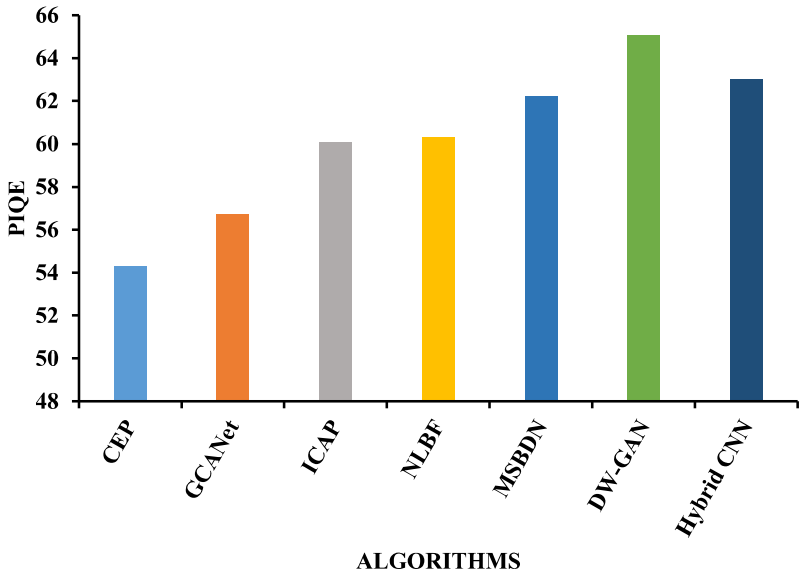


Fig. 4.12 Comparative analysis of hybrid CNN with existing dehazing techniques based on $PIQE$

It has been observed that CEP and GCANet produce the best and second best results as 54.31 and 56.71, respectively. They are followed by ICAP, NLBF, MSBDN, DW-GAN, and hybrid CNN as 60.09, 60.29, 62.2, 63.02, and 65.09, respectively. A model's computational complexity is evaluated based on two major factors namely, performance metrics and computation time. Table 4.6 displays the comparative analysis of computation time obtained from hybrid CNN.

Table 4.6 Comparative analysis of hybrid CNN with existing dehazing techniques based on computation time (in seconds)

Images	FFA-Net	CEP	GCANet	ICAP	NLBF	CCDI	DW-GAN	Hybrid CNN
IMGN24	0.73	0.52	0.82	0.85	0.64	0.61	1.24	0.58
IMGN25	0.75	0.54	0.86	0.82	0.67	0.64	1.26	0.56
IMGN26	0.71	0.50	0.87	0.84	0.66	0.68	1.22	0.58
IMGN27	0.76	0.57	0.82	0.87	0.71	0.64	1.27	0.57
IMGN28	0.71	0.55	0.88	0.79	0.68	0.62	1.23	0.59
IMGN29	0.77	0.57	0.84	0.83	0.63	0.67	1.28	0.60
IMGN30	0.70	0.53	0.89	0.84	0.59	0.61	1.22	0.57
IMGN31	0.74	0.56	0.84	0.83	0.64	0.66	1.28	0.57
IMGN32	0.77	0.51	0.89	0.92	0.68	0.68	1.21	0.56
IMGN33	0.82	0.54	0.85	0.87	0.66	0.70	1.28	0.58
IMGN34	0.74	0.57	0.87	0.85	0.61	0.69	1.23	0.55
IMGN35	0.76	0.58	0.82	0.85	0.69	0.64	1.25	0.59
IMGN36	0.77	0.53	0.85	0.89	0.72	0.61	1.18	0.55
IMGN37	0.72	0.56	0.87	0.82	0.70	0.62	1.24	0.56
IMGN38	0.77	0.54	0.82	0.86	0.64	0.61	1.22	0.59
IMGN39	0.73	0.58	0.88	0.91	0.63	0.65	1.19	0.60
IMGN40	0.78	0.52	0.85	0.83	0.62	0.66	1.20	0.58
IMGN41	0.74	0.58	0.81	0.88	0.69	0.63	1.25	0.59
Avg	0.7483	***0.5472**	0.8516	0.8527	0.6588	0.6455	1.2361	**#0.5761**

It has been observed that the average time to compute the dehazed results obtained using CEP and hybrid CNN are 0.5472 and 0.5761 seconds, respectively. They are followed by CCDI, NLBF, FFA-Net, GCANet, ICAP, and DW-GAN, with computation times of 0.6455, 0.6588, 0.7483, 0.8516, 0.8527, and 1.2361 seconds, respectively.

4.2.4 Ablation Studies

A number of ablation experiments support the contributions made by each step in the proposed model. The different components from each training had been removed for these tests. It primarily entails the following actions:

(1) In Stage-I, the original image has been reconstructed without the retinex filter.

(2) In Stage-II, the original image has been restored without using a wavelet filter.

(3) In Stage-III, the original image has been recovered without using an inverted wavelet filter.

(4) The original image has been recovered without using the inverted wavelet transform and course-scale feature map in Stage-IV.

(5) The original image has been recovered without using wavelet transform and fine-scale feature map in Stage-V.

(6) All filters and feature maps have been used to restore the original image in Stage-VI.

The qualitative analysis of ablation outcomes for IMGN42 through IMGN46 is illustrated in Figure 4.13. All modifications have been proven effective in restoring visibility in areas of high pollution. However, the images restored using Stages-II and III have degraded in visual quality. The images restored using Stage-IV have color and texture distortion. Additionally, IE, $PIQE$, and $NIQE$ have been calculated for IMGN42 through IMGN46 by eliminating various filters and feature maps (see Tables 4.6 – 4.8).

As depicted in Table 4.7, the best results of IE for RW-2 and RW-4 are obtained using Stage VI as 7.4489 and 7.1475, respectively. Also, the best results for RW-1 and RW-3 are obtained using Stage IV at 7.5432 and 7.5496, respectively. Furthermore, the highest value of IE for RW-5 is produced by Stage-III. Similarly, the $PIQE$ of IMGN42 through IMGN46 is depicted in Table 4.8. The least $PIQE$ of IMGN42, IMGN43, IMGN45, and IMGN46 are 42.2697, 55.3237, 54.3325, and 40.4397, respectively.

Also, Stage-I has least $PIQE$ for IMGN44 (i.e., 55.7367). Table 4.9 depicts the $NIQE$ for IMGN42 through IMGN46. It has been observed that IMGN42 and IMGN46 have least $NIQE$ values of 3.1265 and 2.8322, respectively when Stage-VI has been implemented. The least $NIQE$ for IMGN43 (i.e., 3.5361), IMGN44 (i.e., 3.7471), and IMGN45 (i.e., 4.2038) have been obtained when Stage-IV, Stage-I, and Stage-II, respectively, are implemented.

4.2.5 Summary

In grey wolf optimization algorithm, a random population of wolves and their respective positions have been initialized. In this technique, canny edge filter has been used to evaluate the fitness function. The positions of the wolves have been updated using the fitness function. It has been found superior than existing dehazing techniques namely, color attenuation prior, improved color attenuation prior, color ellipsoid prior, and structure patch decomposition technique in terms of similarity index metric, peak

Algorithm	IMGN42	IMGN43	IMGN44	IMGN45	IMGN46
Without Retinex Transform					
Without Wavelet Transform					
Without Inverted Wavelet Transform					
Without Inverted Wavelet Transform and Course Map Estimation					
Without Wavelet Transform and Fine Map Estimation					
With All Modules and Network					

Fig. 4.13 Qualitative analysis of ablation results for hybrid CNN

Table 4.7 *IE* of ablation results evaluated for hybrid CNN

Algorithm	IMGN42	IMGN43	IMGN44	IMGN45	IMGN46
Without Retinex Transform	7.2311	7.3978	7.5456	7.032	6.9857
Without Wavelet Transform	7.1874	6.6838	6.4207	7.1414	6.9857
Without Inverted Wavelet Transform	7.1874	5.9002	5.7179	5.1123	**7.3195**
Without Inverted Wavelet Transform and Course Map Estimation	**7.5432**	7.4171	**7.5496**	6.9865	7.1739
Without Wavelet Transform and Fine Map Estimation	7.2426	7.4403	7.4936	7.0324	6.9857
With All Modules and Network	7.1216	**7.4489**	7.4391	**7.1475**	6.8086

Table 4.8 $PIQE$ of ablation results evaluated for hybrid CNN

Algorithm	IMGN42	IMGN43	IMGN44	IMGN45	IMGN46
Without Retinex Transform	43.7162	59.8639	**55.7367**	55.6248	44.0336
Without Wavelet Transform	43.8546	56.9956	61.062	56.1164	44.0336
Without Inverted Wavelet Transform	43.8546	61.8351	60.474	66.8527	45.8136
Without Inverted Wavelet Transform and Course Map Estimation	43.7372	57.6029	57.3497	55.4913	45.5324
Without Wavelet Transform and Fine Map Estimation	43.765	60.0567	56.1142	55.3294	44.0336
With All Modules and Network	**42.2697**	**55.3237**	57.3075	**54.3325**	**40.4397**

Table 4.9 $NIQE$ of ablation results evaluated for hybrid CNN

Algorithm	IMGN42	IMGN43	IMGN44	IMGN45	IMGN46
Without Retinex Transform	3.3146	3.6106	**3.7471**	4.3065	2.8611
Without Wavelet Transform	3.3415	4.539	4.4753	**4.2038**	2.8611
Without Inverted Wavelet Transform	3.3415	4.873	5.2629	4.829	3.044
Without Inverted Wavelet Transform and Course Map Estimation	3.2226	**3.5361**	3.7994	4.3645	2.9441
Without Wavelet Transform and Fine Map Estimation	3.2868	3.6178	3.8234	4.311	2.8611
With All Modules and Network	**3.1265**	3.5876	3.8013	4.2896	**2.8322**

signal-to-noise, and naturalness quality. It has been observed that GWO produced five best and six second-best results out of 13 results.

In another approach, a retinex filter has been used to reduce the effect of incident light and strengthen the effect of reflected light on the image. A wavelet filter has been used to estimate the horizontal, vertical, diagonal, and approximate components of the hazy image for feature estimation. A multi-scale network has been proposed such that a coarse scale-CNN and fine scale-CNN have been used to produce two individual transmission maps, which have been further combined to obtain a refined map. The proposed architecture has been tested on various benchmark datasets such as RESIDE, O-HAZE, DENSE-HAZE, NH-HAZE, FRIDA, FRIDA, FRIDA2, and HUDRS.

CHAPTER-5
AETHRA-NET: SINGLE IMAGE AND VIDEO DEHAZING USING AUTOENCODER

A fast and efficient video dehazing system with low computational complexity has a huge demand among drivers during hazy winter nights. There are only a few video dehazing models that exist in the literature. Video dehazing requires the sequential extraction and processing of frames. The processed frames must be restored in the same sequence as the original video. However, the existing video dehazing algorithms suffer from color distortion due to the continuous processing of frames. They are not suitable for videos with dense haze. In this chapter, an image and video dehazing system called Aethra-Net has been developed. Also, a real-time dehazing framework is proposed that uses a GoPro camera to capture the hazy scene and a display screen to observe the results. Furthermore, Aethra-Net has been used in the proposed framework to perform real-time dehazing.

5.1 Motivation

Aethra is the Goddess of vision or bright sky in Greek mythology. The existing dehazing algorithms for a single image and video processing have a number of computational challenges, including color deviation, edge distortion, and halo effects, as well as memory storage problems. Aethra-Net comprises five units namely, an encoder, a gush enhancement unit, three blocks of ResNet-101, a decoder, and a vessel enhancement unit. An autoencoder is a convolution neural network (CNN) with encoder and decoder units. An autoencoder detects anomalies in images and simplifies feature detection and extraction.

The encoder downsamples the original hazy image to minimize the spatial size of the feature map for faster computation. The decoder has been used to restore the original spatial size after image processing operations have been completed. The feature maps of encoder and decoder with the same spatial dimensions have been connected via residual network blocks. Light enters the human eye through a variety of pathways, each with a different intensity.

The light processing is influenced by the angle of incidence, intensity, glare, and other lighting characteristics. The human eye, on the other hand, processes the different

routes concurrently to produce a complete image. This motivated the addition of a gush enhancement unit between the encoder and decoder units. The gush enhancement unit generates three feature maps with varying spatial information from the encoded map.

A group of unweighted layers obtained from a residual network of 101 deep layers, named ResNet-101 [196], has been used to connect the encoder and the decoder unit to overcome the vanishing gradient problem [197]. This architecture reduced the accuracy of the restored and predicted images. In some cases, the decoder unit may lose information retrieved from the important features. Thus, a second feature map has been obtained by implementing the vessel enhancement filter.

The vessel enhancement filter has been applied to compensate for the loss of extracted characteristics [198]. Despite being a biomedical imaging filter, it improves the visibility of objects of large value in roadside images that are partially obscured by objects of lesser importance. This filter has been used to locate and identify objects at large depths. It identifies and highlights objects with geometric shapes that are similar to tubular structures. As a result, even if the objects are present at a large depth in the image, or may not be visible at first glance owing to clustering, the edges of the objects are preserved.

5.2 Proposed Dehazing Model

This section discusses the proposed dehazing algorithm. Airlight have been estimated using mean filter. Furthermore, the transmission map has been estimated using a deep neural network and a biomedical imaging filter. They have been discussed in detail in subsections as follows:

5.2.1 Estimation of Atmospheric Light

The mean filter has been used to evaluate airlight. The proposed approach computes average pixel intensity of each color channel c to estimate atmospheric light. The average value of the atmospheric light has been calculated by implementing the average filter to each channel separately. The significance of calculating atmospheric light is to estimate the transmission map and restore image visibility. As a result, airlight is computed as follows:

$$A = \underbrace{avg}_{c\in\{r,g,b\}} \left(\frac{I^c(x,y)}{255}\right) \tag{5.1}$$

where $I^c(x,y)$ is the hazy image in color channel c.

5.2.2 Computation of Transmission Map

Due to the presence of haze particles, an individual's visibility is reduced to 1km. The optimized Aethra map (i.e., $t(x, y)$) has been obtained by mathematically combining a course Aethra map and a tubular enhanced feature map.

- **Estimation of Transmission Map**

The transmission map has been estimated by using the neural network's encoder, gush enhancer, and decoder. The design of the gush enhancement-based autoencoder is depicted in Figure 5.1. The encoder unit has been used to minimize the size of the feature map and perform feature enhancement operations. It reduces the dimension of trainable parameters and the computation time. The encoder unit comprises three pairs of convolutions (Conv) and rectified linear unit (ReLU) layers. Also, the maximum pooling layer has been employed to reduce the size of the feature map during the transition from one pair to another. Table 5.1 shows a layer-wise analysis of the encoder unit of the proposed architecture.

The gush enhancer has been used to enhance the features of the encoded feature map. It computes three feature maps with varying spatial dimensions using the same encoded map. The global-average pooling layer, average pooling layer, and max pooling layer have been used to extract high-level, medium-level, and low-level characteristics, respectively. To acquire high-level features, the global-average pooling layer takes the average value of all the pixels in the feature map. The average pooling layer has been utilized to generate the average pixel value for medium-level features using the kernel of dimension 4×4. A 2×2 kernel has been used to obtain low-level features using the max pooling layer.

These are all followed by individual pairs of Conv and ReLU with varying padding and stride. Light entering the human eye from different paths has been represented by feature maps with diverse spatial dimensions. The spatial information refers to the size of the feature maps after dimensional redundancy. An upsampling layer has been used to make the spatial dimensions the same on all three feature maps. These feature maps have been combined to form a single feature map with multiple filters. Table 5.2 depicts a layer-by-layer analysis of the gush enhancer unit of Aethra-Net architecture.

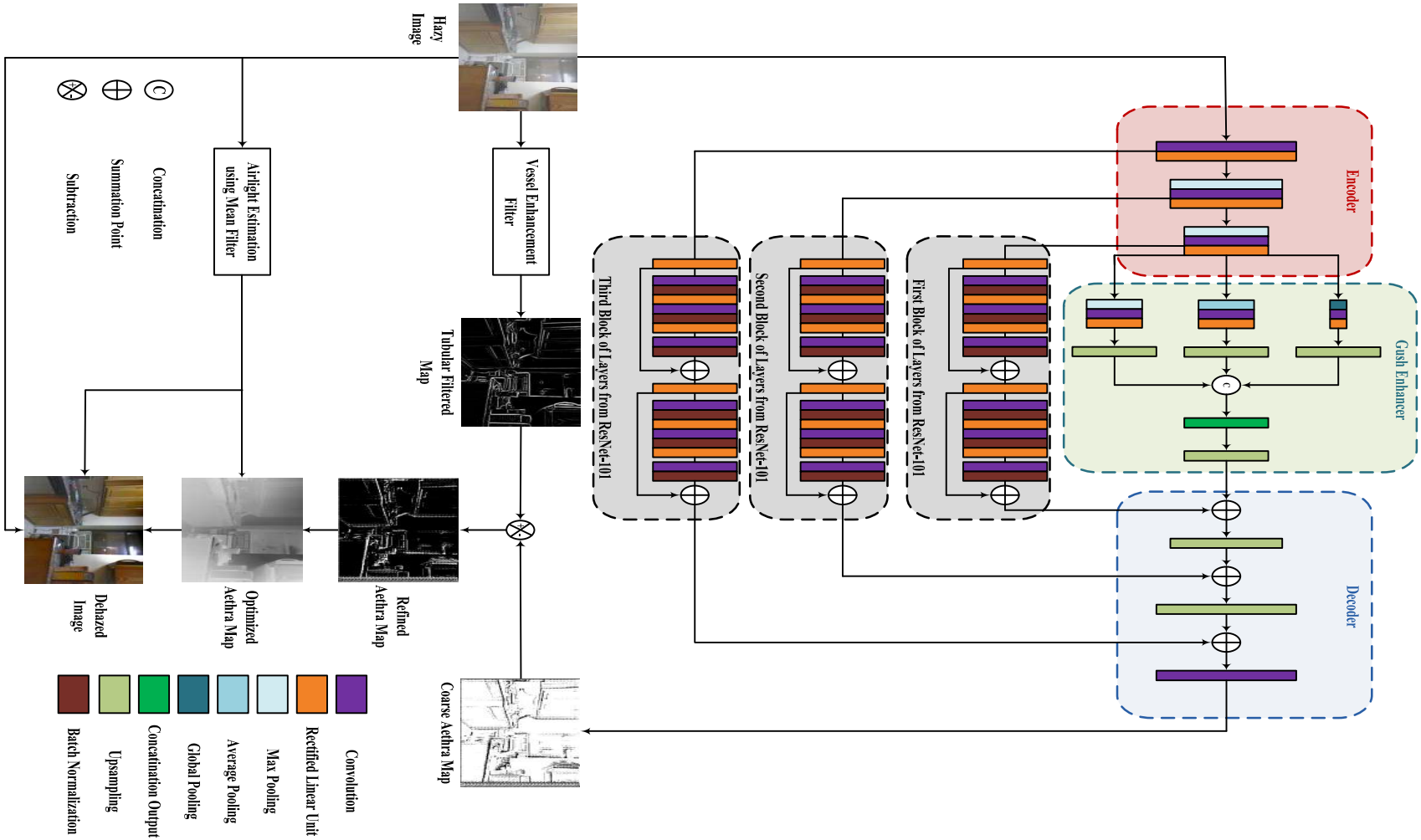


Fig. 5.1 Aethra-Net architecture

Table 5.1 Layer-wise analysis of encoder in Aethra-Net architecture

Type	No. of Filters	Size of Filters	Stride	Padding	Activations	Learnable Parameters	Total Learnables
Input Layer	-	-	-	-	256×256×1	-	0
Convolution	8	3×3×1	[1 1]	Same	256×256×8	Weights: 3×3×1×8 Bias: 1×1×8	80
ReLU	-	-	-	-	256×256×8	-	0
Max Pooling	-	-	[2 2]	[0 0 0 0]	128×128×8	-	0
Convolution	8	3×3×8	[1 1]	Same	128×128×8	Weights: 3×3×8×8 Bias: 1×1×8	584
ReLU	-	-	-	-	128×128×8	-	0
Max Pooling	-	-	[2 2]	[0 0 0 0]	64×64×8	-	0
Convolution	8	3×3×8	[1 1]	Same	64×64×8	Weights: 3×3×8×8 Bias: 1×1×8	584
ReLU	-	-	-	-	64×64×8	-	0

The decoder unit consists of one transpose convolution layer and two upsampling layers. The transpose convolution layer has been used to reduce the trainable parameters by optimizing the number of filters. The upsampling layers increase the size of the feature map because the target size of the map is the same as the size of the input image.

Table 5.3 depicts the decoder unit of the Aethra-Net model. The enhanced feature map has been obtained by connecting encoder and decoder units using three blocks of unweighted layers from the deep residual network with 101 layers called ResNet-101 [199]. The number of filters in these layers was reduced to lessen the computational complexity of the model. It also decreases the number of learnable parameters in order to save memory during model training. The primary advantage of the residual network is the elimination of the vanishing gradient problem. The residual connections reduce the possibility of data loss in low-level features during forward propagation. As a result, the color and edge distortions are minimized. Tables 5.4 – 5.6 show a detailed layer-wise analysis of three ResNet-101 blocks.

- **Vessel Enhancement Filtering**

Angiography is a medical imaging technique that visualizes the flow of blood through blood vessels in the human body [200]. It has been performed to enhance the

Table 5.2 Layer-wise analysis of gush enhancer in Aethra-Net architecture

Type	No. of Filters	Size of Filters	Stride	Padding	Activations	Learnable Parameters	Total Learnables
High-Level Feature Extractor							
Global Average Pooling	-	1×1	-	-	1×1×8	-	0
Convolution	8	1×1	[1 1]	Same	1×1×8	Weights: 1×1×8×8 Bias: 1×1×8	72
ReLU	-	-	-	-	1×1×8	-	0
Upsampling	8	64×64	[2 2]	-	64×64×8	Weights: 64×64×8×8 Bias: 1×1×8	262152
Medium-Level Feature Extractor							
Average Pooling	-	4×4	[2 2]	[0 0 0 0]	31×31×8	-	0
Convolution	8	3×3	[1 1]	Same	31×31×8	Weights: 3×3×8×8 Bias: 1×1×8	584
ReLU	-	-	-	-	31×31×8	-	0
Upsampling	8	4×4	[2 2]		64×64×8	Weights: 4×4×8×8 Bias: 1×1×8	1032
Low-Level Feature Extractor							
Max Pooling	-	2×2	[2 2]	[0 0 0 0]	32×32×8	-	0
Convolution	8	3×3	[1 1]	Same	32×32×8	Weights: 3×3×8×8 Bias: 1×1×8	584
ReLU	-	-	-	-	32×32×8	-	0
Upsampling	8	2×2	[2 2]		64×64×8	Weights: 2×2×8×8 Bias: 1×1×8	264
Feature Combination							
Depth Concatenation	-	-	-	-	64×64×24	-	0
Upsampling	8	1×1	[1 1]	-	64×64×8	Weights: 1×1×8×24 Bias: 1×1×8	200

vessel structures that are obtained during vessel segmentation [198]. The vascular morphology approach has been used to maximize the intensity of the vessels in a tubular structure. Edge-detection filters such as the Canny filter, Prewitt filter, and Sobel filter are used to segment images into discontinuity regions [201].

In this chapter, vessel enhancement filtering has been performed to visualize and analyze the objects that are located at a long distance away from the camera lens. It is a type of edge detection filter that is used in biomedical image processing. It can be used

to generate feature maps in non-biomedical applications because it is mainly used to detect the items represented as tubular vessels. It highlights the geometry of items such as cars, buildings, and trees in the form of tubular structures to enhance the visibility of items that are partially hidden due to clustering. It is computed as [198]:

$$I(\rho_0 + \delta\rho_0, s) \approx I(\rho_0, s) + \delta\rho_0^T \nabla_{o,s} + \delta\rho_0^T \, \mathrm{H}_{o,s}\delta\rho_0 \tag{5.2}$$

where $\nabla_{o,s}$ and $\mathrm{H}_{o,s}$ are the gradient vector and Hessian matrix obtained from the hazy image I, respectively. These parameters are computed in ρ_0 at scale s.

The differential operators of I are inspired from the linear scale space model [202], [203].

The convolution of a hazy image with derivatives of Gaussian is called differentiation of a hazy image. It is computed as follows [198]:

$$\frac{\partial}{\partial\rho} I(\rho, s) = s^{\gamma_{nd}} F(\rho) * \frac{\partial}{\partial\rho} G(\rho, s) \tag{5.3}$$

where γ_{nd} represents the family of normalized derivatives [204].

It is set to unity when no scale is preferred. D_1-dimensional Gaussian is defined as:

$$G(\rho, s) = \frac{1}{\sqrt{(2\pi s^2)}^{D_1}} e^{\frac{-||\rho||^2}{2s^2}} \tag{5.4}$$

Hessian matrix is represented as:

$$\delta\rho_0^T \, \mathrm{H}_{o,s}\delta\rho_0 = \left(\frac{\partial}{\partial\delta\rho_0}\right)\left(\frac{\partial}{\partial\delta\rho_0}\right) I(\rho_0, s) \tag{5.5}$$

For a 2D image, the Hessian matrix's eigenvalues λ_1 and λ_2 are calculated.

Table 5.3 Layer-wise analysis of decoder in Aethra-Net architecture

Type	No. of Filters	Size of Filters	Stride	Padding	Activations	Learnable Parameters	Total Learnables
Addition	-	-	-	-	64×64×8	-	0
Upsampling	8	2×2	[2 2]	-	128×128×8	Weights: 2×2×8×8 Bias: 1×1×8	264
Addition	-	-	-	-	128×128×8	-	0
Upsampling	8	2×2	[2 2]	-	256×256×8	Weights: 2×2×8×8 Bias: 1×1×8	264
Addition	-	-	-	-	256×256×8	-	0
Convolution	1	3×3	[1 1]	Same	256×256×1	Weights: 3×3×8 Bias: 1×1	73
Tanh	-	-	-	-	256×256×1	-	0
Regression Output	-	-	-	-	256×256×1	-	0

Table 5.4 Layer-wise analysis of first block of ResNet-101 in Aethra-Net architecture

Type	No. of Filters	Size of Filters	Stride	Padding	Activations	Learnable Parameters	Total Learnables
ReLU	-	-	-	-	64×64×8	-	0
Convolution	256	1×1	[1 1]	[0 0 0 0]	64×64×256	Weights: 1×1×8×256 Bias: 1×1×256	2304
Batch Normalization	-	-	-	-	64×64×256	Offset: 1×1×256 Scale: 1×1×256	512
ReLU	-	-	-	-	64×64×256	-	0
Convolution	256	3×3	[1 1]	[1 1 1 1]	64×64×256	Weights: 3×3×256×256 Bias: 1×1×256	590080
Batch Normalization	-	-	-	-	64×64×256	Offset: 1×1×256 Scale: 1×1×256	512
ReLU	-	-	-	-	64×64×256	-	0
Convolution	8	1×1	[1 1]	[0 0 0 0]	64×64×8	Weights: 1×1×256×8 Bias: 1×1×8	2056
Batch Normalization	-	-	-	-	64×64×8	Offset: 1×1×8 Scale: 1×1×8	16
Addition	-	-	-	-	64×64×8	-	0
ReLU	-	-	-	-	64×64×8	-	0
Convolution	256	1×1	[1 1]	[0 0 0 0]	64×64×256	Weights: 1×1×8×256 Bias: 1×1×256	2304
Batch Normalization	-	-	-	-	64×64×256	Offset: 1×1×256 Scale: 1×1×256	512
ReLU	-	-	-	-	64×64×256	-	0
Convolution	256	3×3	[1 1]	[1 1 1 1]	64×64×256	Weights: 3×3×256×256 Bias: 1×1×256	590080
Batch Normalization	-	-	-	-	64×64×256	Offset: 1×1×256 Scale: 1×1×256	512
ReLU	-	-	-	-	64×64×256	-	0
Convolution	8	1×1	[1 1]	[0 0 0 0]	64×64×8	Weights: 1×1×256×8 Bias: 1×1×8	2056
Batch Normalization	-	-	-	-	64×64×8	Offset: 1×1×8 Scale: 1×1×8	16
Addition	-	-	-	-	64×64×8	-	0

Table 5.5 Layer-wise analysis of second block of ResNet-101 in Aethra-Net architecture

Type	No. of Filters	Size of Filters	Stride	Padding	Activations	Learnable Parameters	Total Learnables
ReLU	-	-	-	-	128×128×8	-	0
Convolution	256	1×1	[1 1]	[0 0 0 0]	128×128×256	Weights: 1×1×8×256 Bias: 1×1×256	2304
Batch Normalization	-	-	-	-	128×128×256	Offset: 1×1×256 Scale: 1×1×256	512
ReLU	-	-	-	-	128×128×256	-	0
Convolution	256	3×3	[1 1]	[1 1 1 1]	128×128×256	Weights: 3×3×256×256 Bias: 1×1×256	590080
Batch Normalization	-	-	-	-	128×128×256	Offset: 1×1×256 Scale: 1×1×256	512
ReLU	-	-	-	-	128×128×256	-	0
Convolution	8	1×1	[1 1]	[0 0 0 0]	128×128×8	Weights: 1×1×256×8 Bias: 1×1×8	2056
Batch Normalization	-	-	-	-	128×128× 8	Offset: 1×1×8 Scale: 1×1×8	16
Addition	-	-	-	-	128×128×8	-	0
ReLU	-	-	-	-	128×128×8	-	0
Convolution	256	1×1	[1 1]	[0 0 0 0]	128×128×256	Weights: 1×1×8×256 Bias: 1×1×256	2304
Batch Normalization	-	-	-	-	128×128×256	Offset: 1×1×256 Scale: 1×1×256	512
ReLU	-	-	-	-	128×128×256	-	0
Convolution	256	3×3	[1 1]	[1 1 1 1]	128×128×256	Weights: 3×3×256×256 Bias: 1×1×256	590080
Batch Normalization	-	-	-	-	128×128×256	Offset: 1×1×256 Scale: 1×1×256	512
ReLU	-	-	-	-	128×128×256	-	0
Convolution	8	1×1	[1 1]	[0 0 0 0]	128×128×8	Weights: 1×1×256×8 Bias: 1×1×8	2056
Batch Normalization	-	-	-	-	128×128×8	Offset: 1×1×8 Scale: 1×1×8	16
Addition	-	-	-	-	128×128×8	-	0

Table 5.6 Layer-wise analysis of third block of ResNet-101 in Aethra-Net architecture

Type	No. of Filters	Size of Filters	Stride	Padding	Activations	Learnable Parameters	Total Learnables
ReLU	-	-	-	-	256×256×8	-	0
Convolution	256	1×1	[1 1]	[0 0 0 0]	256×256×256	Weights: 1×1×8×256 Bias: 1×1×256	2304
Batch Normalization	-	-	-	-	256×256×256	Offset: 1×1×256 Scale: 1×1×256	512
ReLU	-	-	-	-	256×256×256	-	0
Convolution	256	3×3	[1 1]	[1 1 1 1]	256×256×256	Weights: 3×3×256×256 Bias: 1×1×256	590080
Batch Normalization	-	-	-	-	256×256×256	Offset: 1×1×256 Scale: 1×1×256	512
ReLU	-	-	-	-	256×256×256	-	0
Convolution	8	1×1	[1 1]	[0 0 0 0]	256×256×8	Weights: 1×1×256×8 Bias: 1×1×8	2056
Batch Normalization	-	-	-	-	256×256× 8	Offset: 1×1×8 Scale: 1×1×8	16
Addition	-	-	-	-	256×256×8	-	0
ReLU	-	-	-	-	256×256×8	-	0
Convolution	256	1×1	[1 1]	[0 0 0 0]	256×256×256	Weights: 1×1×8×256 Bias: 1×1×256	2304
Batch Normalization	-	-	-	-	256×256×256	Offset: 1×1×256 Scale: 1×1×256	512
ReLU	-	-	-	-	256×256×256	-	0
Convolution	256	3×3	[1 1]	[1 1 1 1]	256×256×256	Weights: 3×3×256×256 Bias: 1×1×256	590080
Batch Normalization	-	-	-	-	256×256×256	Offset: 1×1×256 Scale: 1×1×256	512
ReLU	-	-	-	-	256×256×256	-	0
Convolution	8	1×1	[1 1]	[0 0 0 0]	256×256×8	Weights: 1×1×256×8 Bias: 1×1×8	2056
Batch Normalization	-	-	-	-	256×256×8	Offset: 1×1×8 Scale: 1×1×8	16
Addition	-	-	-	-	256×256×8	-	0

The tubular structure appears bright in a dark environment, or dark in a bright environment based on the numerical values of the eigenvalues. Table 5.7 summarizes the relationship between eigenvalues, environment, and the type of tubular structure.

- **Optimization of Transmission Map**

The course Aethra map and tubular enhanced filters have been mathematically subtracted to obtain the refined Aethra map. The subtraction has been performed to eliminate the bright features that interfere with the airlight computation. Thus, the computational complexity of the model has been reduced. The estimated airlight and refined Aethra map have been utilized to obtain the optimized Aethra map.

5.2.3 Visibility Restoration in Image

After estimating atmospheric light and obtaining the transmission map, the hazy image has been recovered by limiting the amount of light contained in the image, which is known as gamma. Assuming $t(x, y)$ as minimum exposure of light, the computed image is mathematically defined as:

$$J(x,y) = \frac{I(x,y)-A}{[max(t(x,y),t_0(x,y))]^{gamma}} + A \tag{5.6}$$

5.3 Video Dehazing

This section discusses the analysis of hazy videos that includes frame extraction, frame processing, and video reconstruction. The proposed architecture has been tested on two systems called Syst-1 and Syst-2. Syst-1 is equipped with 8GB of RAM and a 64-bit operating system powered by an Intel (R) Core (TM) i7-6700 CPU. Syst-2 is equipped with 8GB of RAM and a 64-bit operating system powered by an Intel (R) Core (TM) i5-1035G1 CPU. The computer vision toolbox and image processing toolbox of MATLAB R2021a have been utilized to perform the desired task.

In video processing, frames have been extracted from the hazy video. They have been computed by implementing the proposed approach to generate dehazed results.

Table 5.7 Relationship between eigenvalues and the type of tubular structure

Eigenvalue 1 (λ_1)	Eigenvalue 2 (λ_2)	Environment	Type of tubular structure
Low and positive	High and negative	Dark	Bright
Low and positive	High and positive	Bright	Dark

The proposed dehazing method has been implemented for one frame at a time. The frames have been processed in the correct sequence to reconstruct the video.

5.4 Results and Discussion

This section discusses the image and video dehazing results obtained from Aethra-Net for both real and synthetic test sets. The comparative analysis has been done with the existing dehazing techniques such as GCANet, NLD, NLBF, CCDI, MSBDN, AECL, and MAMLP.

5.4.1 Training Details

The proposed model was trained using 1900 images from different benchmark datasets, including FRIDA, FRIDA-2, RESIDE, NH-HAZE, O-HAZE, and HUDRS. The images have been resized to 256×256 pixels. It has been further tested on 700 images categorized into seven test-sets namely, IMSET-9, IMSET-10, IMSET-11, IMSET-12, IMSET-13, IMSET-14, and IMSET-15.

Aethra-Net has been used to analyze the intermediate steps of the proposed dehazing approach on sample hazy images IMGN51 through IMGN55 from IMSET-9, as shown in Figure 5.2. IMSET-10 consists of FRIDA sample test images IMGN56 and IMGN57, as depicted in Figure 5.3. IMSET-11 consists of IMGN58 and IMGN59 from FRIDA2 as depicted in Figure 5.4. IMSET-12 consists of NH-HAZE sample test images IMGN60 and IMGN61, as depicted in Figure 5.5. IMSET-13 consists of IMGN62 and IMGN63 from RESIDE as depicted in Figure 5.6. IMSET-14 is made up of IMGN64 and IMGN65 from HUDRS, as seen in Figure 5.7. IMSET-15 is made up of IMGN27 and IMGN29 from O-HAZE, as seen in Figure 5.8.

To discuss the versatility of the proposed architecture, different challenging scenarios have been discussed. The images IMGN56 through IMGN59 have non-homogenous haze. The images IMGN27, IMGN29, IMGN60, and IMGN61, are densely hazed. The performance of existing techniques has also been evaluated on less hazy images such as IMGN62, IMGN63, IMGN64, and IMGN65.

In this model, the Adam optimizer has been used with mean square error (MSE), an initial learning rate of 0.005, and a batch size of 5. MSE is commonly calculated in dehazing architectures due to its low computational cost. It can be calculated either

exclusively [205] or in combination with other losses [188]. The model has been trained for 300 epochs.

5.4.2 Analysis of Hazy Images

The qualitative and quantitative analysis of various hazy images from IMSET-9 to IMSET-15 has been performed. The computed results have been compared with the existing dehazing techniques.

- **Qualitative Analysis**

Figures 5.2 – 5.8 show the results computed by implementing Aethra-Net. The visual comparison with the existing dehazing techniques has been conducted. It has been observed that the results obtained by NLBF and CCDI are darker than their corresponding ground truth images. The results obtained by NLBF and GCANet suffer from color distortion. The dehazed images computed by NLD, CCDI, MSBDN, and AECL are brighter than their respective ground truth images.

The distant objects, such as trees and houses in IMGN56 to IMGN59, are only visible in the restored images when the proposed architecture has been used. The second-best result closer to ground truth is obtained using CCDI. However, the mountain in IMGN58 is not restored for any dehazing algorithm. Because of the dense haze, no mountain can be seen in the original hazy image. The depth of images IMGN27, IMGN29, IMGN60, and IMGN61, i.e., visibility obtained using NLD, NLBF, MSBDN, AECL, and MAMLP is almost the same. Although the results obtained using GCANet and CCDI are clearer as compared to other state-of-the-art algorithms. The dehazed result of IMGN62 computed using Aethra-Net has a gradient artifact in the sky region.

The results computed by Aethra-Net may suffer from color deviation. It is because low, medium, and high features have been extracted and upsampled to the exact dimensions before concatenation. The hazy image has been restored using a restoration factor called gamma. It has a high value for better dehazing. However, a much higher value of gamma can cause texture, color, or edge distortion. A smaller value of color difference implies less color distortion.

The comparative analysis of color deviation of the images computed by implementing Aethra-Net and other existing techniques for IMGN27, IMGN29, and IMGN56 through IMGN65 is depicted in Table 5.8. It has been observed that color

deviation for IMGN27, IMGN56, IMGN57, IMGN58, IMGN59, and IMGN63 is the least with values 12.645, 12.645, 0, 0, 2.42, and 25.238, respectively when Aethra-Net is implemented. The color deviation for IMGN58 and IMGN59 is also zero when CCDI is implemented. The value of color deviation for IMGN60 and IMGN61 is 34.867 and 30.992, which is the least when MAMLP is implemented. The least average color deviation for all sample images is 20.6581 when Aethra-Net is implemented.

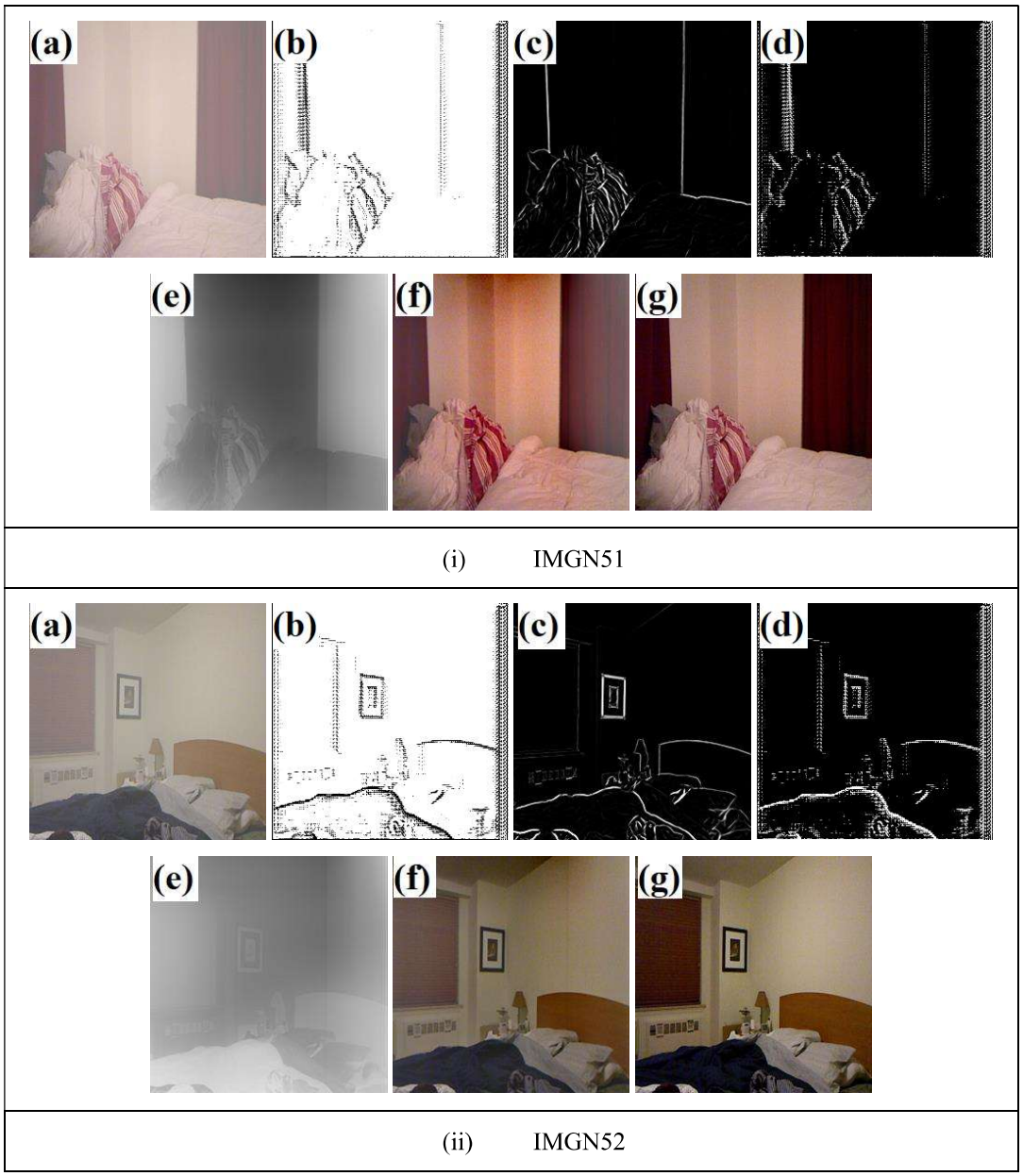
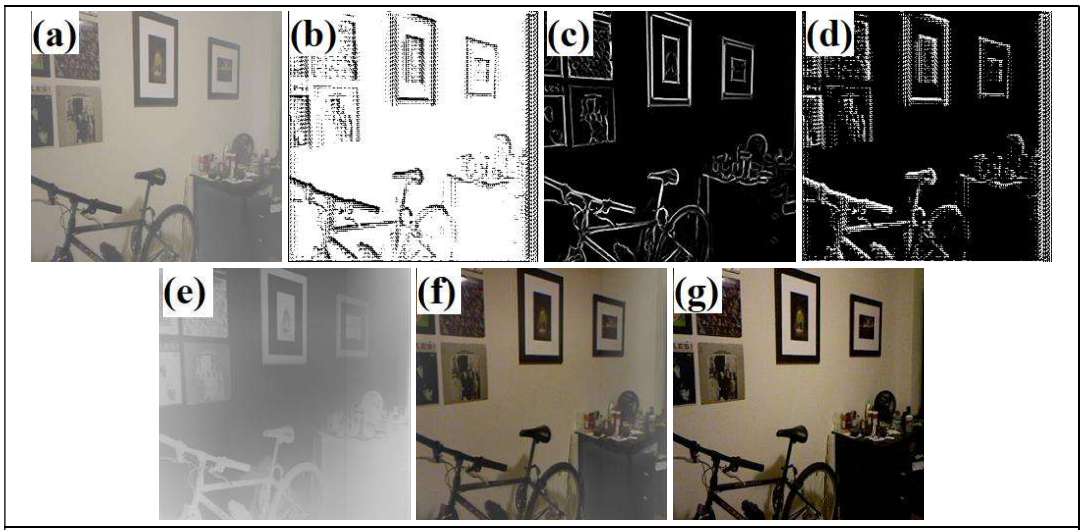
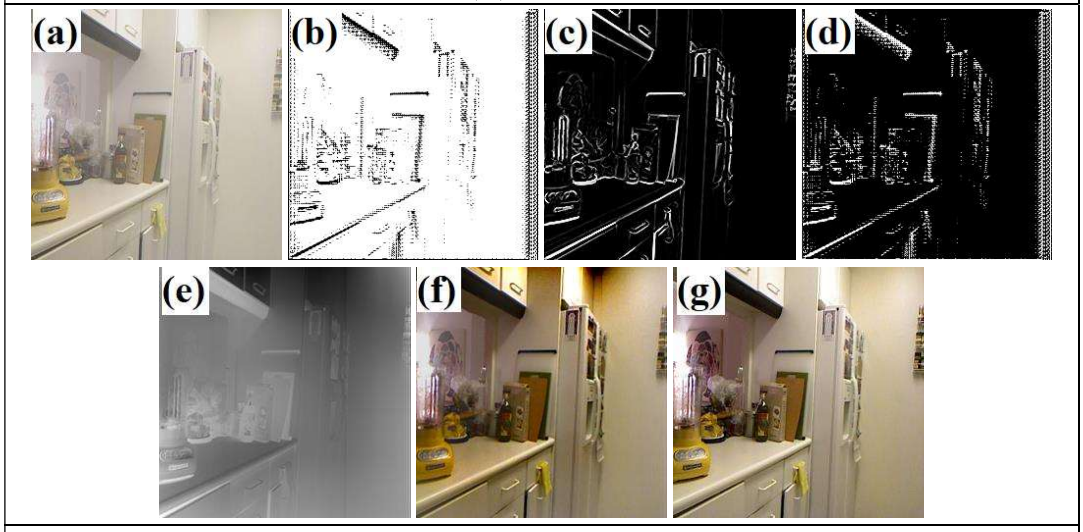


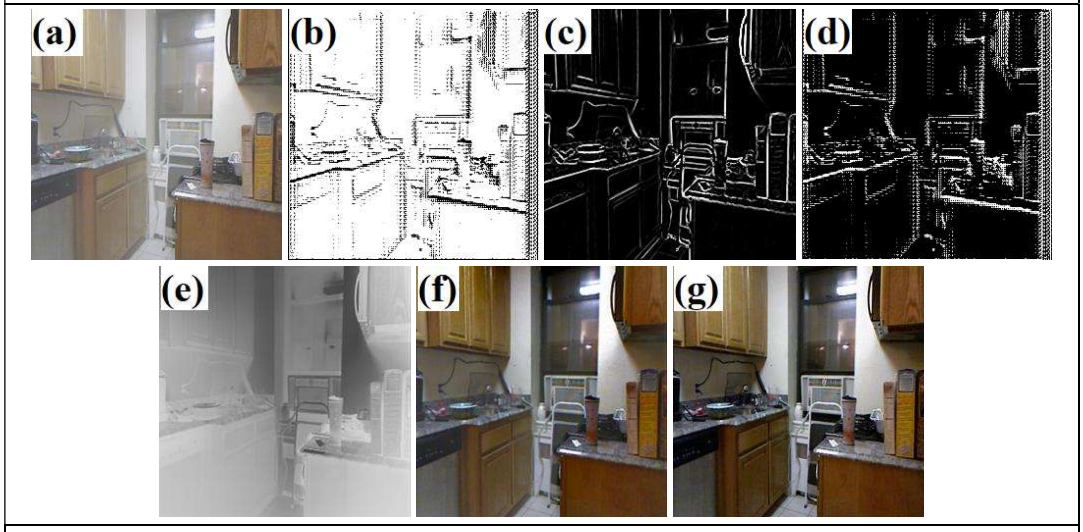
Fig. 5.2 Dehazing results of sample images from IMSET-9 using Aethra-Net (a) Hazy images, (b) Course Aethra map, (c) Tubular enhanced map, (d) Refined Aethra map, (e) Optimized Aethra map, (f) Dehazed output, and (g) Ground truth



(iii) IMGN53

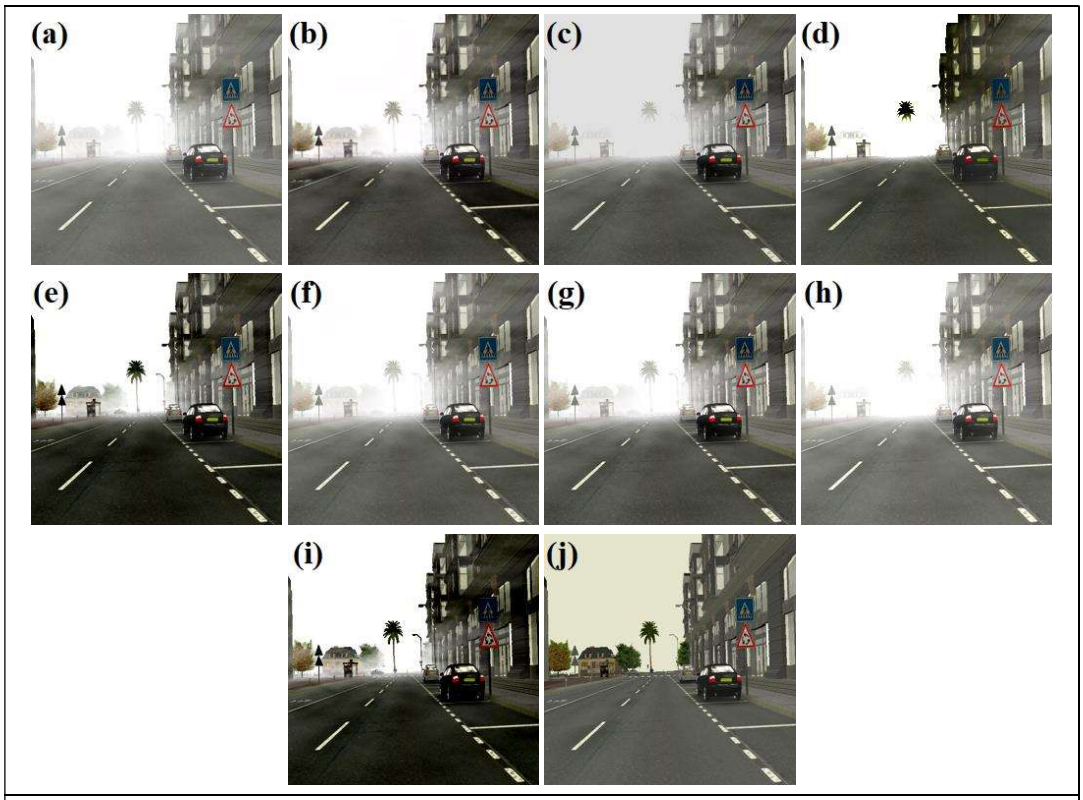


(iv) IMGN54

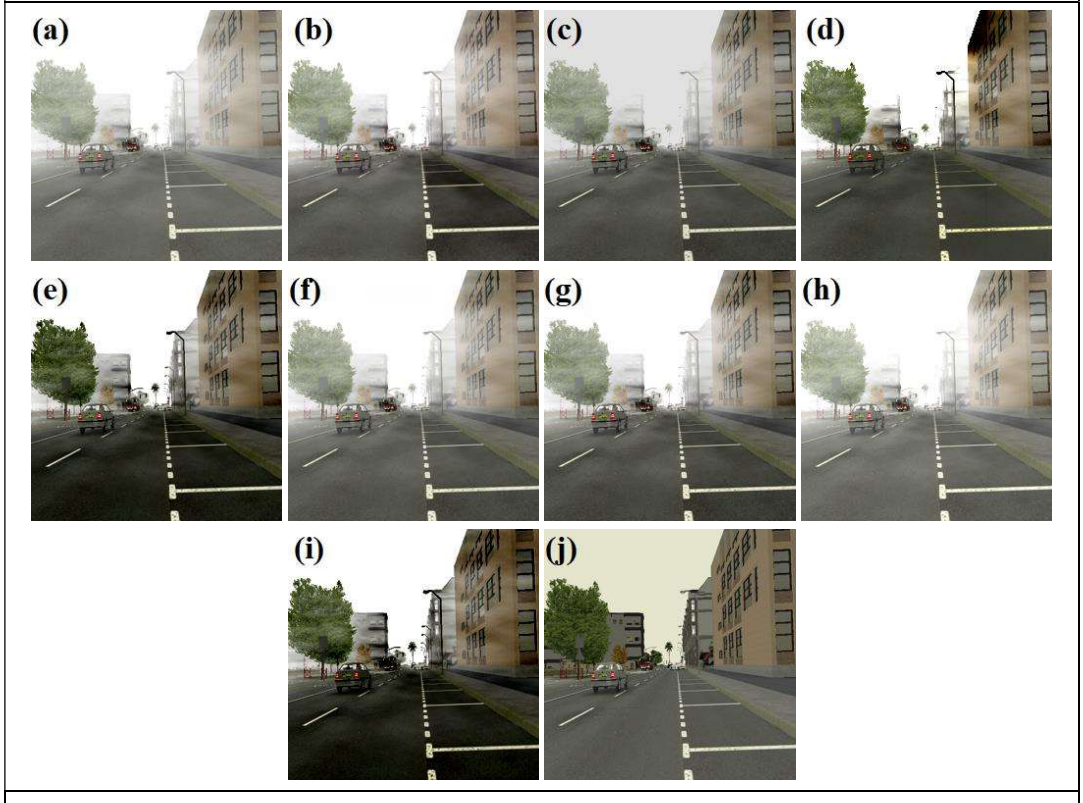


(v) IMGN55

Fig. 5.2 (continue) Dehazing results of sample images from IMSET-9 using Aethra-Net (a) Hazy images, (b) Course Aethra map, (c) Tubular enhanced map, (d) Refined Aethra map, (e) Optimized Aethra map, (f) Dehazed output, and (g) Ground truth

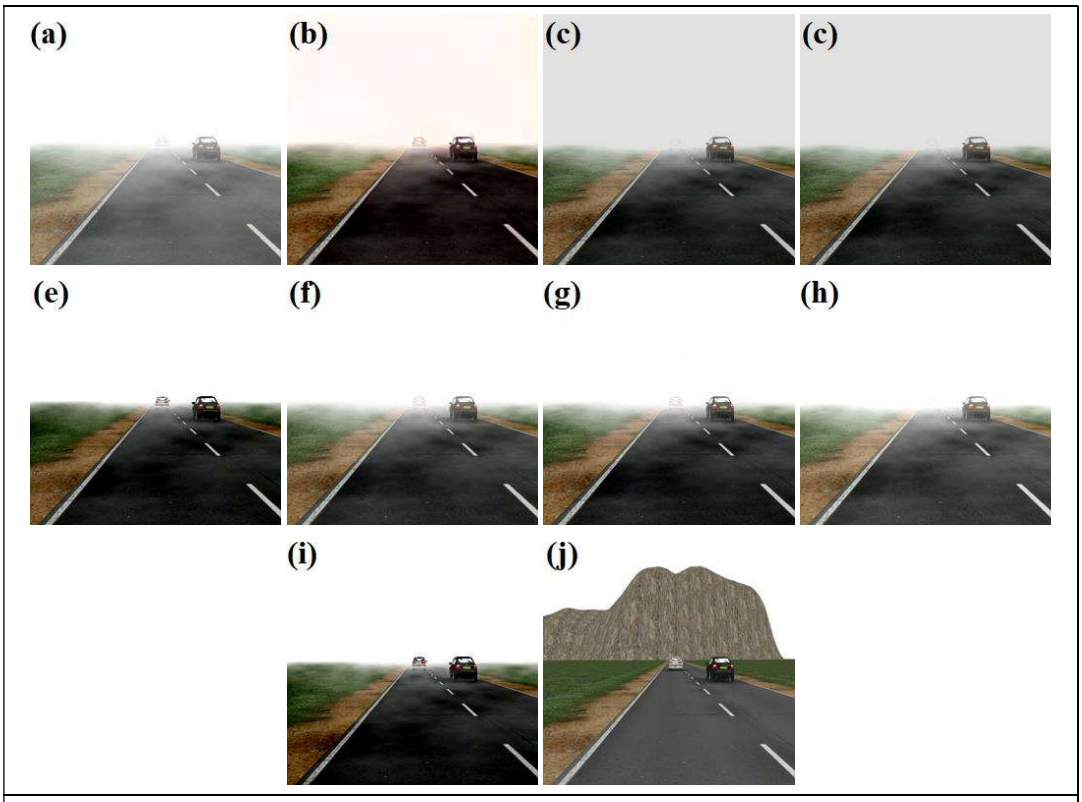


(i) IMGN56

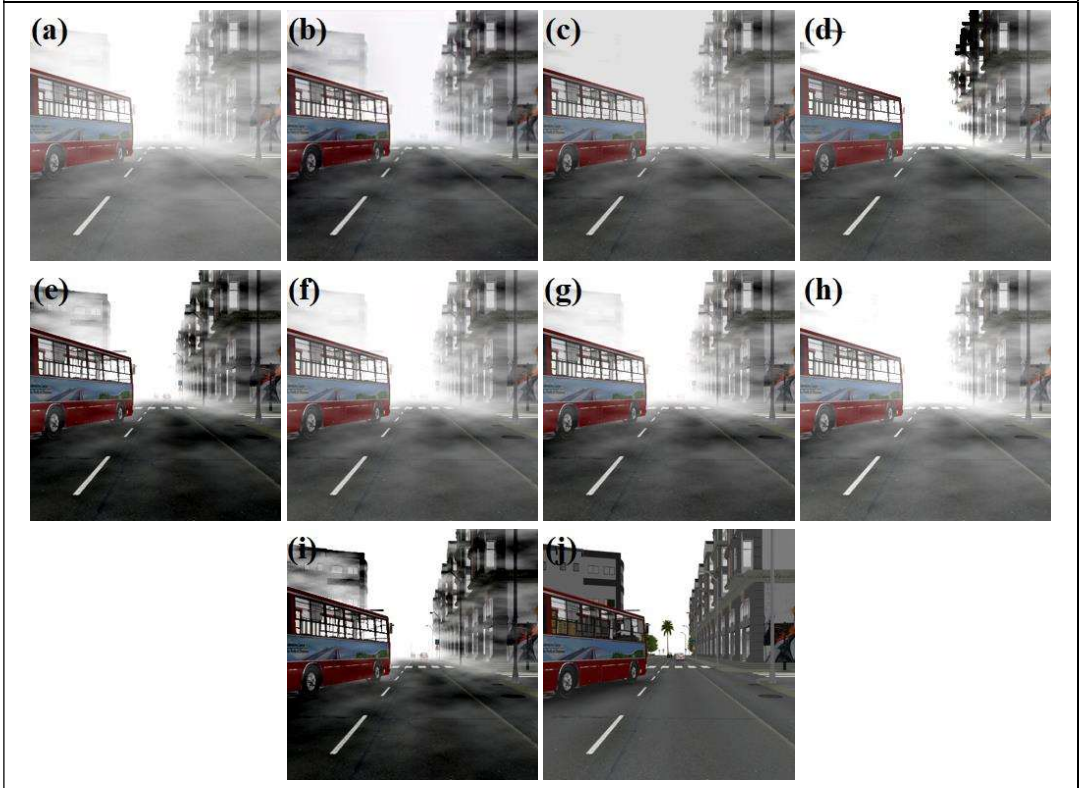


(ii) IMGN57

Fig. 5.3 Visual comparison of dehazed sample images from IMSET-10 computed by implementing Aethra-Net and existing dehazing techniques (a) Hazy image (b) GCANet (c) NLD (d) NLBF (e) CCDI (f) MSBDN (g) AECL (h) MAMLP (i) Aethra-Net and (j) Ground truth

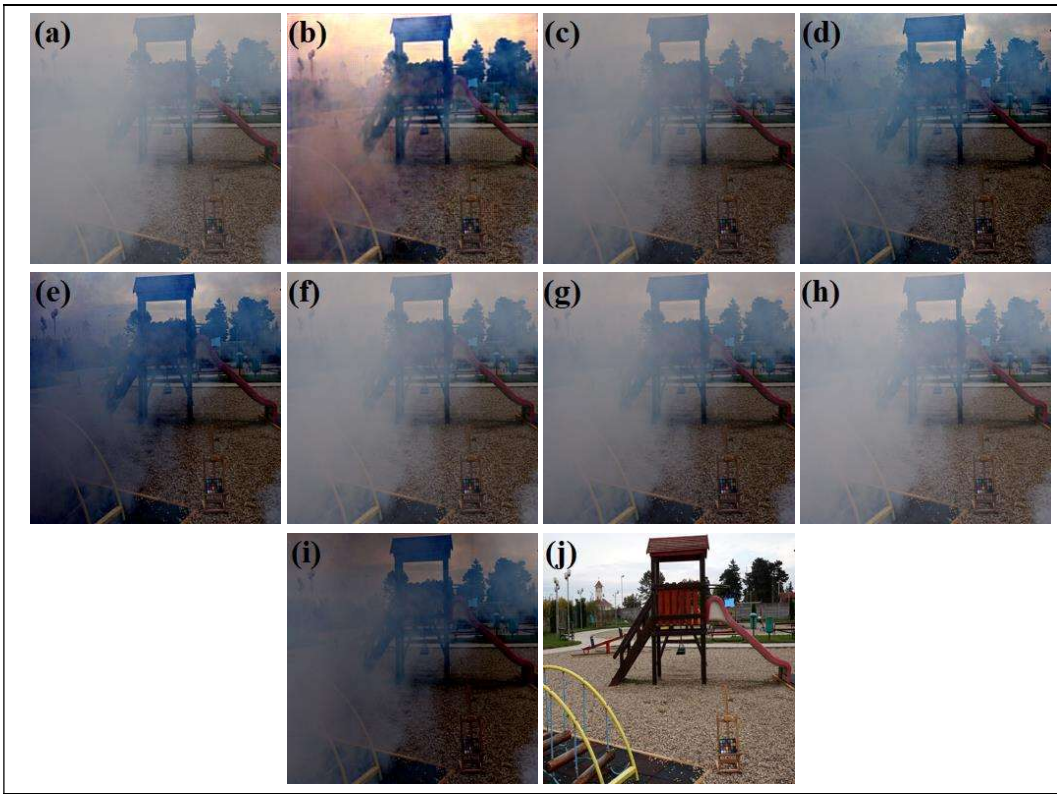


(i) IMGN58

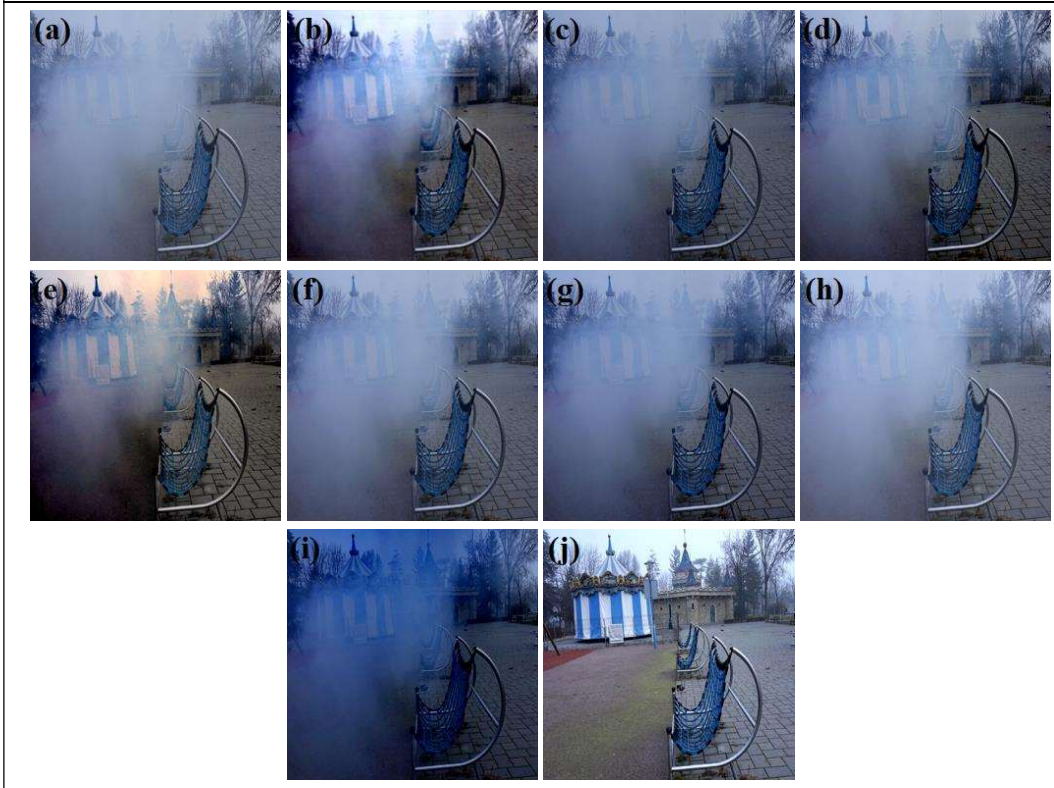


(ii) IMGN59

Fig. 5.4 Visual comparison of dehazed sample images from IMSET-11 computed by implementing Aethra-Net and existing dehazing techniques (a) Hazy image (b) GCANet (c) NLD (d) NLBF (e) CCDI (f) MSBDN (g) AECL (h) MAMLP (i) Aethra-Net and (j) Ground truth



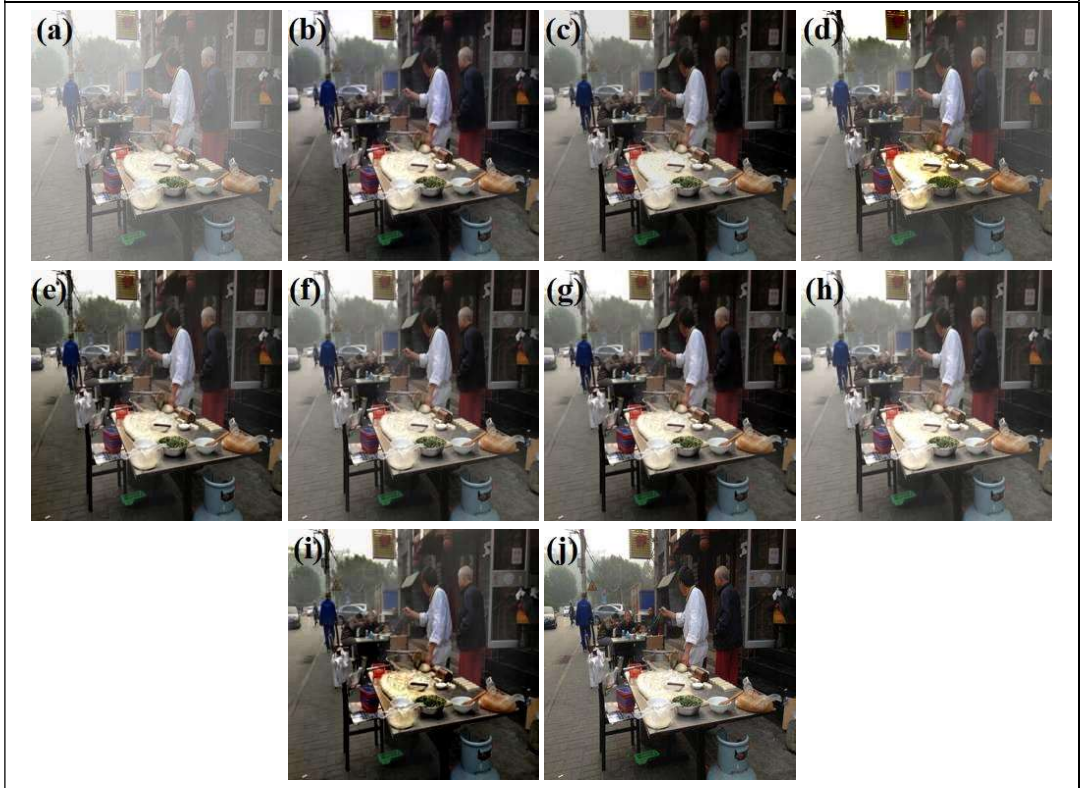
(i) IMGN60



(ii) IMGN61

Fig. 5.5 Visual comparison of dehazed sample images from IMSET-12 computed by implementing Aethra-Net and existing dehazing techniques (a) Hazy image (b) GCANet (c) NLD (d) NLBF (e) CCDI (f) MSBDN (g) AECL (h) MAMLP (i) Aethra-Net and (j) Ground truth

(i) IMGN62



(ii) IMGN63

Fig. 5.6 Visual comparison of dehazed sample images from IMSET-13 computed by implementing Aethra-Net and existing dehazing techniques (a) Hazy image (b) GCANet (c) NLD (d) NLBF (e) CCDI (f) MSBDN (g) AECL (h) MAMLP (i) Aethra-Net and (j) Ground truth

(i) IMGN64

(ii) IMGN65

Fig. 5.7 Visual comparison of dehazed sample images from IMSET-14 computed by implementing Aethra-Net and existing dehazing techniques (a) Hazy image (b) GCANet (c) NLD (d) NLBF (e) CCDI (f) MSBDN (g) AECL (h) MAMLP (i) Aethra-Net and (j) Ground truth

(i) IMGN27

(ii) IMGN29

Fig. 5.8 Visual comparison of dehazed sample images from IMSET-15 computed by implementing Aethra-Net and existing dehazing techniques (a) Hazy image (b) GCANet (c) NLD (d) NLBF (e) CCDI (f) MSBDN (g) AECL (h) MAMLP (i) Aethra-Net and (j) Ground truth

Table 5.8 Comparative analysis of Aethra-Net with existing dehazing techniques based on color deviation using CIE

Images	GCANet	NLD	NLBF	CCDI	MSBDN	AECL	MAMLP	Aethra-Net
IMGN56	20.412	14.21	18.645	18.645	18.899	19.013	18.645	**12.645**
IMGN57	20.065	14.21	18.645	18.645	18.809	19.213	18.645	**12.645**
IMGN58	5.675	10.469	0.005	**0**	6.666	1.011	6.545	**0**
IMGN59	9.075	10.469	0.005	**0**	7.225	2.313	7.667	**0**
IMGN60	39.537	46.208	51.15	66.748	42.377	45.866	**34.867**	59.348
IMGN61	44.62	38.005	39.362	50.592	34.229	39.322	**30.992**	46.333
IMGN62	16.554	9.624	**6.302**	12.599	7.891	8.431	12.478	11.285
IMGN63	17.397	16.469	6	8.42	7.658	7.656	6	**2.42**
IMGN64	14.505	14.854	15.335	27.266	**13.431**	16.432	19.844	22.717
IMGN64	15.616	**13.743**	16.435	25.434	15.643	15.643	21.745	19.764
IMGN27	26.804	29.759	30.966	31.938	28.232	27.474	37.187	**25.238**
IMGN29	**28.461**	44.269	45.056	46.448	47.534	50.962	36.115	35.503
Average	21.5600	21.8574	20.6588	25.5612	20.7161	21.1113	20.8941	**20.6581**

It is followed by NLBF and MSBDN, both of which have an average color deviation of 20.6588 and 20.7161, respectively. Thus, it can be said that desired qualitative results for all the images are not obtained by one technique. Thus, performance metrics have been evaluated for Aethra-Net and the existing SOTA techniques.

- **Quantitative Analysis**

The proposed model's performance has been compared to GCANet, NLD, NLBF, CCDI, MSBDN, AECL, and MAMLP. Figures 5.9 and 5.10 represent the comparative analysis of Aethra-Net and existing dehazing techniques based on $SSIM$ and $PSNR$ respectively. Various performance metrics such as $SSIM$, $PSNR$, and $BRISQUE$ have been calculated for IMGN27, IMGN29, and IMGN56 through IMGN65.

The highest average value of $SSIM$ computed for various dehazing algorithms is 0.709 when the proposed approach has been used. It is because the items in computed images (i.e., IMGN27, IMGN29, and IMGN56 through IMGN65) are more visible as compared to the results computed using the existing SOTA techniques. It is followed by MSBDN, NLD, and MAMLP with the average value of $SSIM$ as 0.642, 0.6309, and 0.6301, respectively.

The highest average value of $PSNR$ obtained by implementing Aethra-Net for various hazy images is 16.408. It has been observed from Table 5.9 that the results computed using Aethra-Net has less noise and higher signal strength because ResNet-101 blocks minimize the effect of vanishing gradient. It is followed by AECL, MSBDN, and GCANet with the average value of $PSNR$ as 16.168, 15.947, and 15.746, respectively.

An image with high contrast gain and a large percentage of saturated pixels does not seem natural and has a large $BRISQUE$. As depicted in Table 5.9, the value of $BRISQUE$ computed by implementing Aethra-Net is the least for images IMGN61, IMGN63, IMGN64, and IMGN27 as 6.4074, 9.2371, 15.2541, and 22.1465, respectively. The value of $BRISQUE$ computed by implementing AECL is the least for IMGN58, IMGN62, and IMGN65 as 27.4081, 21.8937, and 40.5595, respectively. The value of $BRISQUE$ computed by implementing MSBDN is the least for IMGN56, IMGN57, and IMGN59 as 32.5858, 23.931, and 23.1495, respectively. The value of $BRISQUE$ computed by implementing GCANet is the least for IMGN60 and IMGN29 as 12.7153 and 42.8586, respectively. Aethra-Net has the least value of $BRISQUE$ which is 28.12205, followed by AECL, MSBDN and GCANet, which have average values of 32.429, 33.435, and 36.987, respectively.

5.4.3 Analysis of Hazy Videos

In this section, the analysis of the dehazed frames computed by Aethra-Net for hazy videos has been presented. Hazy videos have been obtained by browsing the internet. The frames have been extracted from the videos and processed in the same order. Figure 5.11 depicts the sampled dehazed frames of the test set IMSET-16 extracted by video processing using the proposed approach.

- **Calculation of Natural Losses**

$BRISQUE$ has been calculated for dehazed frames obtained by processing the frames obtained from hazy videos. Figure 5.12 depicts the comparative analysis of the computational time of frames obtained from Syst-1 and Syst-2 for identical images with varied resolutions. Table 5.10 depicts the computed value of $BRISQUE$ for dehazed frames. It has been observed that the average value of $BRISQUE$ is 53.832. It is significantly greater than the value of $BRISQUE$ obtained from IMSET-10 to IMSET-15. This is due to the test videos were shot in real-world dense fog.

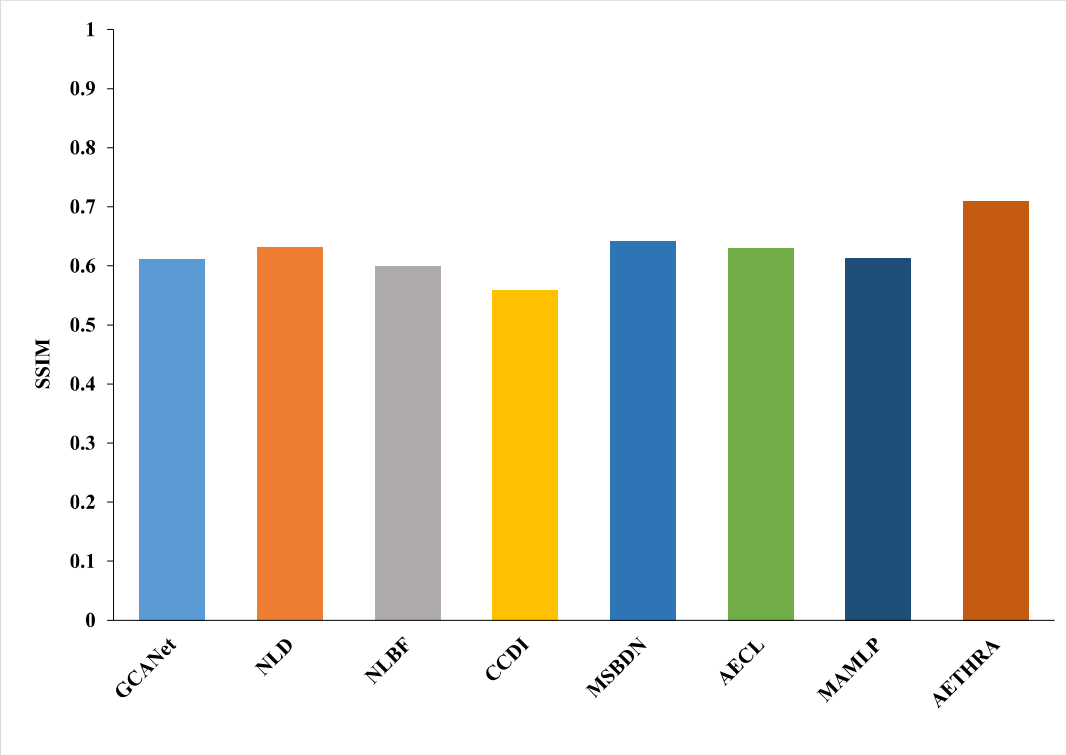


Fig. 5.9 Comparative analysis of Aethra-Net with existing dehazing techniques based on $SSIM$

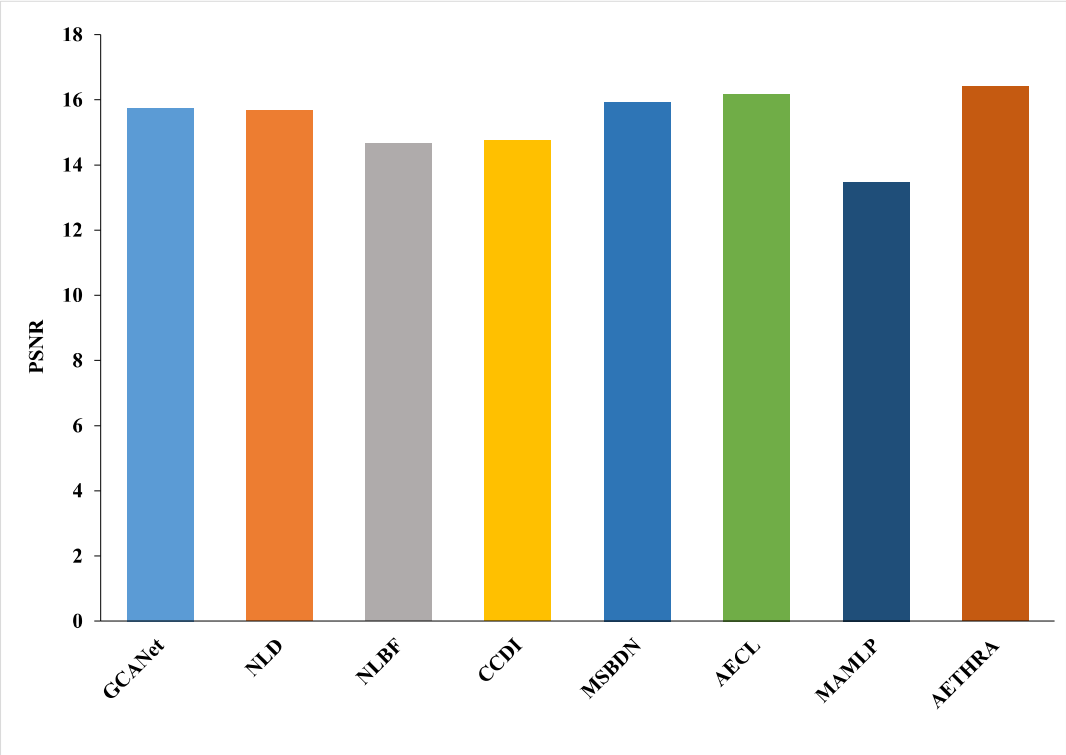


Fig. 5.10 Comparative analysis of Aethra-Net with existing dehazing techniques based on the $PSNR$

Table 5.9 Comparative analysis of Aethra-Net with existing dehazing techniques based on $BRISQUE$

Images	GCANet	NLD	NLBF	CCDI	MSBDN	AECL	MAMLP	Aethra-Net
IMGN56	34.7128	43.3153	44.6778	44.7283	**32.5858**	32.59	44.0459	34.9742
IMGN57	37.1077	37.6681	37.3667	40.982	**23.931**	24.9426	38.0251	33.5738
IMGN58	52.5747	56.9723	56.8638	56.9257	38.233	**27.4081**	57.0401	46.6313
IMGN59	27.3486	38.9704	43.6099	40.2135	**23.1495**	23.4972	39.3145	28.6664
IMGN60	**12.7153**	29.0579	27.9988	23.6109	29.0887	27.1434	29.0887	17.2569
IMGN61	12.5592	29.4949	18.0134	24.743	29.614	22.7496	29.614	**6.4074**
IMGN62	44.2867	44.2716	45.9147	35.1614	27.3491	**21.8937**	44.2734	33.6829
IMGN63	41.0683	24.5858	19.9402	31.9768	34.3927	38.8156	24.8748	**9.2371**
IMGN64	46.3942	52.6739	56.4474	27.7616	31.0776	38.8012	52.8703	**15.2541**
IMGN65	47.2707	56.9212	58.9468	42.335	41.5371	**40.5595**	57.1177	46.1758
IMGN27	44.9555	43.9389	44.1037	43.1931	44.1723	44.6891	43.9387	**22.1465**
IMGN29	**42.8586**	46.0894	46.0894	53.4581	46.0898	46.0581	46.0894	43.4582
Average	36.98769	41.99664	41.66438	38.75745	33.43505	32.42901	42.19105	**28.12205**

- **Computation Analysis**

Image is represented as the matrix of M rows and N columns such that each element of the matrix is called a pixel. As a result, the resolution is equal to the product of the number of rows and columns in an image. The image resolution and computation time have a linear relationship, which means that as the spatial dimension of the image increases, so will the computation time. The processing speed of computer vision problems is largely influenced by the random access memory (RAM) of the system. A system with less RAM takes longer to compute video dehazing as compared to a system with more RAM. The comparative analysis of the computational time for IMGN66, IMGN67, IMGN68, IMGN69, and IMGN70 has been presented in Figure 5.12. The results have been computed by using Syst-1 and Syst-2 for different resolutions. It has been observed that Syst-1 takes less time (in seconds) than Syst-2 to process the same frames.

5.5 Ablation Studies

Several ablation experiments have been carried out to validate the contribution of different feature extractors, namely high-level, medium-level, and low-level feature

Video Frames	Hazy Frames	Dehazed Frames
IMGN66		
IMGN67		
IMGN68		
IMGN69		
IMGN70		

Fig. 5.11 Computed results of hazy frames using Aethra-Net

Table 5.10 Calculation of $BRISQUE$ for dehazed video frames computed using Aethra-Net

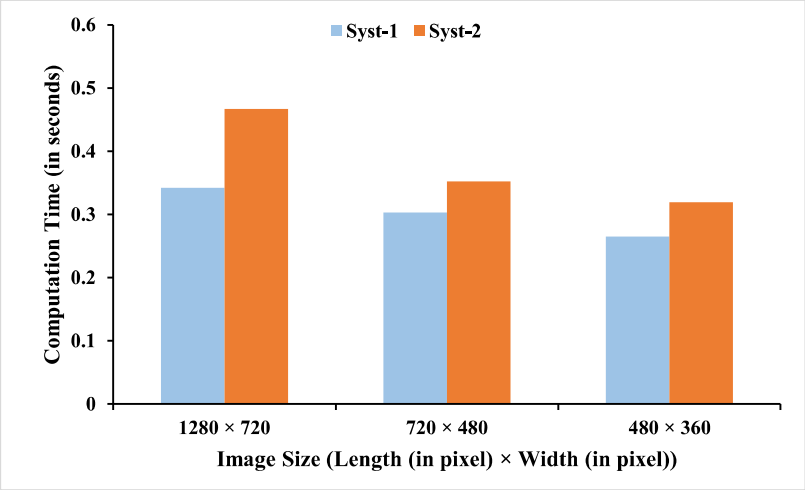
Video Frames	IMGN66	IMGN67	IMGN68	IMGN69	IMGN70
$BRISQUE$	52.349	52.209	50.577	58.091	55.937

extractors in the proposed model. In these experiments, different components were removed in each training. It mainly includes the following steps:

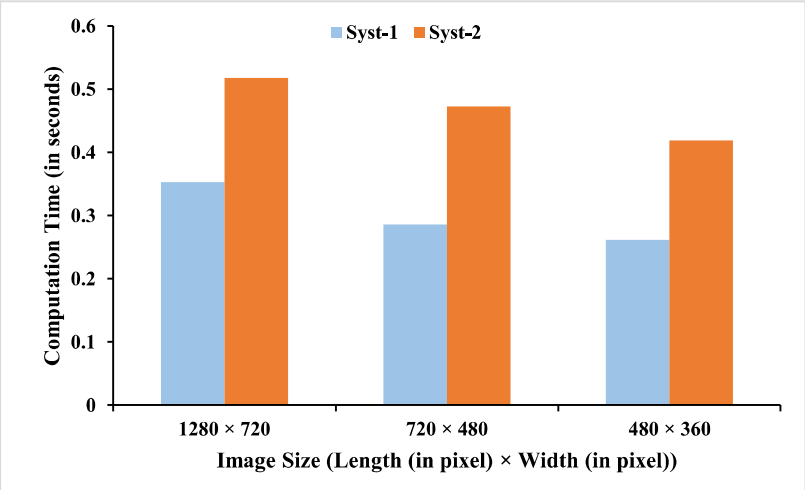
(1) The original image has been restored with extraction of high-level features only.

(2) The original image has been restored with extraction of medium-level features only.

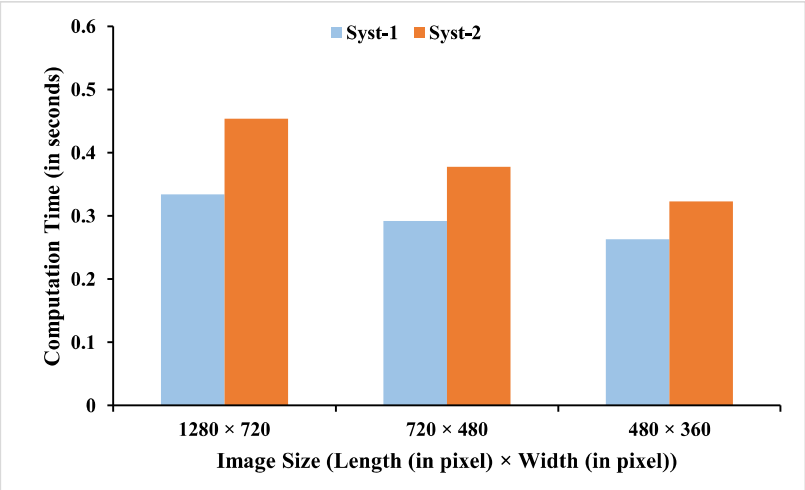
(3) The original image has been restored with extraction of low-level features only.



(a) IMGN66

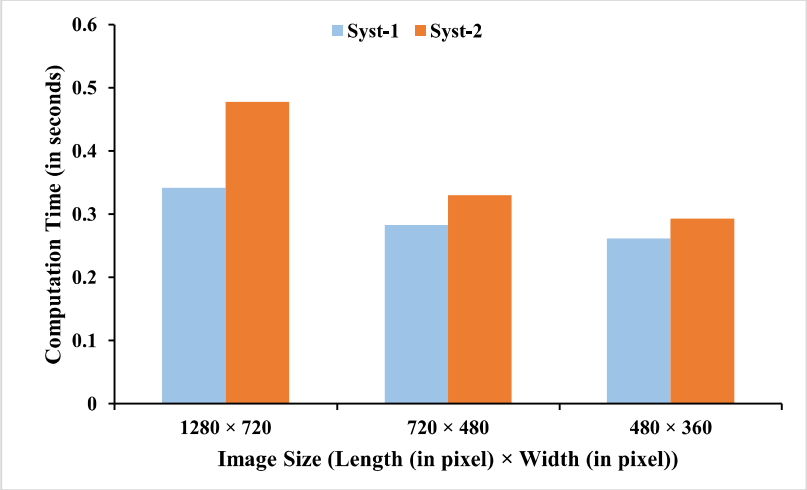


(b) IMGN67

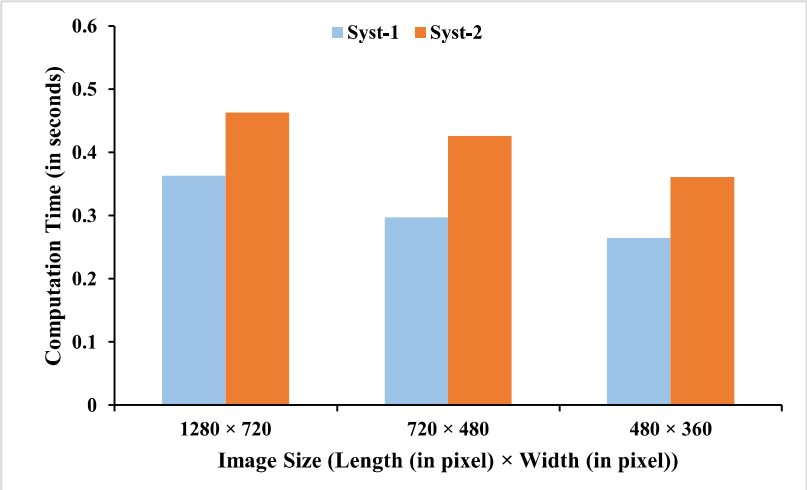


(c) IMGN68

Fig. 5.12 Computation time of sample video frames (in seconds) computed by implementing Aethra-Net



(d) IMGN69



(e) IMGN70

Fig. 5.12 (continue) Computation time of sample video frames (in seconds) computed by implementing Aethra-Net

(4) The original image has been restored with extraction of high-level and medium-level features.

(5) The original image has been restored with extraction of medium-level and low-level features.

(6) The original image has been restored with extraction of high-level and low-level features.

(7) The original image has been restored with all feature extraction.

The detailed ablations findings for IMSET-17 (i.e., IMGN63, IMGN64, and IMGN27) have been obtained. It has been observed from Figure 5.13 and Table 5.11 that all algorithms are negative for visibility restoration in dense haze except when all feature extractors are used.

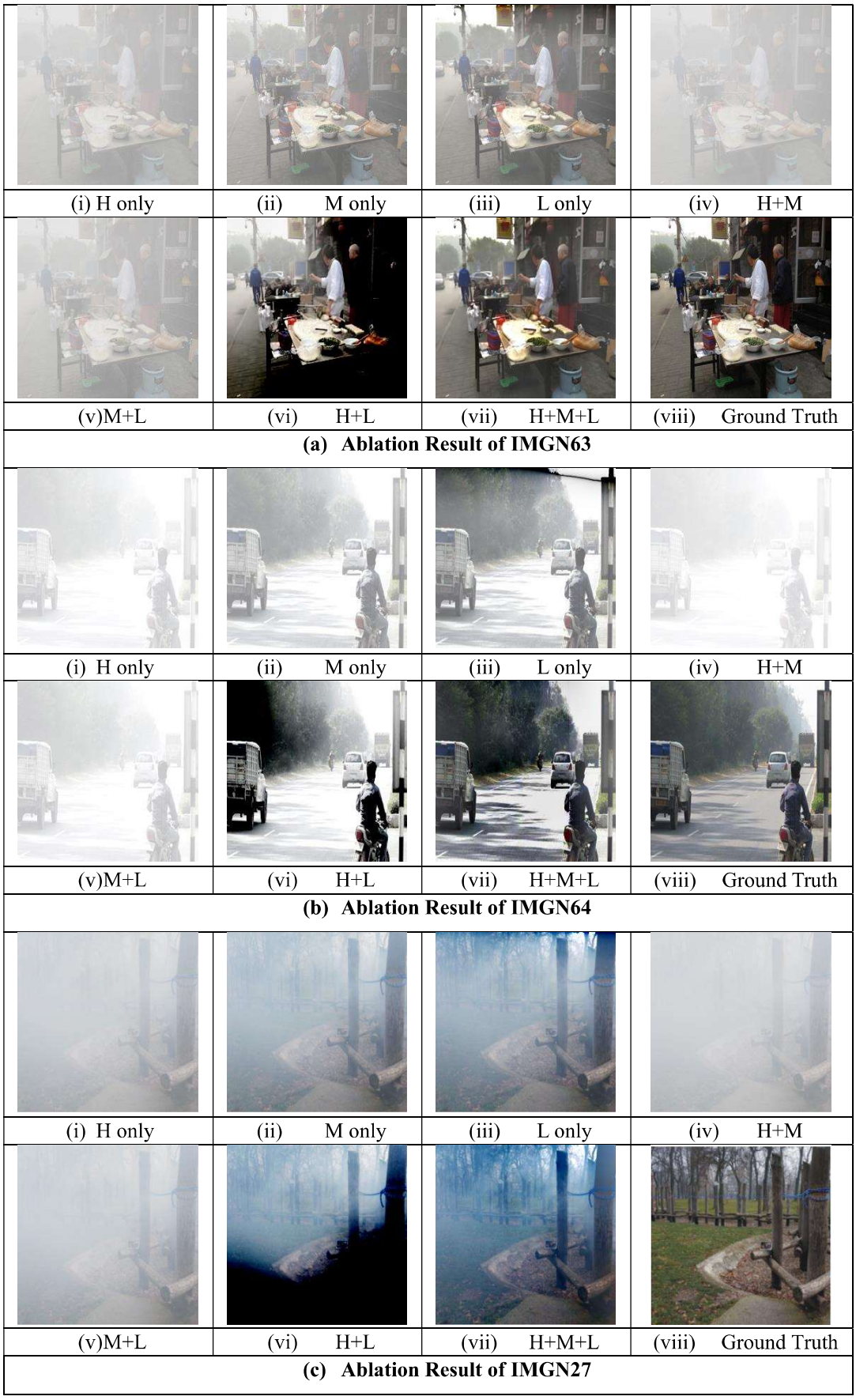


Fig. 5.13 Qualitative analysis of ablation results (H: High-Level Feature Extractor; M: Medium-Level Feature Extractor; L: Low-Level Feature Extractor)

Table 5.11 Quality metrics of ablation results of RRNet

Algorithm	Visibility Restoration in Dense Haze	Color Distortion	Naturalness
Only High-Level Features	✗	✓	✓
Only Medium-Level Features	✗	✓	✓
Only Low-Level Features	✗	✓	✓
High-Level + Medium-Level Features	✗	✓	✓
Medium-Level + Low-Level Features	✗	✗	✓
High-Level + Low-Level Features	✗	✗	✗
High-Level + Medium-Level + Low-Level Features	✓	✓	✓

Table 5.12 Quantitative analysis of ablation results of RRNet

Algorithm	*PSNR*	*BRISQUE*	*PIQE*	*IE*
Only High-Level Features	7.2016	33.4017	44.7905	5.958633
Only Medium-Level Features	8.1769	35.3114	44.129	6.729167
Only Low-Level Features	9.329567	30.0046	40.15977	7.217667
High-Level + Medium-Level Features	6.000133	34.37763	50.0539	5.3141
Medium-Level + Low-Level Features	6.575533	34.25457	44.8893	5.910433
High-Level + Low-Level Features	8.548067	37.66603	45.02867	5.768033
High-Level + Medium-Level +Low-Level Features	**11.8263**	**15.5459**	**35.39637**	**7.435233**

When only a low-level feature extractor is used, the images are partially dehazed. Furthermore, the images restored using a combination of low-level and high-level feature extractors are darker than images restored using all the feature extractors combined. Furthermore, when a low-level feature extractor is combined with only a medium-level feature extractor and a high-level feature extractor, the restored images suffer from color distortion. Furthermore, the images restored using the combination of high and low-level feature extractors do not appear natural.

As depicted in Table 5.12, the best average values of $PSNR$, $BRISQUE$, $PIQE$, and IE have been obtained for IMGN63, IMGN64, and IMGN27 as 11.8263, 15.5459, 35.39637, and 7.435233, respectively, when all the feature extractors are utilized together.

5.6 Summary

In this chapter, a gush enhancement unit has been used to predict the transmission map in a deep learning-based autoencoder. The gush enhancer utilizes three different

spatial feature maps for image refinement. A gush enhancement module has been developed to demonstrate the nature of light entering the human eye through different paths. It extracts the high-level, medium-level, and low-level features. Multiple Resnet-101 blocks have been used to overcome the vanishing gradient problem. The vessel enhancement technique has been used on geometrical objects resembling tubular structures. It improves the qualitative characteristics of the restored images, such as halo effect minimization, gradient artifacts reduction, and edge preservation. The proposed approach has been used for video dehazing. The video frames have been extracted and processed to obtain the dehazed results. The restored frames have been arranged in the correct order. The proposed approach has been found superior to existing dehazing techniques in terms of $BRISQUE$ by 15.29%. The average time of frame dehazing is 0.11 second.

CHAPTER-6
AN EFFICIENT DEHAZING FRAMEWORK FOR ROAD SAFETY

In this chapter, a fast and efficient dehazing algorithm called the Bounding Function for Gray-world Kernel Prior (BFGKP) technique has been proposed to obtain dehazed images. The atmospheric light has been obtained using an ordered filter to evenly distribute the pixel values for edge smoothing. Also, a dynamic real-time dehazing framework has been proposed that uses a GoPro camera to capture the hazy scene and a tablet to observe the results. The framework also consists of a bluerigger chord and camlink to transfer the acquired frames from camera to screen. Various visibility restoration-based challenges have been considered while implementing this framework for real-time applications.

6.1 Motivation

A novel dehazing technique called the bounding function for gray-world kernel prior (BFGKP) algorithm has been designed to minimize the distortion in dehazed images. It estimates the key properties of hazy images using a mean filter and kernel assumption technique. The edges have been preserved in the resultant images along with the reduction of flickering effects. Also, the computation time has been minimized without affecting the resolution of the image. In contrast to [176], the atmospheric light has also been updated every frame.

The majority of the existing dehazing frameworks were static and tested on pre-acquired video datasets. With the advancement in technology, it is possible to collect real-time data from cloud storage using fog computing and edge computing techniques [206], [207]. The cloud storage can only be accessed using an IP address generated for a specified task.

6.2 Proposed Dehazing Model

In this section, the transmission map and airlight have been estimated to obtain a restored image. The video frames have been segregated one at a time and processed using the proposed approach, as depicted in Figure 6.1.

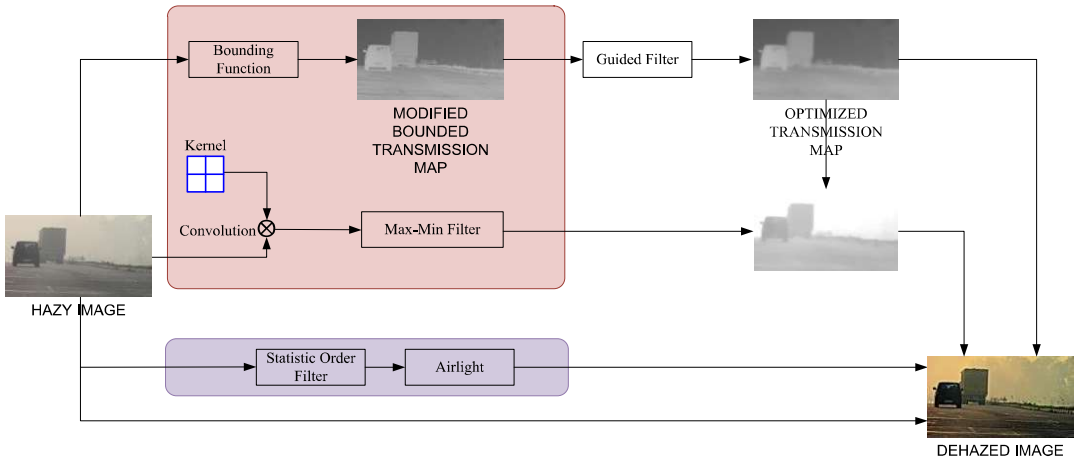


Fig. 6.1 Detailed illustration of Bounding Function for Gray-world Kernel Prior dehazing technique

6.2.1 Estimation of Airlight

In real-time hazy images, non-homogenous haze is present in three color channels (Red, Green, and Blue). Each channel has been represented in the form of 2D matrix. The atmospheric light has been estimated using a 2D weighted order filter (WOF) such that the color channel with the least non-zero pixel intensity has been selected.

6.2.2 Initialization of Bounding Function

The map has been produced using hazy image and estimated airlight. In the hazy image, the color channel with minimum intensity ($I_{minch}(x,y)$) has been selected. The maximum value ($I_{minch_{max}}(x,y)$) of selected color channel has been found. Also, the mean value of airlight estimated for different color channels has been calculated as follows [133]:

$$A_{mean} = \left(\frac{A_{ch_R}+A_{ch_G}+A_{ch_B}}{3}\right) \tag{6.1}$$

where A_{ch_R}, A_{ch_G}, and A_{ch_B} are the estimated airlight of color channels (R, G, and B), respectively.

The upper and lower boundaries of depth map to be estimated have been obtained by calculating the difference of pixel intensities of $I_{minch}(x,y)$ and A_{mean}. The bounded transmission map ($t_{bounded}(x,y)$) has been estimated using a lower bound function as follows [133]:

$$t_{bounded}(x,y) \geq \frac{1}{1+\frac{I_{minch_{max}}(x,y)\times 10^{-0.05\times\varsigma}}{A_{mean}-I_{minch}(x,y)}} \tag{6.2}$$

where ς is the bounding function given as:

$$\varsigma = \frac{\chi}{\sqrt{I_{minch}(x,y)}} \tag{6.3}$$

where χ is a quality control constant used to reduce the edge distortion present due to non-linearity in the computed image. If the value of χ is low, noise is introduced in the dense hazy regions of the reconstructed images.

The minimum pixel intensity of $J(x,y)$ is higher than $I(x,y)$ due to additive airlight. Also, depth of $I(x,y)$ is lower due to transmission map. Thus, constraints have been set for $I(x,y)$ using airlight. The modified bounded transmission map has been obtained using Eq. (6.4).

$$t_{bounded_{modified}}(x,y) = |t_{bounded}(x,y) \times lb(x,y)| + \frac{|ub(x,y)\times t_{bounded}(x,y)|}{max|ub(x,y)\times t_{bounded}(x,y)|} \tag{6.4}$$

where $t_{bounded_{modified}}$ is the modified bounded transmission map. ub and lb are upper and lower bounds of the estimated transmission map, respectively.

$$ub(x,y) = 1, lb(x,y) = A_{mean}, \quad if\ I_{minch}(x,y) \geq A_{mean} \tag{6.5a}$$

$$lb(x,y) = 1, lb(x,y) = A_{mean} \quad if\ I_{minch}(x,y) < A_{mean} \tag{6.5b}$$

The optimized transmission map ($t(x,y)$) has been obtained by using guided filter on modified map ($t_{bounded_{modified}}$). It has been used to improve texture and reduce edge distortion.

6.2.3 Gray-world Kernel Prior

The features have been automatically extracted using the Gray-world assumption (GWA) to generate a saliency map [208]. The concept of saliency describes the visual attention needed to evaluate the contrast of an item present in a hazy image as compared to its surroundings. The saliency map depicts the importance of degree of haze in different regions of the image. It highlights the objects of greater significance that are ignored in a clutter.

In this method, convolution of each channel of modified bounded map and a window k of size 75×75 has been performed. The maximum intensity for the color channel with the fewest pixel values in the convolved output has been determined using the max-min filter as depicted in Eq. (6.6).

$$t1_{max} = \max\left(\min\left(I(x,y) * k\right)\right) \tag{6.6}$$

where $*$ is the convolution operation.

6.2.4 Scene Enhancement Map

To enhance the drivers' visibility, a scene enhancement map (SEM) has been calculated for the reconstructed video. As the depth increases, light absorption becomes stronger. The light absorption coefficient ($\mathcal{L}(x,y)$) is given below [139]:

$$\mathcal{L}(x,y) = 1 - \sigma \frac{d(x,y)}{\max\left(d(x,y)\right)} \tag{6.7}$$

$\sigma = 0.15$ is an optimization coefficient.

$d(x,y)$ can be calculated by rewriting Eq. (1.2) as:

$$d(x,y) = -\frac{\ln\left(t(x,y)\right)}{\beta} \tag{6.8}$$

By substituting Eq. (6.8) in Eq. (6.7), $\mathcal{L}(x,y)$ can be rewritten as:

$$\mathcal{L}(x,y) = 1 - \sigma \frac{\ln\left(t(x,y)\right)}{\max\left(\ln\left(t(x,y)\right)\right)} \tag{6.9}$$

Furthermore, the airlight has been considered to improve the exposure of light. Thus, scene recovery factor (SEM) has been given as:

$$SEM = A \times \mathcal{L}(x,y) \tag{6.10}$$

In Eq. (6.9), to ease the calculation, the logarithmic functions have been replaced by a rational function and $\max\left(\ln\left(t(x,y)\right)\right)$ has been replaced by $t1_{max}$. Therefore, Eq. (6.10) can be rewritten as:

$$SEM = A \times \left(1 - \sigma \frac{\frac{1}{p_2 + t(x,y)}}{\frac{1}{p_2 + \max\left(\ln\left(t(x,y)\right)\right)}}\right) = A \times \left(1 - \sigma \frac{p_2 + t1_{max}}{p_2 + t(x,y)}\right) \tag{6.11}$$

where $p_1 = -0.3$ and $p_2 = 0.2$ are optimization factors.

6.2.5 Haze Restoration Model

The components estimated by implementing the proposed approach has been used for visibility restoration. The restored image has been obtained as follows:

$$J(x,y) = \frac{I(x,y) - A(1 - t(x,y))}{t(x,y) \times \text{SEM}} \tag{6.12}$$

The dehazed frames have been generated on screen in the same sequence in which they are acquired and processed.

6.3 Qualitative Analysis

The intermediate results of IMSET-18 (i.e., IMGN71 to IMGN73), namely transmission map, optimized map, and scene enhancement map obtained by implementing BFGKP have been presented in Figure 6.2. It has been observed that the transmission maps use depth maps to depict the salient features of the hazy images. These maps have been optimized using airlight. It has been concluded that the intensity of pixels in the optimized maps is proportional to the thickness of haze in the original images, i.e., dark regions in optimized maps indicate high density fog in the input images.

Figures 6.3 – 6.5 depict the visual comparison of BFGKP with the existing dehazing techniques. In Figure 6.3, the dehazed images for datasets IMSET-19 (i.e., IMGN74 to IMGN79) using DCP, NLBF, and CCDI are cloudy and have low intensity. IDE enhances the illumination of the reconstructed images. CAP, ICAP, and OTM-AAL generate partially dehazed images. The illumination of dehazed results obtained by implementing BFGKP is superior to the existing dehazing techniques.

The existing techniques and proposed dehazing technique have also been used to obtain the dehazed images for the IMSET-20 dataset (i.e., IMGN80 to IMGN94), as depicted in Figure 6.4. The sky regions of the dehazed images of IMGN82, IMGN83, and IMGN84 are partially dehazed for CAP, ICAP, OTM-AAL, and IDE. It has been observed that the results computed using BFGKP are sharper and brighter as compared to DCP, NLBF, CDP, and CCDI. Figure 6.5 shows the visual performance of dehazed results with low, medium, and dense haze from IMSET-21 for various dehazing techniques. The results computed for evening and dense hazy images using the existing algorithms suffer from color deviation and do not produce effective results.

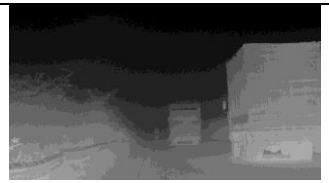
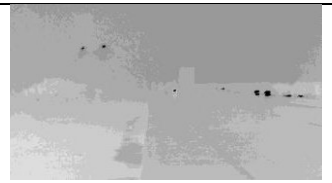
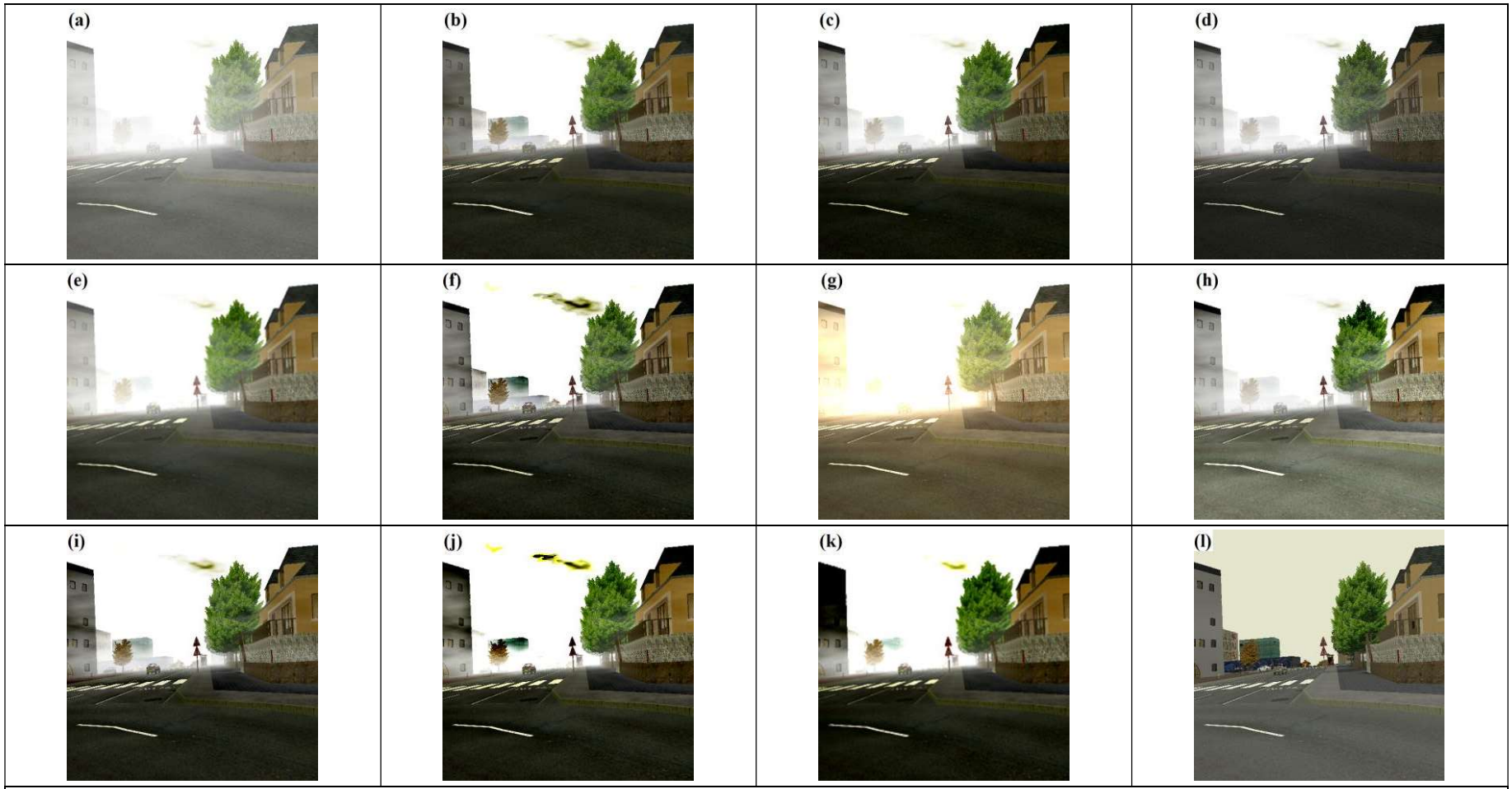
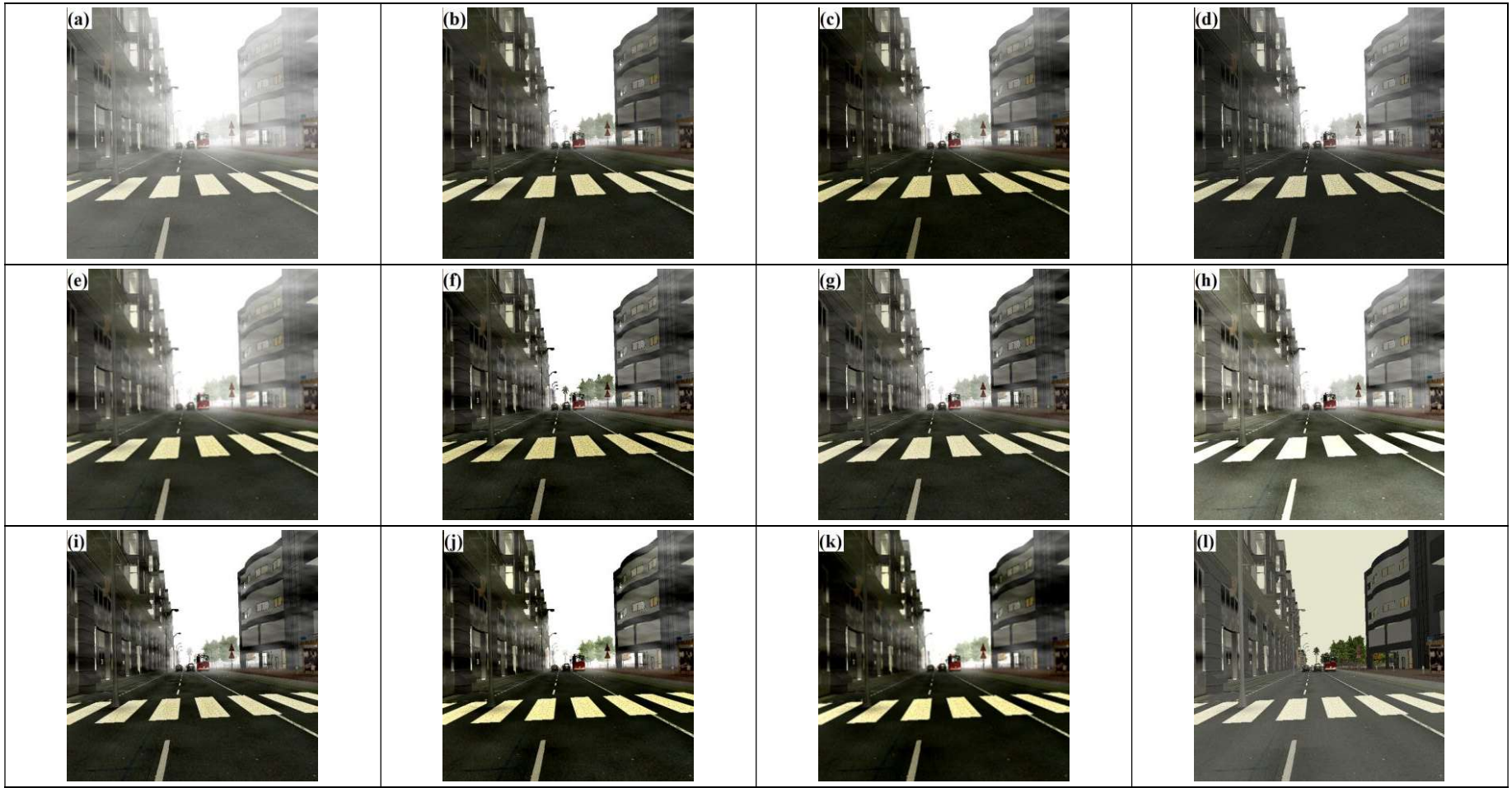
Datasets	IMGN71	IMGN72	IMGN73
Hazy Images			
Transmission Maps			
Optimized Maps			
Scene Enhancement Map			
Dehazed Images			

Fig. 6.2 Intermediate results of BFGKP



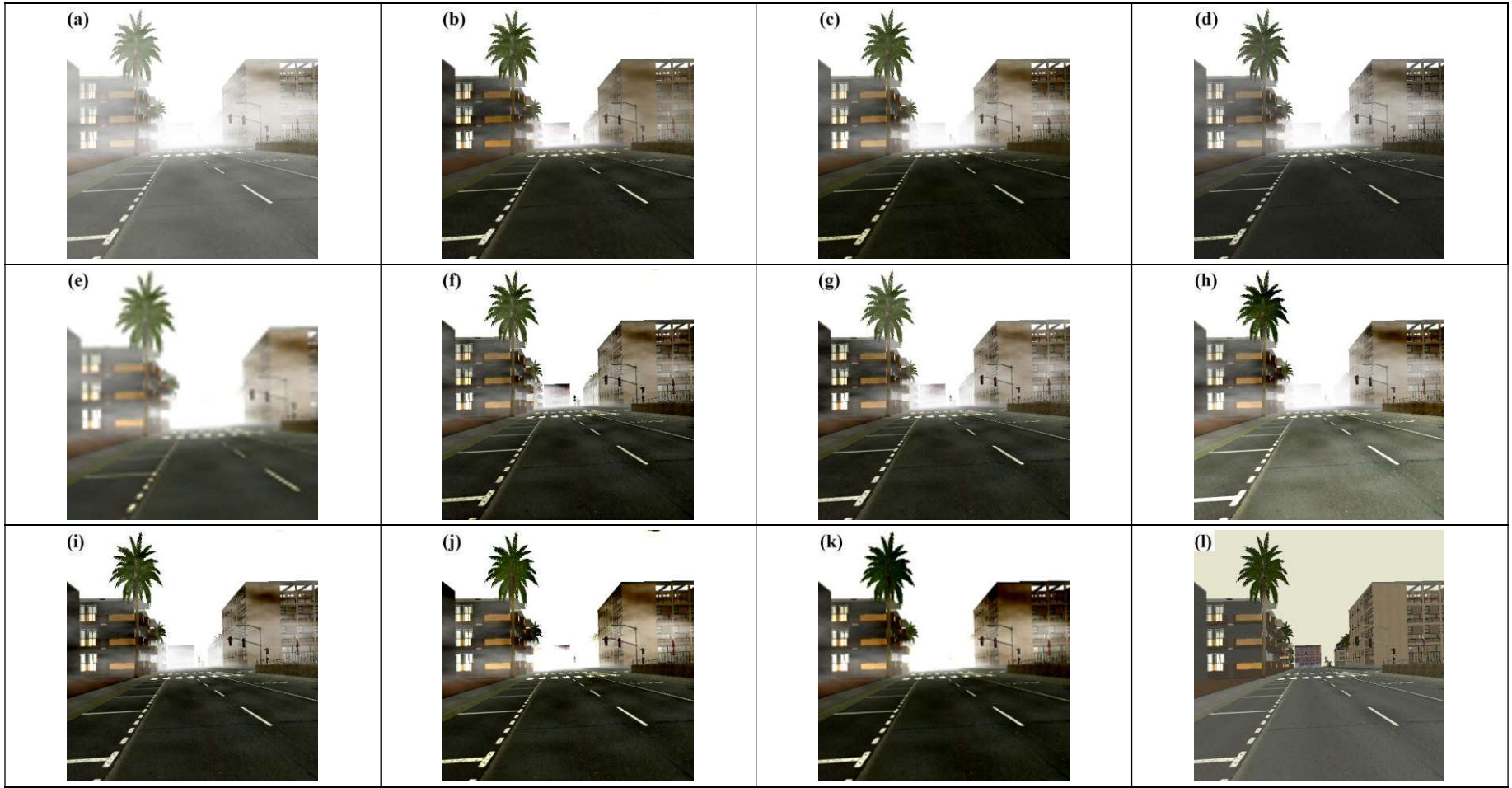
(I) IMGN74

Fig. 6.3 Comparison of images from IMSET-19 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth.



(II) IMGN75

Fig. 6.3 (continue) Comparison of images from IMSET-19 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth



(III) IMGN76

Fig. 6.3 (continue) Comparison of images from IMSET-19 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth

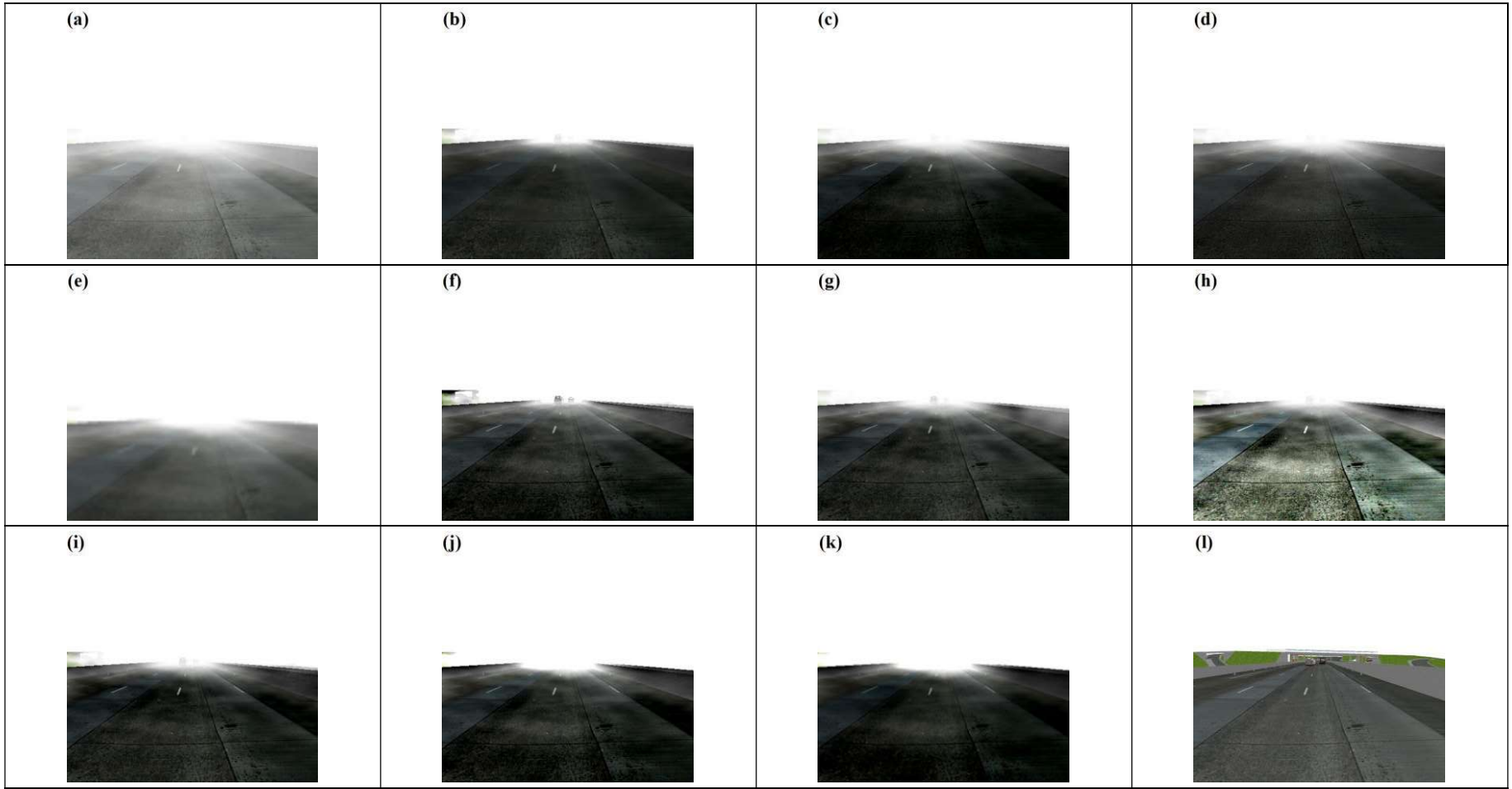


Fig. 6.3 (continue) Comparison of images from IMSET-19 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth

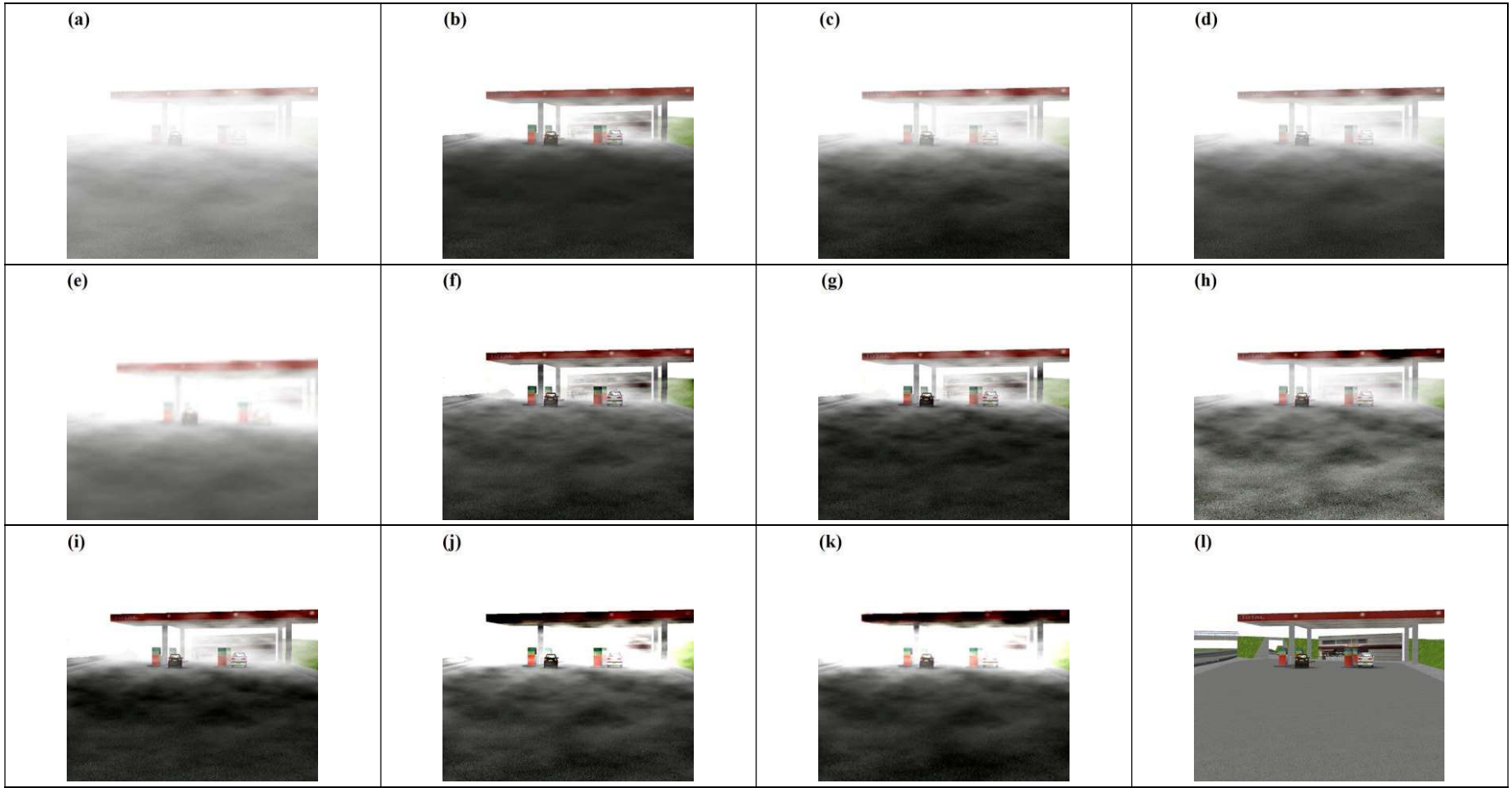


Fig. 6.3 (continue) Comparison of images from IMSET-19 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth

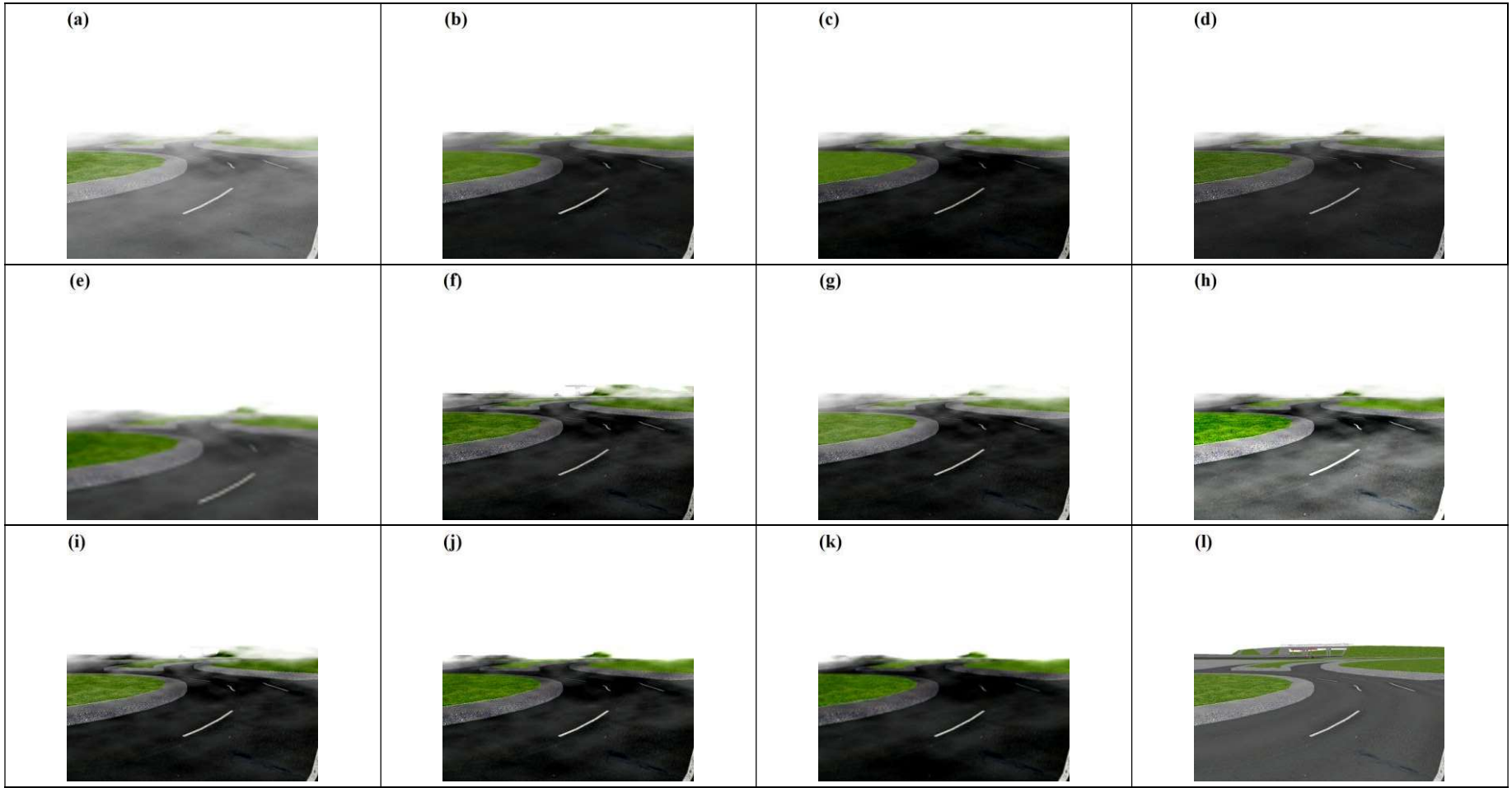
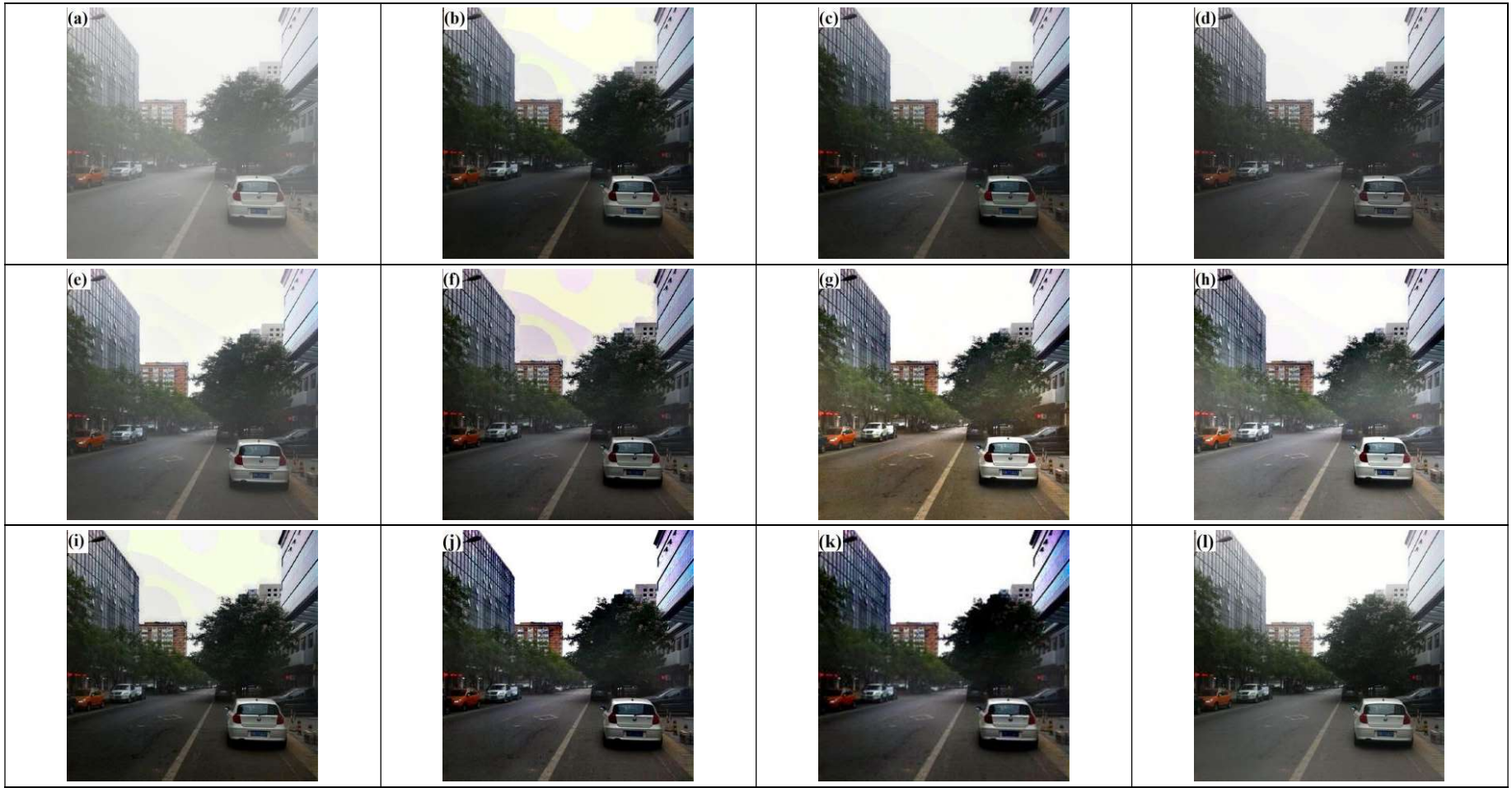
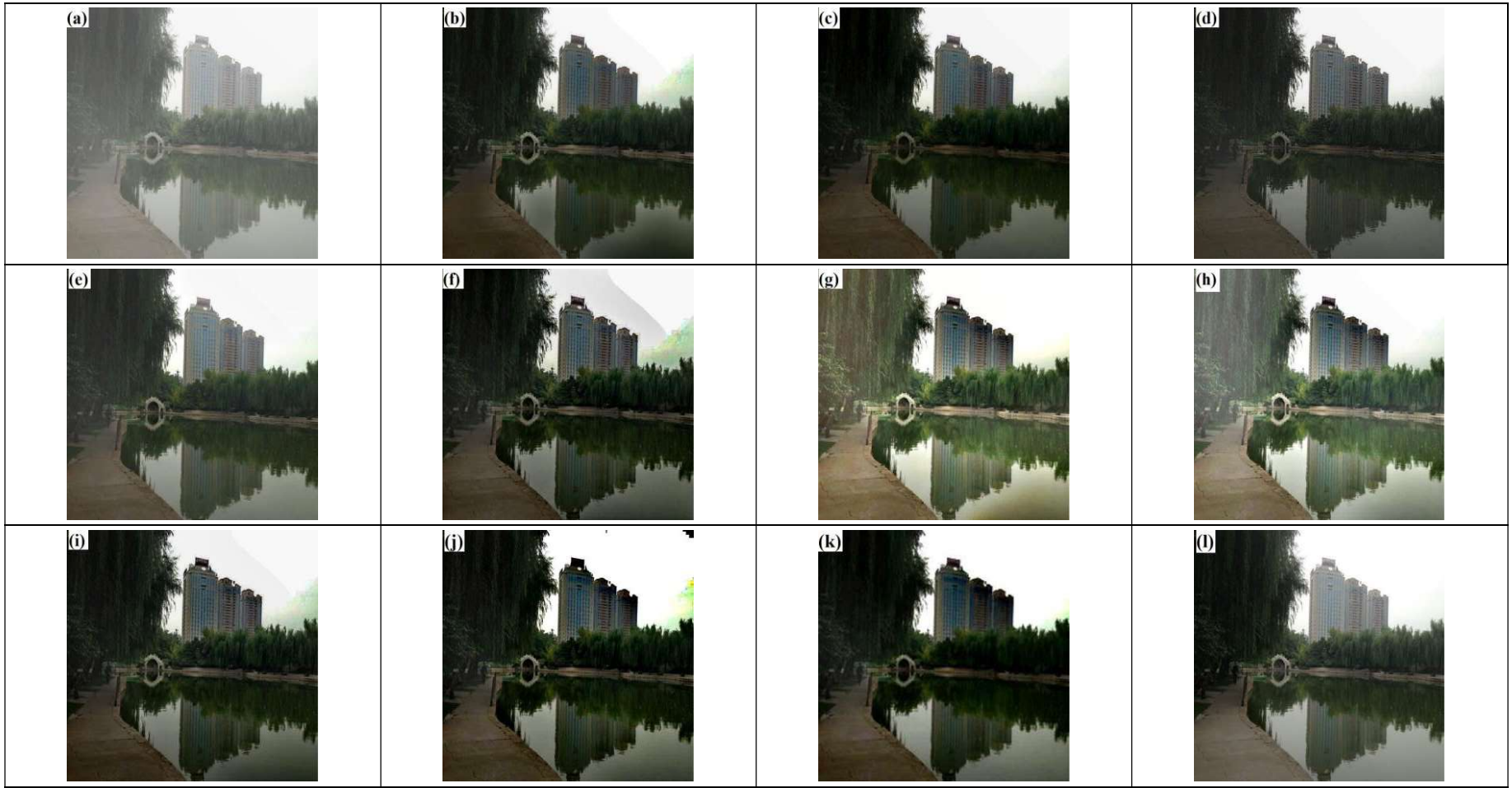


Fig. 6.3 (continue) Comparison of images from IMSET-19 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth



(I) IMGN80

Fig. 6.4 Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth

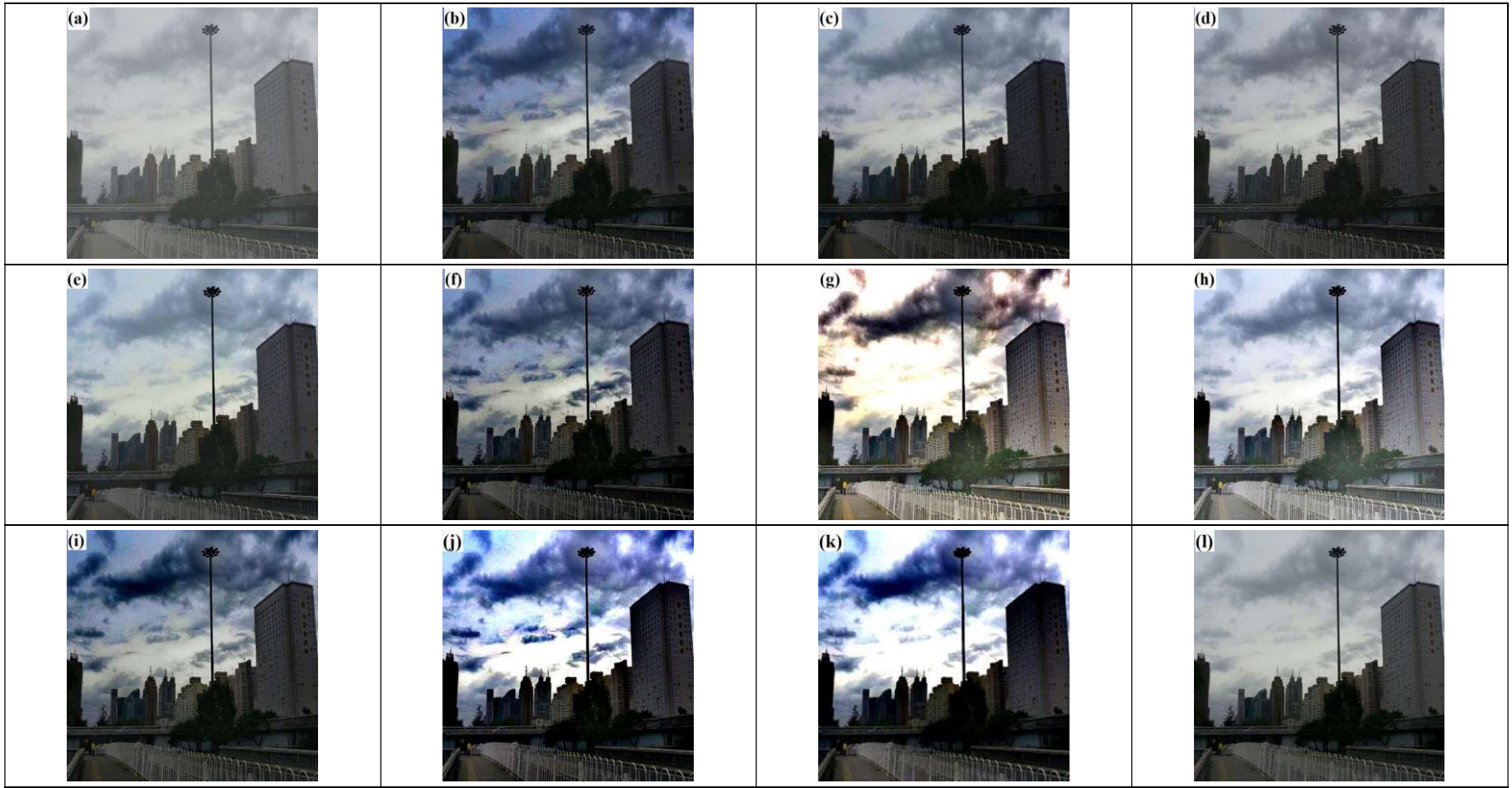


(II) IMGN81

Fig. 6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth

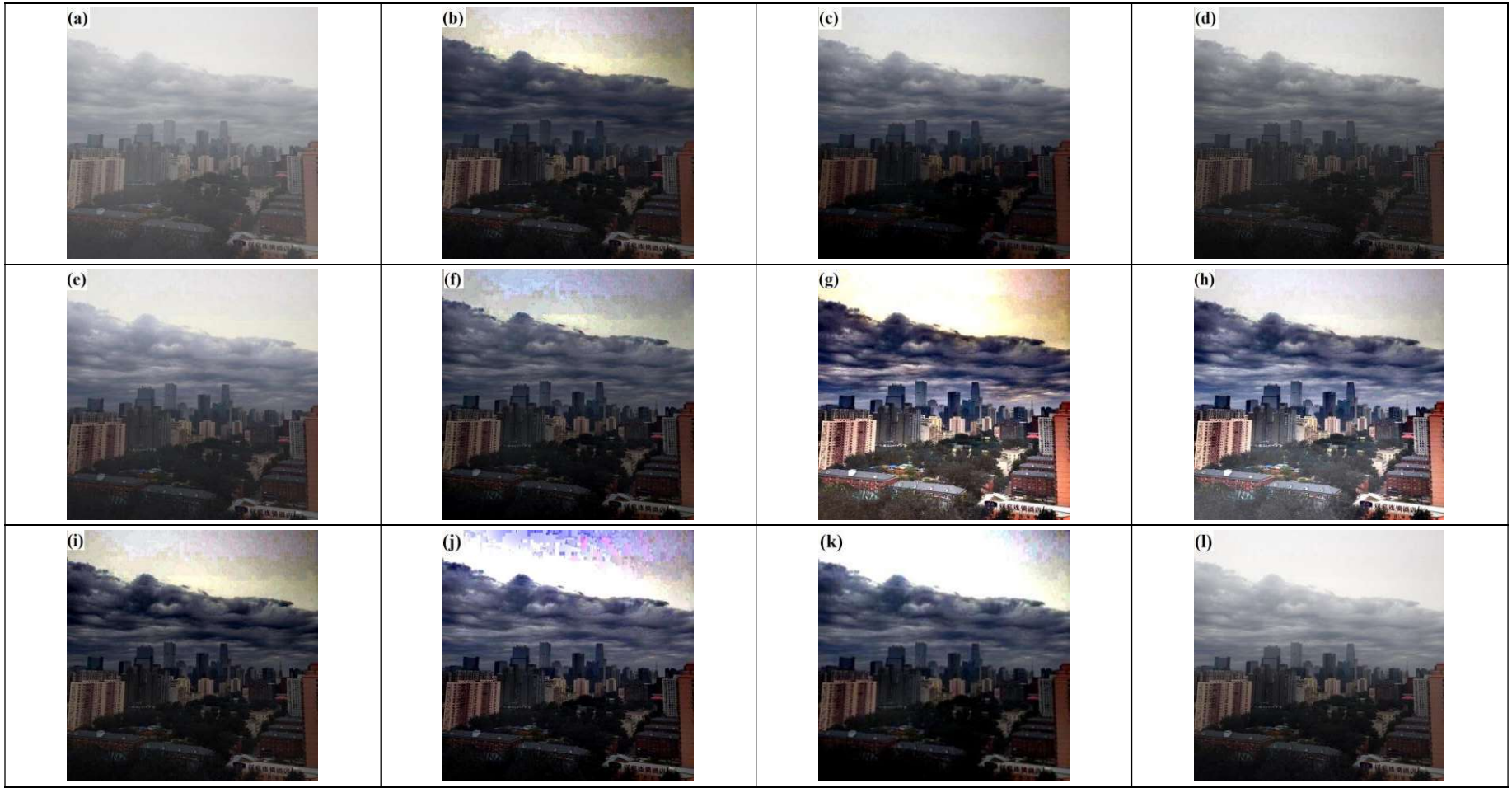
(III) IMGN82

Fig. 6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth



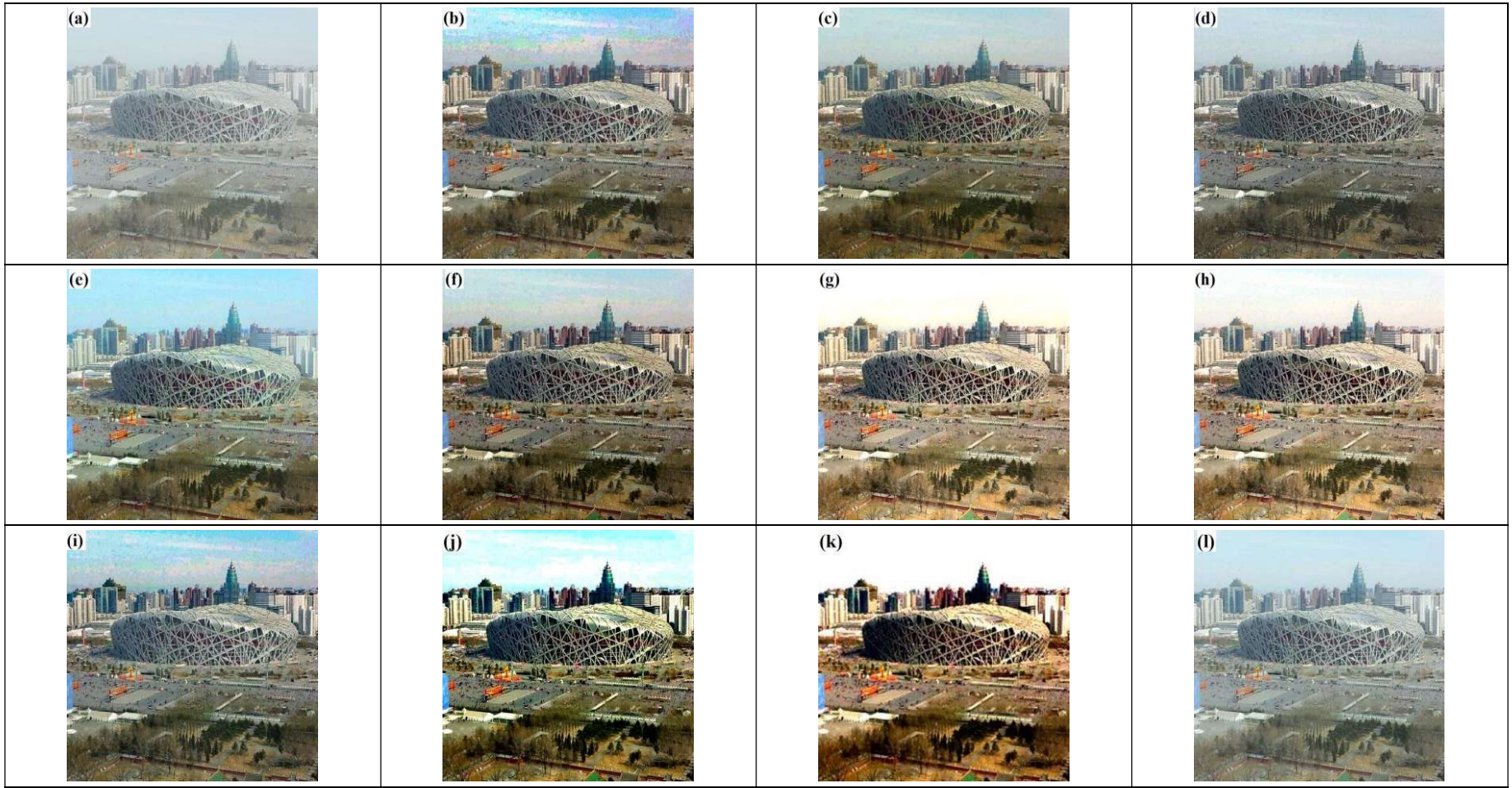
(IV) IMGN83

Fig. 6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth



(V) IMGN84

Fig.6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth

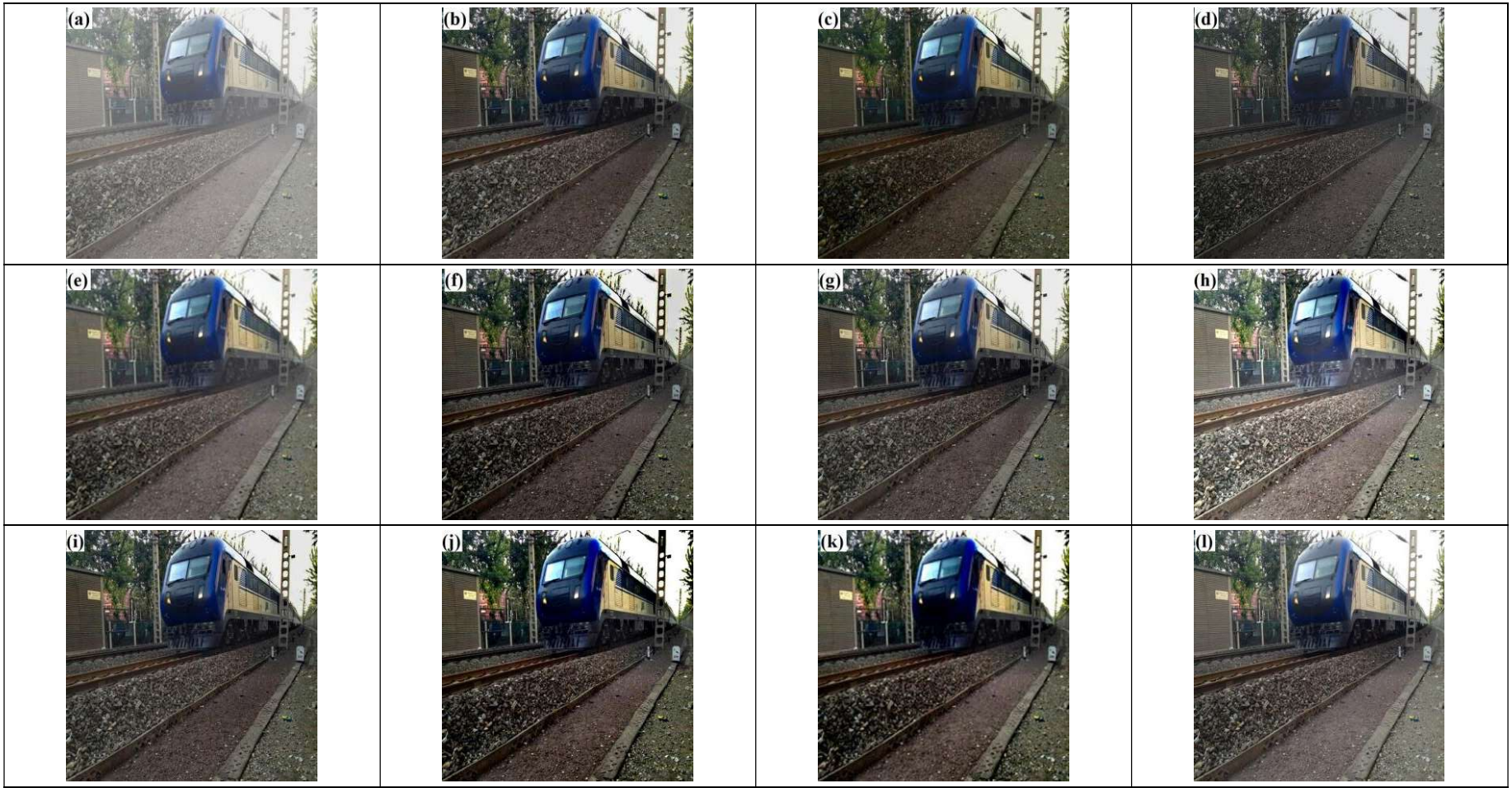


(VI) IMGN85

Fig. 6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth

(VII) IMGN86

Fig. 6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth



(VIII) IMGN87

Fig. 6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth

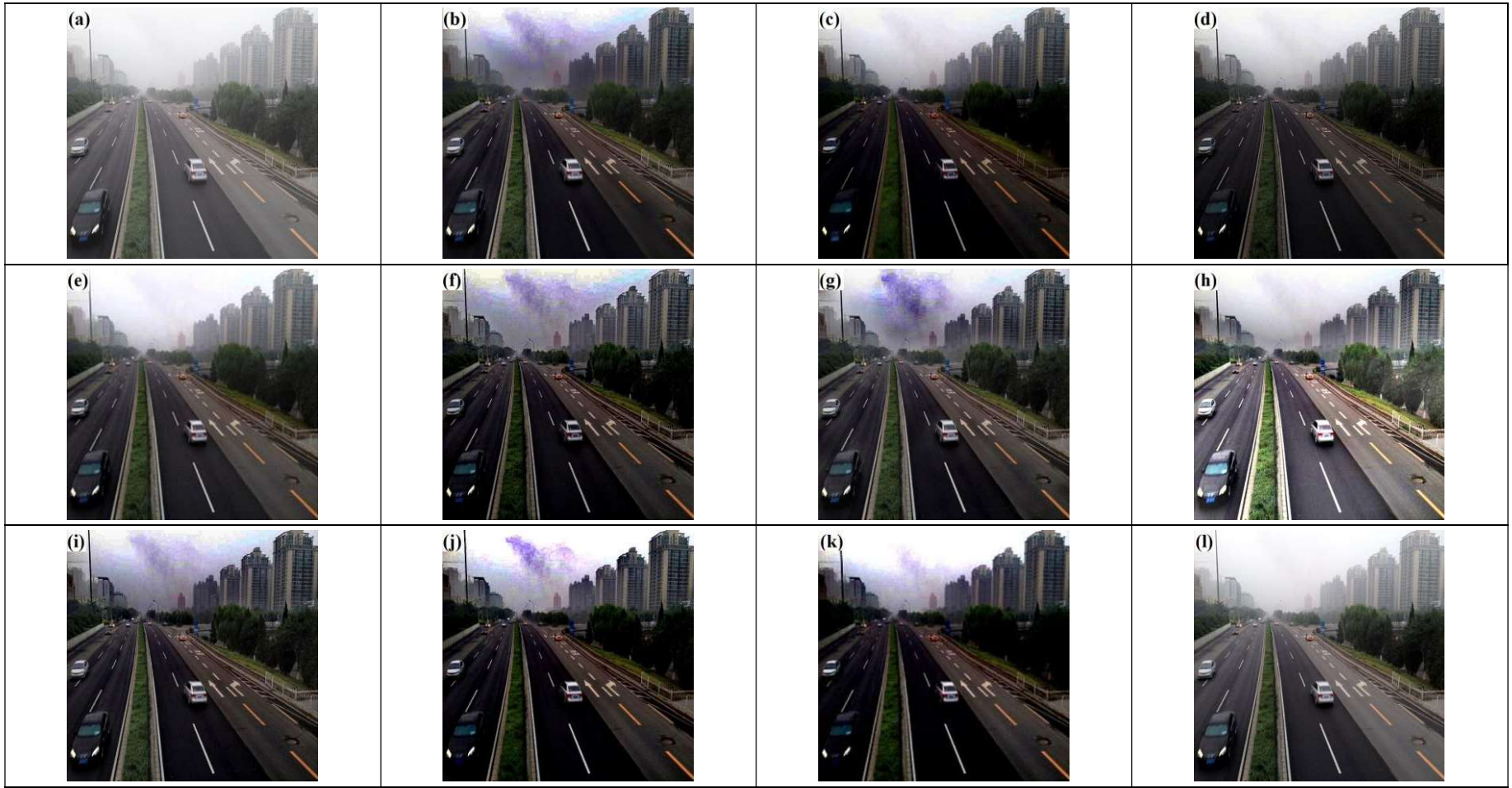
(IX) IMGN88

Fig. 6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth

Fig. 6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth

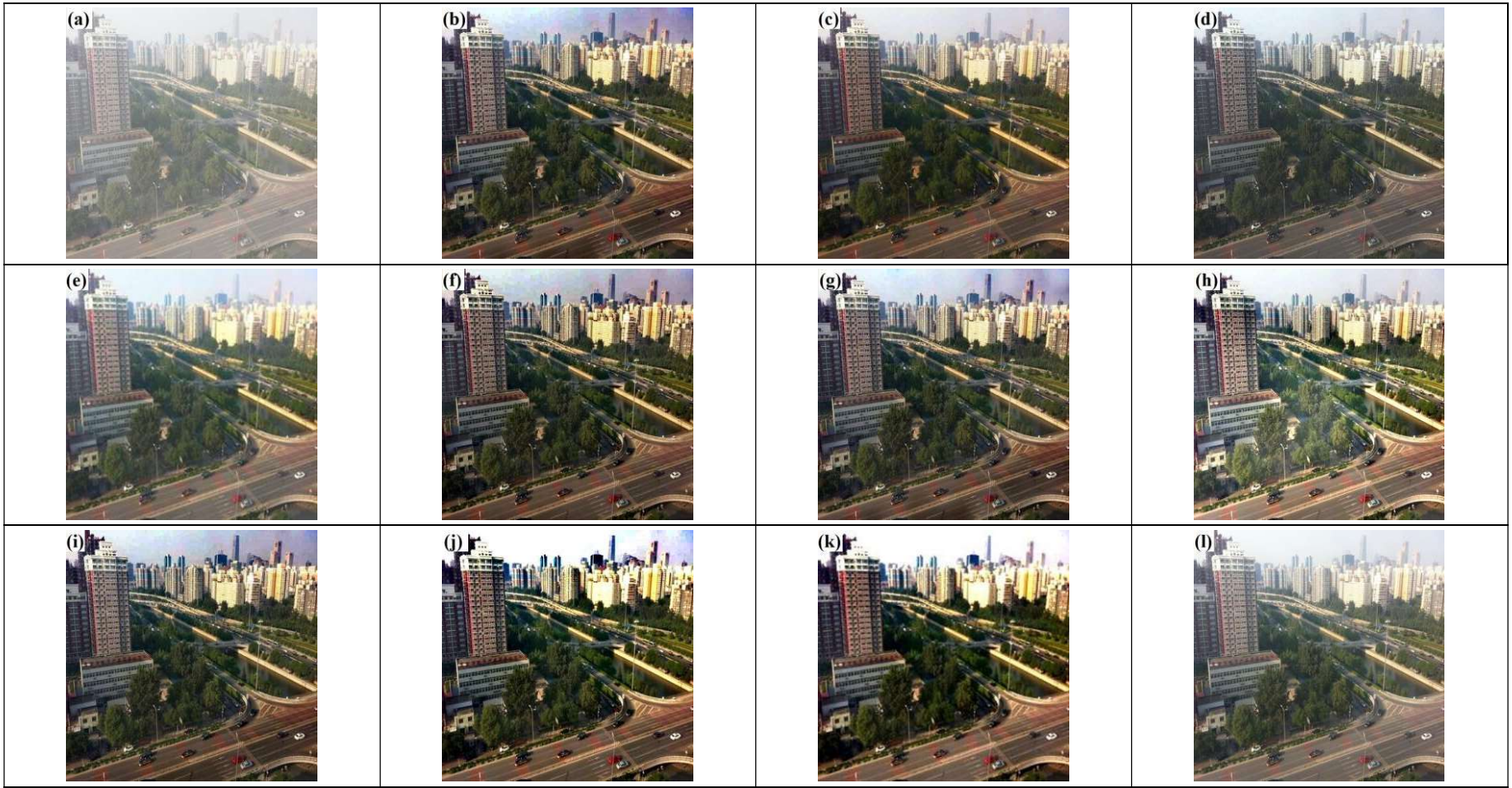
(XI) IMGN90

Fig. 6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth



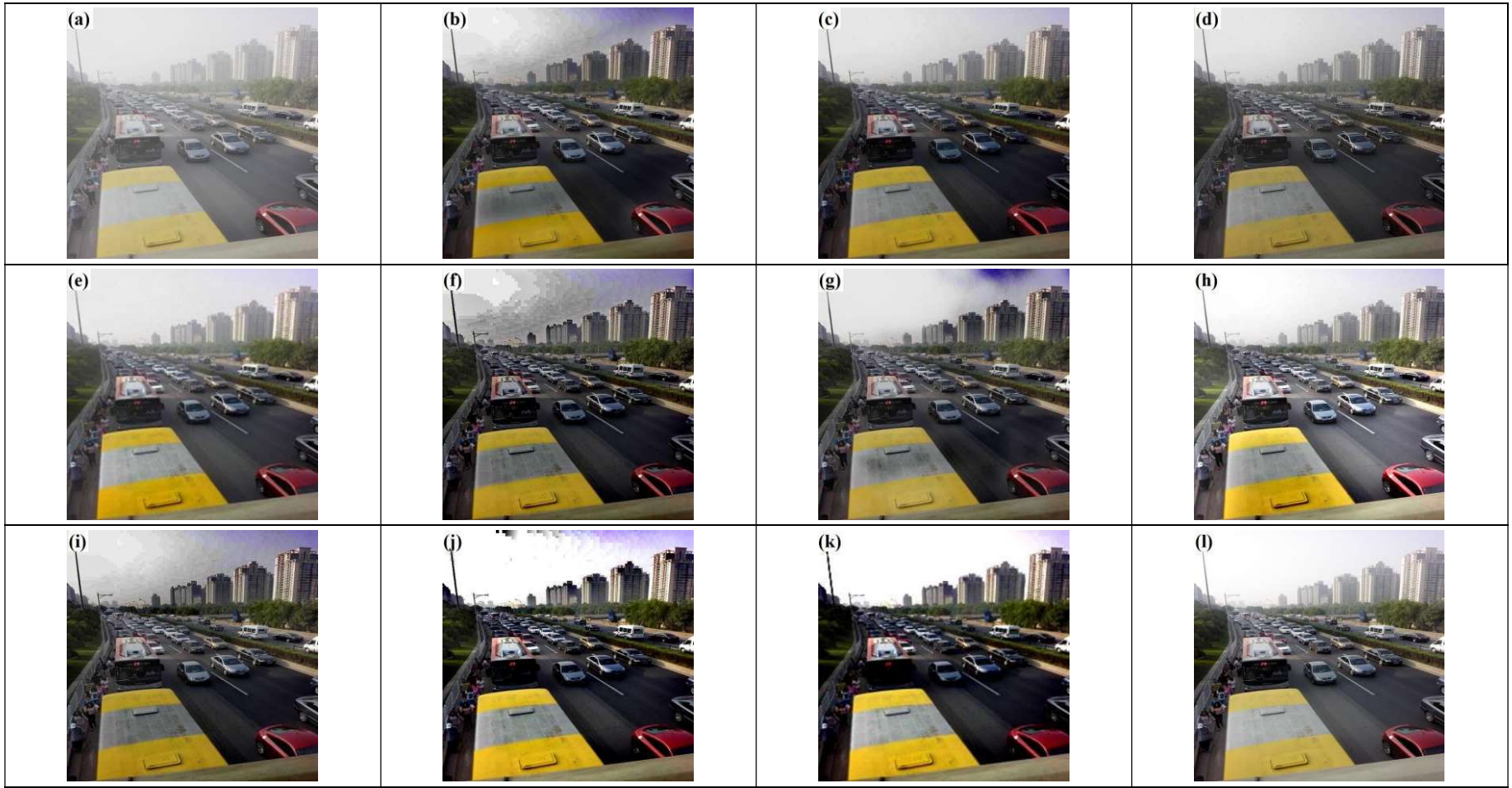
(XII) IMGN91

Fig. 6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth



(XIII) IMGN92

Fig. 6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth



(XIV) IMGN93

Fig. 6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth

(XV) IMGN94

Fig. 6.4 (continue) Comparison of images from IMSET-20 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), (k) BFGKP (256×256), and (l) Ground truth

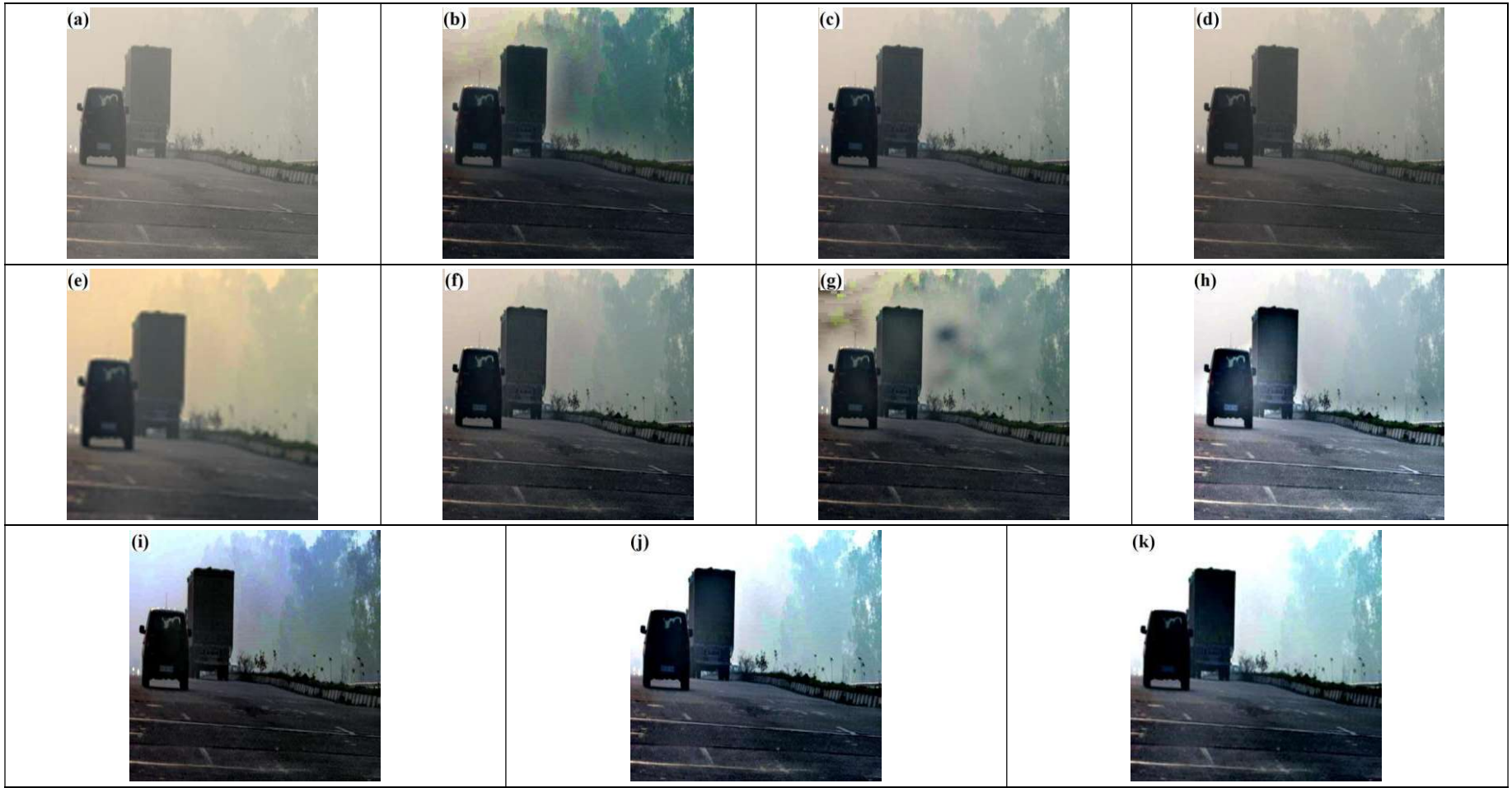
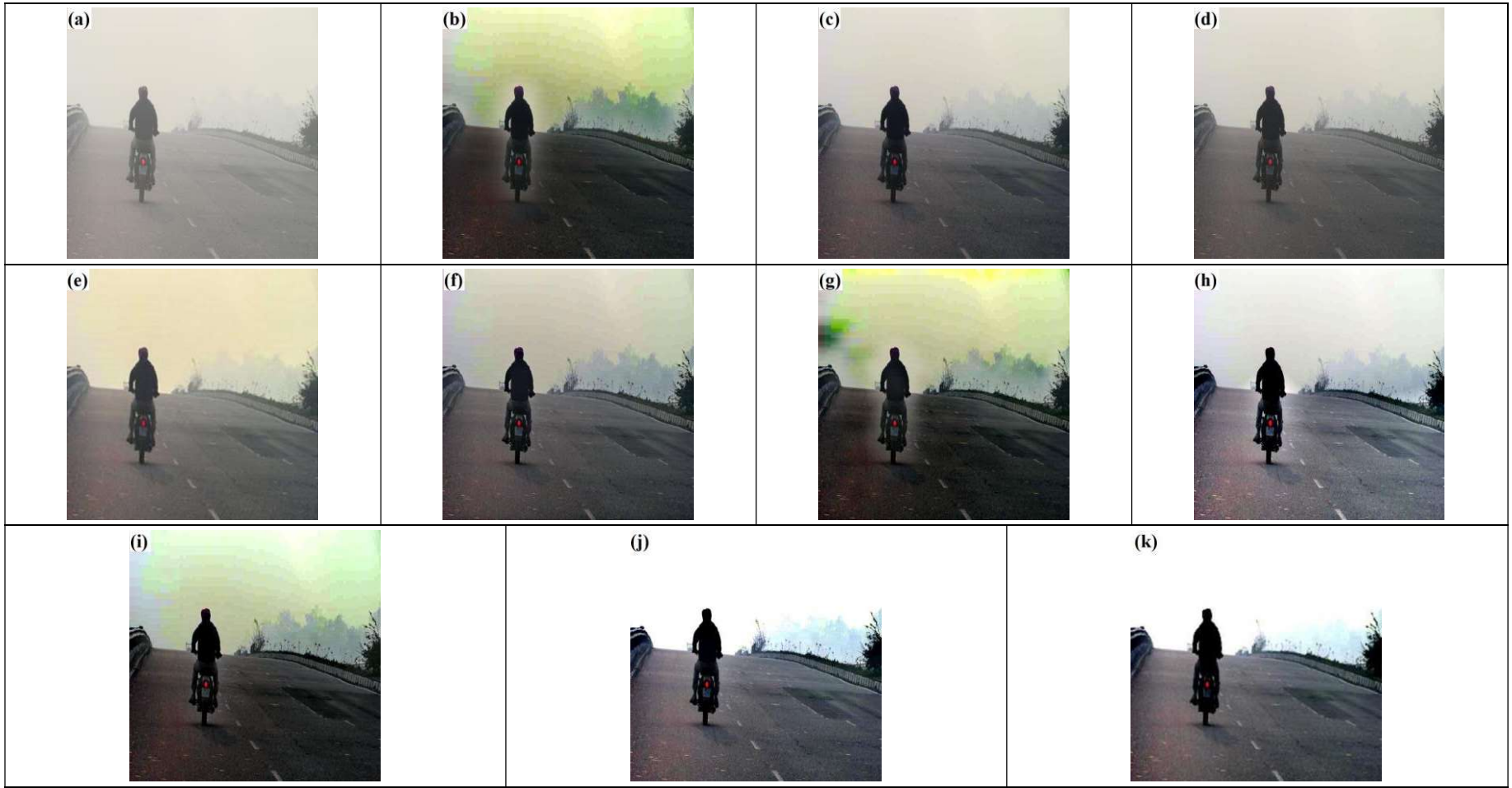


Fig. 6.5 Comparison of images from IMSET-21 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), and (k) BFGKP (256×256)

(II) IMGN95

Fig. 6.5 (continue) Comparison of images from IMSET-21 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), and (k) BFGKP (256×256)

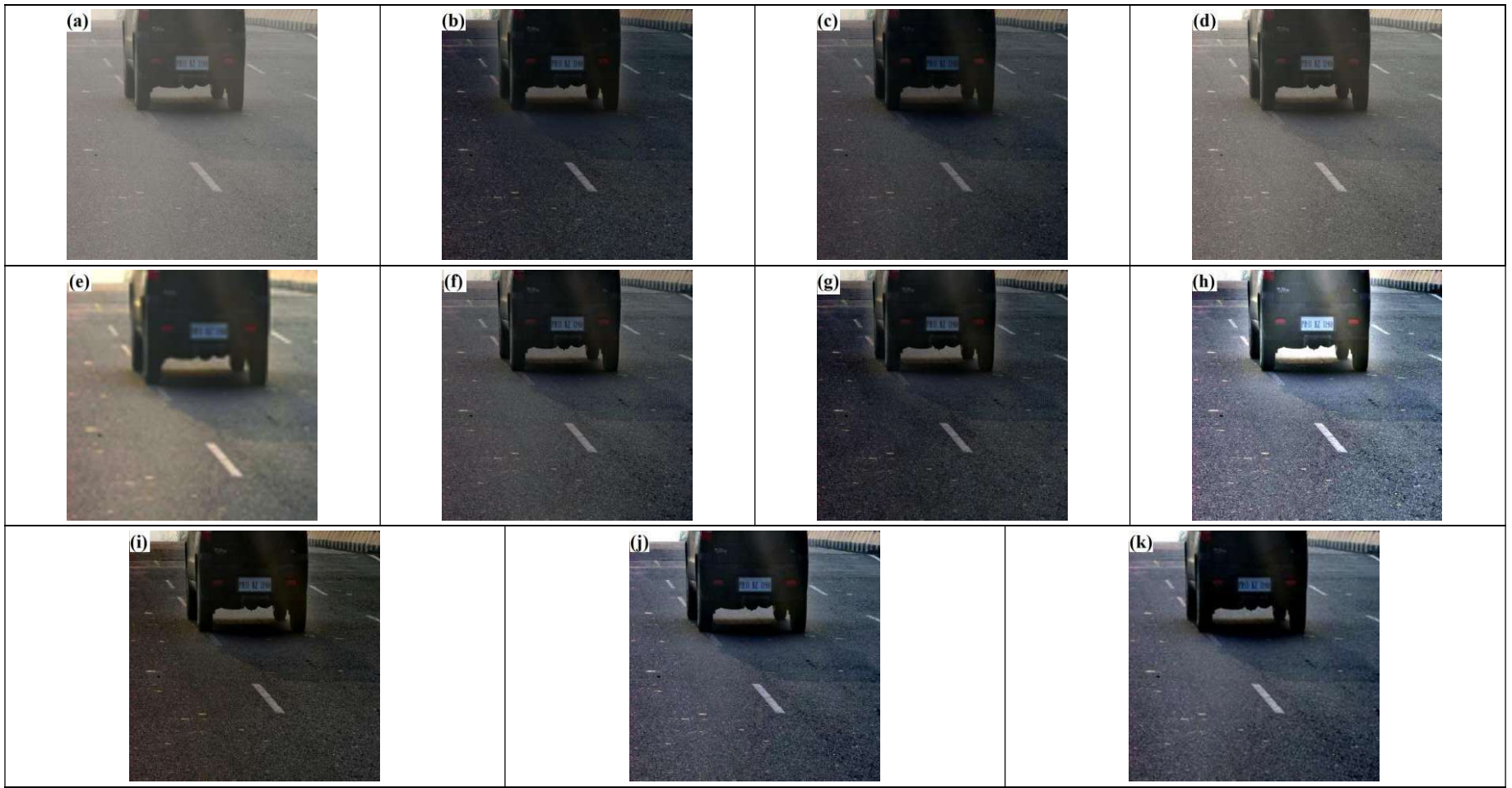


(III) IMGN96

Fig. 6.5 (continue) Comparison of images from IMSET-21 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), and (k) BFGKP (256×256)

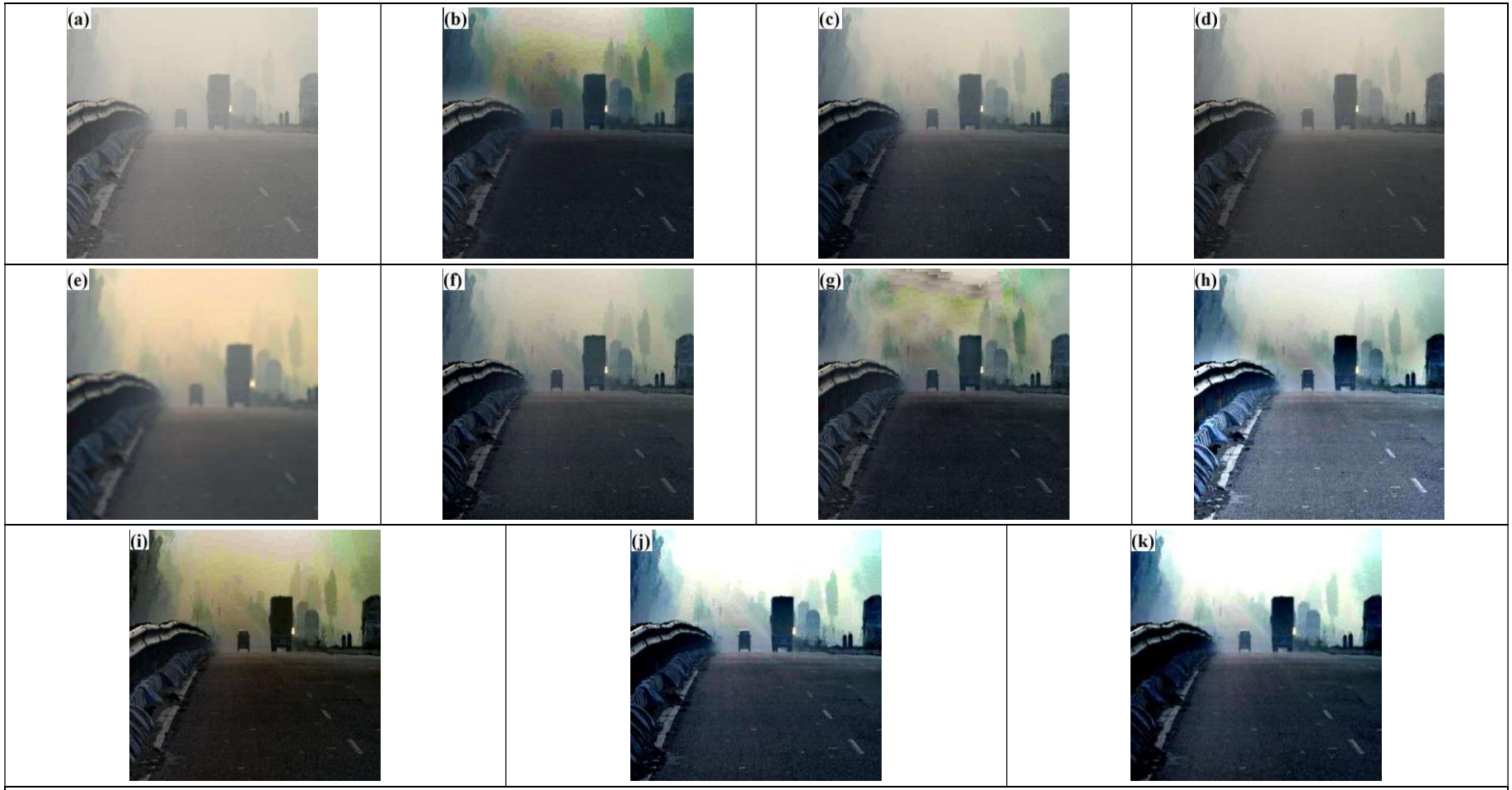
(IV) IMGN97

Fig. 6.5 (continue) Comparison of images from IMSET-21 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), and (k) BFGKP (256×256)



(V) IMGN98

Fig. 6.5 (continue) Comparison of images from IMSET-21 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), and (k) BFGKP (256×256)



(VI) IMGN99

Fig. 6.5 (continue) Comparison of images from IMSET-21 computed by various algorithms (a) Hazy image, (b) DCP, (c) CAP, (d) ICAP, (e) OTM-AAL, (f) NLBF, (g) CDP, (h) IDE, (i) CCDI, (j) BFGKP (640×640), and (k) BFGKP (256×256)

However, the images of IMSET-21 reconstructed using BFGKP suffer from color distortion in sky regions and clear visibility in non-sky regions. Overall, the results produced by BFGKP are often balanced in terms of haze density and illumination as compared to other visibility restoration techniques.

The proposed algorithm has also been tested on roadside videos (IMSET-22) obtained from internet browsing. Figure 6.6 depicts hazy frames extracted from these videos along with the resultant images produced by implementing BFGKP. The resultant images suffer from color distortion in sky regions.

6.4 Quantitative Analysis

BFGKP has been analyzed through various performance metrics, namely the naturalness image quality evaluator ($NIQE$), blind/referenceless image spatial quality evaluator ($BRISQUE$), and fog aware density evaluator ($FADE$). These performance metrics have been calculated for both synthetic and real-world images from various datasets using BFGKP and compared with other visibility restoration techniques. $NIQE$ has been used to obtain the deviation of the features derived from the hazy images for

Fig. 6.6 Video dehazing of IMSET-22 using BFGKP technique

the natural scene statistics. Table 6.1 depicts the $NIQE$ calculated using the proposed and existing algorithms (best and second best performances are indicated by '*' and '#', respectively). It has been observed that the BFGKP outperforms the existing techniques when $NIQE$ for IMGN75, IMGN82, IMGN85, IMGN86, IMGN87, IMGN92, IMGN71, and IMGN99 is estimated as 3.15, 2.62, 3.07, 2.46, 2.93, 3.02, 3.86, and 3.43, respectively. Also, the best results for IMGN74, IMGN76, IMGN81, IMGN83, IMGN88, IMGN95 are obtained using DCP, which are 2.92, 2.96, 2.83, 3.28, 2.96, and 3.66, respectively. BFGKP follows them with values of 2.93, 3.01, 3.01, 3.48, 3.16, and 3.77, respectively. Thus, the proposed approach computes the best $NIQE$ results for IMSET-19 and IMSET-20. Furthermore, the best average value of 3.632 is obtained for $NIQE$ when the DCP is implemented. It is followed by BFGKP and NLBF with value of 4.021 and 4.104, respectively.

$BRISQUE$ does not evaluate block and blur when evaluating image quality. As a result, it has lower computational complexity than $NIQE$. Thus, it is used in real-time applications. Table 6.2 depicts the $BRISQUE$ calculated using the proposed and existing algorithms (best and second best performances are indicated by '*' and '#', respectively). It has been observed that the BFGKP outperforms the existing techniques when $BRISQUE$ is calculated for IMGN78, IMGN77, IMGN92, and IMGN94 as 43.4979, 44.6355, 23.2286, and 12.4897, respectively. The best performances for IMGN75, IMGN96, and IMGN97 are obtained using OTM-AAL algorithm (i.e., 29.4355, 34.5007, and 35.1677, respectively). It is followed by BFGKP, which has a value of 31.822, 35.4655, and 35.3689, respectively. Also, the best performances for IMGN87 and IMGN98 are obtained using NLBF (i.e., 30.2194 and 17.0905, respectively). It is followed by BFGKP, which has a value of 40.8071 and 18.1673, respectively. Thus, the proposed approach computes the best $BRISQUE$ results for IMSET-19. Furthermore, NLBF computes the best average value as 34.10819. BFGKP and DCP come in second and third place, respectively, with average values of 35.91021 and 35.95769.

The validity of BFGKP on real-world images has been proven using $FADE$. It depends on contrast energy, entropy, and saturation of the image instead of human-rated judgments, the position of the camera, and depth maps. If the $FADE$ value is zero, the image is said to as fog-free. Figures 6.7(a)-6.7(c) depict the graphical representation of $FADE$ values calculated for IMSET-19, IMSET-20, and IMSET-21 using proposed

approach and SOTA techniques. It has been observed from figures that the least FADE values for IMGN74 through IMGN76 as 0.3785, 0.2811, and 0.3712, respectively; IMGN77 through IMGN79 as 0.3125, 0.6154, and 0.388, respectively; IMGN80 through IMGN94 as 0.4119, 0.3909, 0.2451, 0.3083, 0.4084, 0.2605, 0.1733, 0.1701, 0.195, 0.3745, 0.3059, 0.2889, 0.2425, 0.3451, and 0.2553, respectively; IMGN71 as 0.7333; IMGN76 as 0.7237, IMGN78 as 0.3898; and IMGN79 as 0.6161. As a result, the proposed approach computes the best $FADE$ results for all datasets.

The performance metrics and computation time are the two main criteria used to assess the computational complexity of a model. The experimental setup, including the workstation and hyperparameters, affects the computational cost. The average frame time to and from the server is 0.72 seconds. Table 6.3 depicts the comparative analysis of BFGKP (original and compressed image size) with the existing techniques in terms of computation time (best and second best performances are indicated by '*' and '#', respectively). It has been observed that BFGKP outperforms the existing algorithms when computed time is calculated for images with sizes of 256×256 and 640×640 as 0.044 second and 0.2567 second, respectively.

6.5 Real-time Dehazing Framework

The existing techniques were not tested on the real-time scenarios except for [176]. However, the atmospheric light was updated after an interval of 5000 frames. The flickering effects were observed in the results produced using optimized contrast enhancement [174]. The computational complexity of few video dehazing algorithms [176], [209], [210] is substantial. This section presents a novel static real-time dehazing framework for road safety.

A static dehazing framework has been designed in the laboratory of Electrical and Instrumentation Engineering Department, Thapar Institute of Engineering and Technology, India (see Figure 6.8). Figure 6.9 depicts the schematic diagram of real-time static dehazing framework that captures and processes the frames simultaneously. The developed dehazing framework consists of a GoPro Hero 7 camera to capture the image, a display screen or a system to view the output, a 4k Camlink to transfer the frames captured by the GoPro in 4k resolution, and a bluerigger chord with one HDMI port connected to Camlink and the other port connected to the GoPro camera. A novel real-time dehazing framework has been proposed using readily available electronic

Table 6.1 Comparative analysis on images using BFGKP and existing dehazing models in terms of $NIQE$

Images	DCP	CAP	ICAP	OTM-AAL	NLBF	CDP	IDE	CCDI	BFGKP
IMGN74	*2.92	3.95	3.76	5.23	2.97	3.45	3.51	3.01	#2.93
IMGN75	#3.26	4.16	4.22	5.4	3.34	3.38	3.51	3.39	*3.15
IMGN76	*2.96	3.96	3.63	6.55	3.13	3.17	3.31	3.22	#3.01
IMGN77	#9.23	9.98	11.63	*6.96	12.75	12.92	16.89	12.18	12.5
IMGN78	*6.27	#7.24	7.5	7.43	8.2	8.55	9.37	7.99	7.42
IMGN79	9.13	#8.91	10.54	*6.78	11.33	10.81	11.73	10.86	9
IMGN80	#2.56	3.53	2.98	2.6	*2.45	2.74	2.59	*2.45	3.28
IMGN81	*2.83	4.44	3.83	3.56	3.18	3.58	3.32	3.34	#3.01
IMGN82	#2.79	4.29	3.4	3.15	2.95	3.15	3.12	3.26	*2.62
IMGN83	*3.28	4.64	3.79	3.49	3.71	3.61	3.59	3.61	#3.48
IMGN84	*2.34	3.84	2.86	2.65	2.58	2.49	#2.38	2.52	2.65
IMGN85	4.45	3.98	4.26	3.87	4.18	4.71	#3.75	4.25	*3.07
IMGN86	#2.48	2.55	2.87	3.34	2.71	2.58	2.56	2.65	*2.46
IMGN87	3.21	#2.97	3.68	4.44	2.98	3.11	3.28	3.29	*2.93

Table 6.1 (continue) Comparative analysis on images using BFGKP and existing dehazing models in terms of $NIQE$

Images	DCP	CAP	ICAP	OTM-AAL	NLBF	CDP	IDE	CCDI	BFGKP
IMGN88	*2.96	*2.96	3.98	5.36	3.3	3.41	3.52	3.35	#3.16
IMGN89	*2.32	#2.39	2.85	3.43	2.51	2.61	2.82	2.63	2.41
IMGN90	*3.33	3.6	3.96	4.58	3.56	#3.48	3.77	3.64	3.65
IMGN91	#2.92	3.34	3.43	3.99	3.44	3.24	*2.85	3.48	3.11
IMGN92	3.37	3.28	3.68	4.45	3.46	#3.25	3.37	3.48	*3.02
IMGN93	*2.42	2.55	3.14	3.89	2.76	2.47	#2.43	2.65	2.75
IMGN94	#2.64	#2.64	2.91	3.2	2.68	2.67	*2.62	2.78	2.75
IMGN71	4.19	4.26	4.5	5.18	4.34	#3.91	4.29	4.29	*3.86
IMGN95	*3.66	3.84	3.96	*3.66	4.15	3.98	4.32	4.22	#3.77
IMGN96	*3.57	3.85	4.17	4.71	3.96	#3.83	4.14	3.99	6.3
IMGN97	*2.95	3.48	3.84	4.3	3.39	#3.27	3.66	3.41	6.02
IMGN98	*2.59	#2.71	2.93	5.44	2.88	2.87	3.13	2.98	2.85
IMGN99	#3.45	3.62	3.84	5.13	3.92	3.91	3.71	3.99	*3.43
Average	*3.632593	4.10963	4.301481	4.547037	4.104074	4.116667	4.353333	4.107778	#4.021852

Table 6.2 Comparative analysis on images using BFGKP and existing dehazing models in terms of $BRISQUE$

Images	DCP	CAP	ICAP	OTM-AAL	NLBF	CDP	IDE	CCDI	BFGKP
IMGN74	44.7184	44.5314	49.5727	*38.2897	45.3569	#40.8148	47.5356	45.8682	44.6169
IMGN75	34.6098	34.3525	35.1737	*29.4355	33.7735	33.552	33.4907	33.7445	#31.822
IMGN76	44.4763	#44.1557	48.0246	60.6971	45.4671	45.526	46.5362	45.6819	*43.4979
IMGN77	47.2573	47.3511	47.7869	48.0805	#46.2523	46.7108	46.6674	46.8443	*44.6355
IMGN78	39.4754	39.4825	40.4203	49.287	*36.4434	#36.7416	38.0178	37.5824	39.1404
IMGN79	46.9248	#45.4647	46.5696	48.4975	*45.3832	45.7994	45.7524	46.013	45.4852
IMGN80	#25.887	38.375	37.011	36.185	*24.562	36.004	34.512	30.726	37.5373
IMGN81	*26.395	40.365	#34.371	35.329	35.919	36.949	36.304	35.33	40.4977
IMGN82	26.868	35.328	35.642	33.61	*20.86	33.595	31.118	#24.714	30.9008
IMGN83	#21.652	32.927	32.104	29.236	28.877	*18.48	30.112	25.027	29.3713
IMGN84	37.07	37.742	24.897	29.568	20.886	32.351	*7.343	#15.85	36.4708
IMGN85	38.773	28.123	#26.263	26.649	31.821	40.018	*23.345	34.244	40.5911
IMGN86	30.2965	32.7754	33.1579	29.9778	*25.9585	29.4168	#26.2098	27.3838	33.7083
IMGN87	42.9811	41.4149	42.4501	42.9149	*30.2193	45.2395	44.5344	44.7172	#40.8071

Table 6.2 (continue) Comparative analysis on images using BFGKP and existing dehazing models in terms of $BRISQUE$

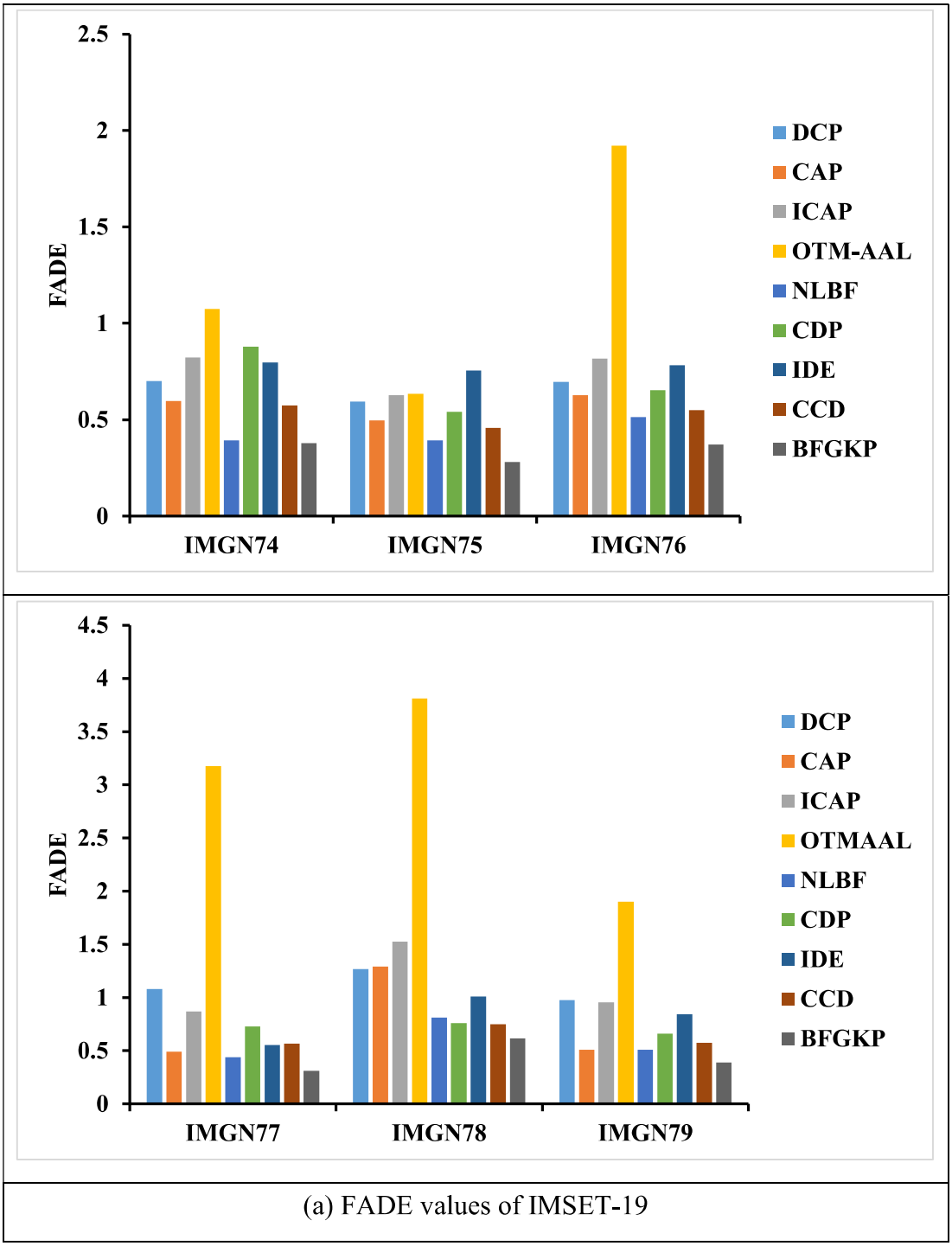
Images	DCP	CAP	ICAP	OTM-AAL	NLBF	CDP	IDE	CCDI	BFGKP
IMGN88	35.6407	*28.5671	#28.7339	38.2713	30.9556	36.7169	40.2931	36.515	34.3594
IMGN89	*29.7719	38.6387	35.6144	38.8694	#30.4286	36.2068	40.2397	37.4371	35.5821
IMGN90	22.3428	23.2025	13.6328	21.9459	*8.0887	15.6827	#9.9225	14.7228	30.8306
IMGN91	31.3945	41.4775	#29.2148	*28.8672	37.9311	32.0652	36.389	36.2701	37.3344
IMGN92	42.2464	35.746	25.0331	#24.8334	38.7648	43.7157	47.1592	41.4712	*23.2286
IMGN93	33.5532	38.3494	32.6383	*26.8095	#29.9114	30.4778	35.9509	34.9775	34.5277
IMGN94	#13.758	24.5777	23.1491	22.2468	27.7893	31.9571	16.6603	20.096	*12.4897
IMGN71	*44.9341	57.7804	52.8908	63.1637	53.0001	#45.9223	57.1896	48.7129	49.8599
IMGN95	37.7948	38.9979	38.9673	*30.0509	40.4597	#31.193	43.0299	43.8041	40.1676
IMGN96	49.5907	43.9881	40.5476	*34.5007	48.9231	44.3635	50.5883	48.8798	#35.4655
IMGN97	50.2249	44.9775	43.9013	*35.1677	45.5674	38.6415	51.2102	52.5379	#35.3689
IMGN98	25.7895	25.7363	26.1915	62.2454	*17.0905	19.3357	20.1975	18.9933	#18.1673
IMGN99	46.4314	47.3097	47.2807	62.9126	#40.2315	44.193	*40.1723	43.457	43.1116
Average	35.95769	38.21248	35.97183	38.0608	*34.10819	35.98774	36.3141	35.98522	#35.91021

Table 6.3 Comparative analysis on images using BFGKP and other existing dehazing techniques in terms of computation time (in seconds)

Images	DCP	CAP	ICAP	OTM-AAL	NLBF	CDP	IDE	CCDI	BFGKP (640×640)	BFGKP (256×256)
IMGN74	0.34	1.25	0.85	3.37	0.52	0.59	0.52	0.67	#0.24	*0.043
IMGN75	0.37	1.28	0.86	3.31	0.56	0.61	0.54	0.64	#0.27	*0.041
IMGN76	0.35	1.24	0.83	3.36	0.57	0.62	0.53	0.65	#0.25	*0.046
IMGN77	0.37	1.25	0.84	3.34	0.55	0.58	0.53	0.64	#0.25	*0.047
IMGN78	0.34	1.26	0.84	3.32	0.52	0.57	0.52	0.67	#0.25	*0.044
IMGN79	0.33	1.27	0.87	3.32	0.53	0.59	0.56	0.65	#0.26	*0.046
IMGN80	0.31	1.26	0.88	3.34	0.54	0.6	0.54	0.69	#0.28	*0.043
IMGN81	0.33	1.23	0.88	3.34	0.53	0.57	0.53	0.65	#0.26	*0.045
IMGN82	0.31	1.27	0.84	3.34	0.56	0.55	0.55	0.63	#0.21	*0.045
IMGN83	0.34	1.24	0.84	3.31	0.57	0.59	0.56	0.68	#0.24	*0.047
IMGN84	0.32	1.25	0.87	3.37	0.52	0.62	0.53	0.62	#0.25	*0.042
IMGN85	0.38	1.27	0.83	3.35	0.58	0.6	0.57	0.64	#0.27	*0.044
IMGN86	0.36	1.26	0.85	3.36	0.55	0.6	0.55	0.63	#0.23	*0.047
IMGN87	0.37	1.23	0.83	3.31	0.55	0.61	0.53	0.63	#0.23	*0.049

Table 6.3 (continue) Comparative analysis on images using BFGKP and other existing dehazing techniques in terms of computation time (in seconds)

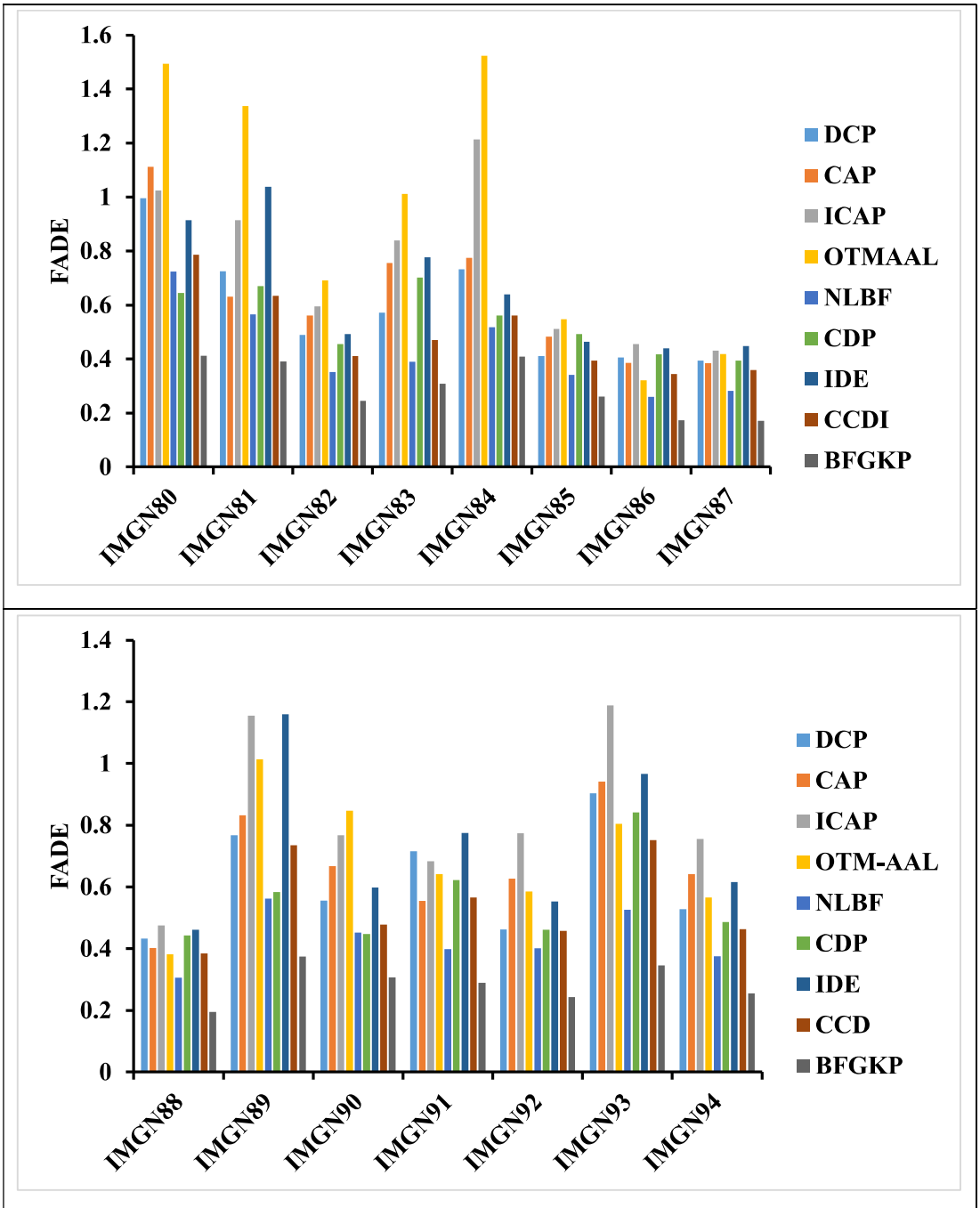
Images	DCP	CAP	ICAP	OTM-AAL	NLBF	CDP	IDE	CCDI	BFGKP (640×640)	BFGKP (256×256)
IMGN88	0.35	1.25	0.86	3.33	0.57	0.64	0.52	0.66	#0.26	*0.047
IMGN89	0.35	1.26	0.83	3.33	0.52	0.59	0.52	0.66	#0.25	*0.044
IMGN90	0.39	1.27	0.85	3.33	0.54	0.58	0.55	0.64	#0.28	*0.044
IMGN91	0.31	1.24	0.88	3.35	0.53	0.61	0.54	0.67	#0.24	*0.043
IMGN92	0.34	1.22	0.86	3.37	0.53	0.62	0.57	0.65	#0.26	*0.046
IMGN93	0.31	1.24	0.84	3.32	0.54	0.6	0.55	0.68	#0.27	*0.046
IMGN94	0.34	1.23	0.83	3.31	0.52	0.59	0.53	0.63	#0.29	*0.045
IMGN71	0.33	1.27	0.85	3.34	0.55	0.59	0.54	0.64	#0.25	*0.045
IMGN95	0.37	1.25	0.81	3.37	0.56	0.58	0.56	0.65	#0.28	*0.047
IMGN96	0.35	1.23	0.86	3.34	0.57	0.61	0.54	0.61	#0.29	*0.042
IMGN97	0.36	1.22	0.88	3.38	0.53	0.62	0.54	0.65	#0.23	*0.043
IMGN98	0.32	1.25	0.89	3.35	0.54	0.6	0.54	0.68	#0.26	*0.043
IMGN99	0.33	1.27	0.83	3.36	0.55	0.62	0.56	0.65	#0.28	*0.044
Average	0.3434	1.2503	0.8511	3.3414	0.5444	0.5981	0.5414	0.6503	#0.2567	*0.0447



(a) FADE values of IMSET-19

Fig. 6.7 Comparative analysis of images using BFGKP and other existing dehazing techniques based on *FADE*

equipment and devices. Tables 6.4 – 6.6 present the technical specifications of the equipment used in the proposed framework [211]–[213]. The movement of this framework is restricted up to the length of bluerigger cable.



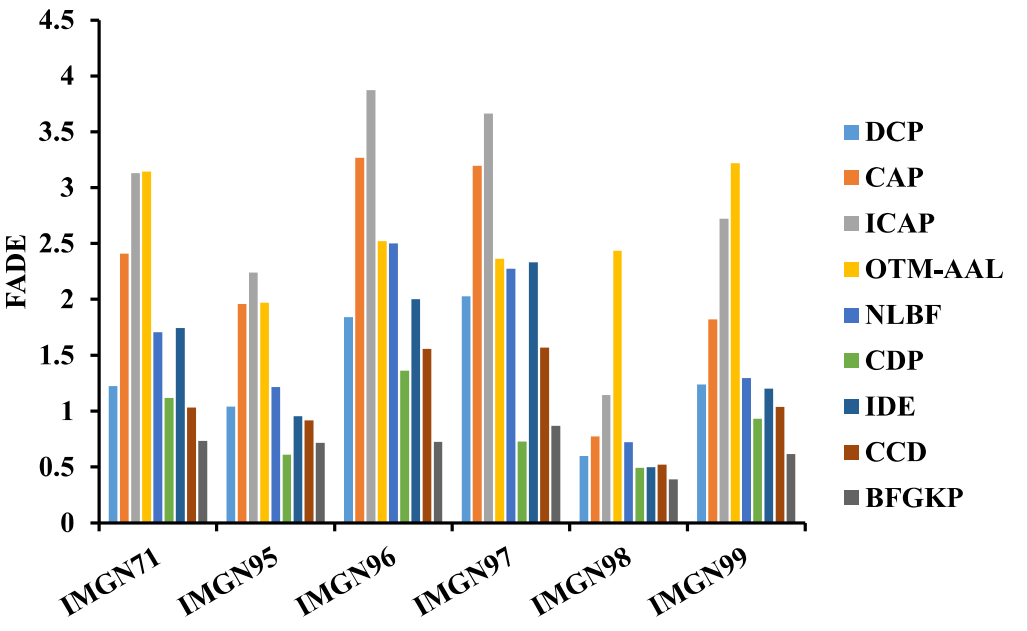
(b) FADE values of IMSET-20

Fig. 6.7 (continue) Comparative analysis of images using BFGKP and other existing dehazing techniques based on *FADE*

6.6 Dynamic Dehazing Framework

Real-time dynamic frames have been captured during the process of video acquisition. The issue of hazy frame reconstruction for road safety is being addressed in this initial phase. In this system, a high-definition GoPro Hero 7 has been used to capture

real-time scenarios of the environment. The frames have been captured and transferred to the display via blurigger chord and Camlink 4k. The maximum resolution of the frame acquired by the camera is 4000×3000 pixels. However, it can be reduced to 720×480



(c) FADE values of IMSET-21

Fig. 6.7 (continue) Comparative analysis of images using BFGKP and other existing dehazing techniques based on *FADE*

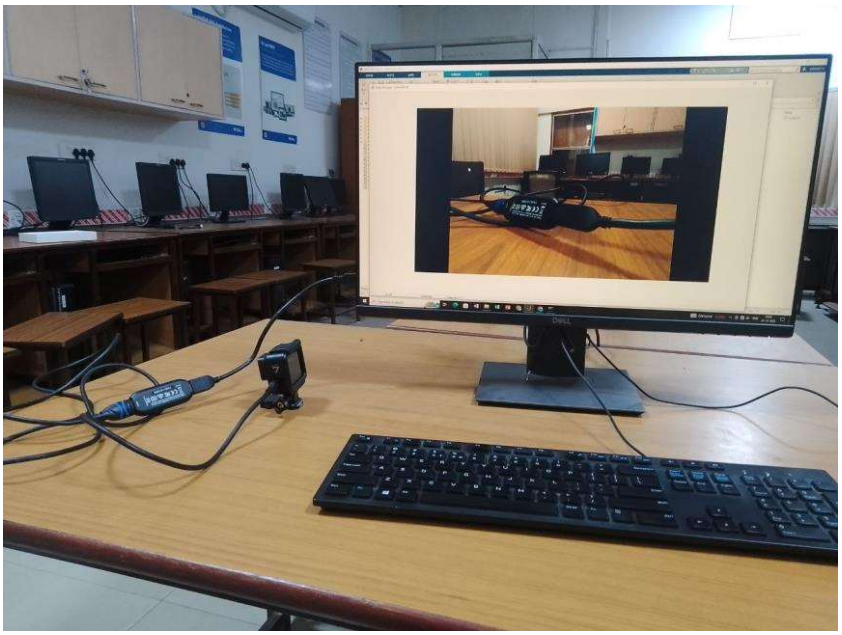


Fig. 6.8 Real-time static dehazing framework

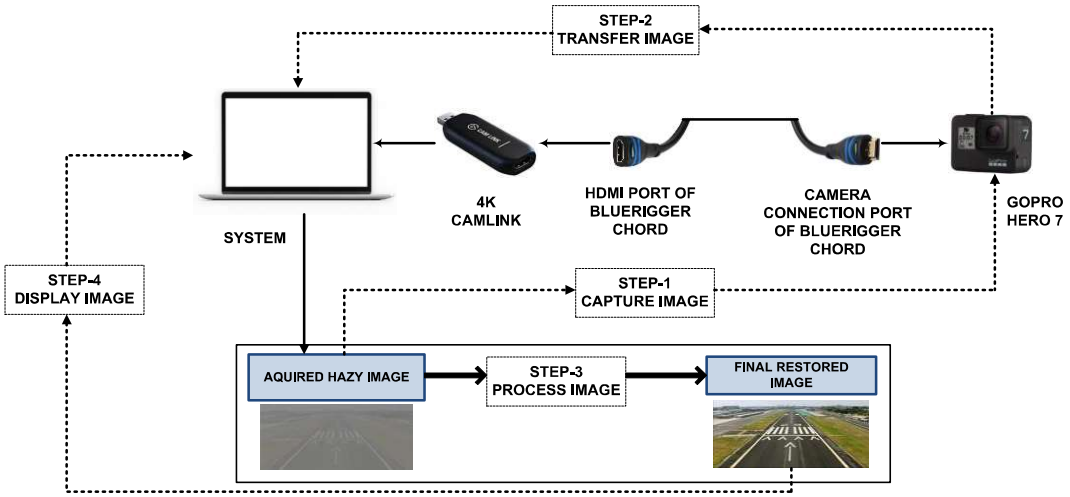


Fig. 6.9 Static Diagram of real-time dehazing framework

Table 6.4 Technical specifications of GoPro Hero 7 [211]

Specification	Value
Effective Pixels	12 megapixels
Sensor Type	CMOS
Image Resolution	4000 × 3000p
Battery Description	1220mAh Lithium-ion
Wi-Fi	Yes
GPS	Yes

Table 6.5 Technical specifications of Bluerigger chord [212]

Specification	Value
Cable	4k HDMI – Ultra HD
Resolution	1080p
Ethernet	Yes

Table 6.6 Technical specifications of Camlink 4k (fps – frames per second) [213]

Specification	Value
Input	HDMI (unencrypted)
Supported Resolutions	720×480p (up to 60fps)
	720×576p (up to 50fps)
	1280×720p (up to 60fps)
	1920×1080p (up to 60fps)
	3840×2160p (up to 30fps)

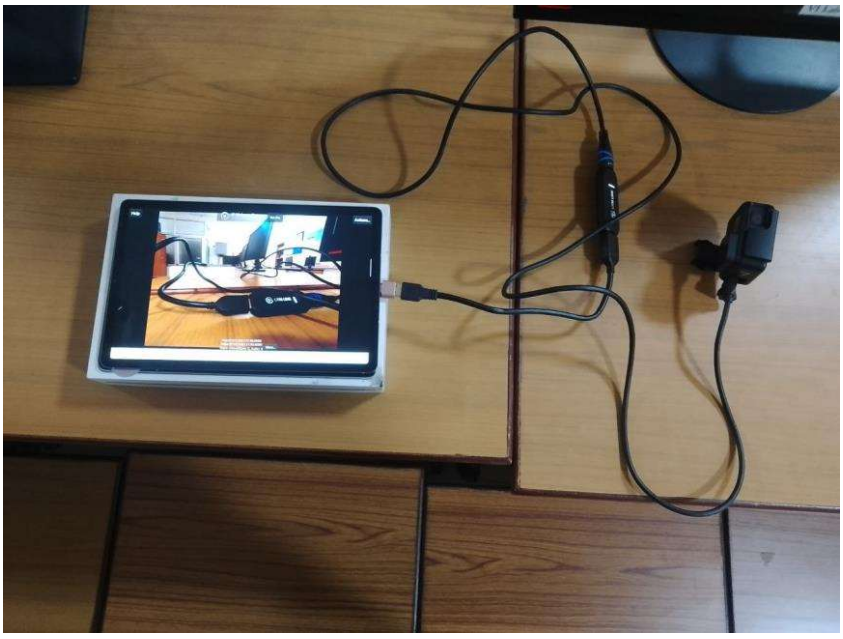


Fig. 6.10 Real-time dynamic defogging framework

pixels with 60 frames per second (fps) using Camlink 4k. MATLAB software supported by PC has the "MATLAB Support Package for IP Cameras" package to access videos from the server. The foggy frames have been sequentially processed using BFGKP.

Another application called Google Remote Desktop designed by Google LLC has been used to remotely access the PC from a Galaxy S6 tablet. The tablet can remotely access the PC irrespective of the user's location. Figure 6.10 depicts a dynamic defogging framework that captures and processes the frames simultaneously.

Figure 6.11 depicts the real-time dehazing scenes with inputs and outputs of IMSET-23 (i.e., IMGN105, IMGN106, and IMGN107) displayed simultaneously. The output of a hazy frame is displayed after 1-2 seconds, the shortest amount of time needed for data to be sent to the server, processed, and then sent back to the display.

6.7 Standalone Application

A standalone application has been proposed for dehazing using edge computing and cloud computing [214], [215]. The results are computed for network-related algorithms such as the Internet of Things (IoT) and big data analysis. If the results are monitored using devices near to the source or consumer of data, then the process is called edge computing. The algorithm designed for single image dehazing has been modified to perform real-time dehazing using MATLAB Mobile application designed by MathWorks, Inc. This application is accessible through the Google play store for android-powered devices. Once the application has been downloaded, the account

Images	IMGN105	IMGN106	IMGN107
Hazy Image			
Dehazed Image			

Fig. 6.11 Visual analysis of real-time hazy frames (IMSET-23) using BFGKP technique

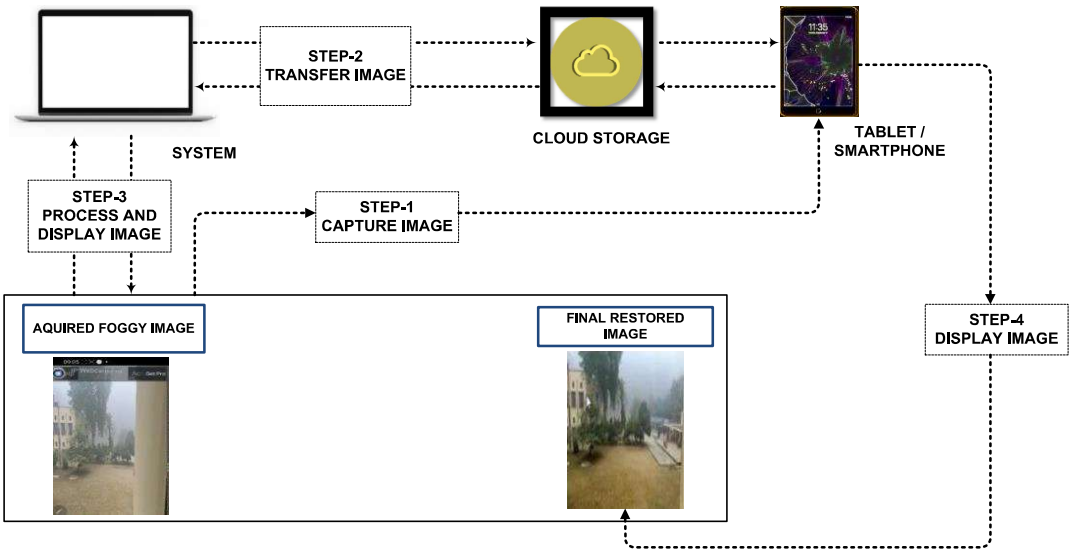


Fig. 6.12 Schematic diagram for web-based defogging application

details for a licenced version of MATLAB must be entered to access the MATLAB drive and different packages available for cloud storage of MATLAB, i.e., MATLAB online. Figure 6.12 depicts schematic diagram of a web-based defogging application.

6.8 Summary

A robust and fast dehazing technique called BFGKP has been proposed in this chapter. It uses gray-world kernel and bounding function to generate the transmission map of hazy images. BFGKP has been found superior to existing prior-based algorithms namely, color cast dependency, image dehazing and exposure, bounded channel difference prior, color attenuation prior, and dark channel prior among others. The computation time has improved by 7.68 times. Also, $FADE$ value has been found better for 92.59% images. In this chapter, communication between a PC and a GoPro Camera is established using a connecting cable and Camlink 4k, so that the real-time frame acquired by the GoPro is instantaneously dehazed using the proposed dehazing algorithm. Also, a web-based application has been developed to generate real-time dehazed results by establishing communication between a Samsung Galaxy S6 tablet and a GoPro Hero 7 Camera.

CHAPTER-7
DESMOGGING OF STILL IMAGES USING RESIDUAL REGRESSION NETWORK AND MORPHOLOGICAL EROSION

Smog is one of the air pollutants that makes it difficult for drivers to see. Smog is a mixture of fog and smoke that produces black fumes and reduces the visibility of drivers within a range of one kilometre. The small size and high density of smog particles, in comparison to other air pollutants, impede drivers' vision on the road. This chapter presents a novel desmogging architecture called the residual regression network for single image desmogging.

7.1 Motivation

The prior-based techniques are used to extract the features of the smoggy image. However, these techniques perform poorly for the visibility restoration of objects with a color similar to that of atmospheric light. For instance, smoggy images usually have uniform distortion in both sky and non-sky regions. Therefore, a regression-based residual CNN architecture followed by morphological erosion has been proposed. The residual regression network (RRNet) estimates the transmission map. Thereafter, it has been refined through erosion-based morphological operations. RRNet preserves texture and edges by computing the thickness of smog around them. It boosts the flow of features by adding the outputs of previously deconvolved layers to reduce the loss of information.

The majority of visibility restoration models utilize classification-based techniques to classify the images into hazy and haze-free classes. CNN-based autoencoders have been used to achieve the task of generating feature maps from derived inputs. It increases the computational time of the network to store all the features in a one-dimensional layer. He et al. [196] discussed the importance of residual connections in a convolution neural network called residual nets (ResNets). The residual connections have been used to prevent the loss of information, reduce the training error, and increase the accuracy of the network. Initially, the performance of plain networks with different layers such as 18-layer and 34-layer plain nets, had been evaluated to reconstruct smoggy image. The models were designed using 1.28 million training images, 50,000 validation images, and 0.1 million test images from the ImageNet

database. However, these networks suffer from vanishing gradient issues, i.e., the weights do not get updated after a definite number of layers. Thus, it becomes difficult to train the model. So, RRNet has been proposed as it comprises the advantages of residual connections and a regression model to obtain the output image with dimensions similar to the input image. The estimated transmission map may contain extensive details. However, only salient features are required to reconstruct the smoggy image. These salient features have been preserved using morphological operations [216]. Erosion is one of the morphological operations that adjusts the intensity of the pixels based on the neighbouring pixel values [217].

7.2 Architecture of Desmogging Model

In the proposed desmogging model, a light residual regression network (RRNet) has been proposed to obtain desmogged images. It has been used to estimate the transmission map from the smoggy image. The erosion operation has been applied to refine the transmission map. The atmospheric light has been estimated using an order statistic filter for edge smoothing. If the smoggy image is completely restored, it will look artificial. The smoggy image has been reconstructed using a modified visibility restoration model. Figure 7.1 depicts the diagrammatic flow of the proposed smog removal technique.

7.2.1 Estimation of Transmission Map

Due to smog gradients, the visibility is reduced to less than 1km, and they may not be uniformly distributed. The different stages have been used to estimate the transmission map from the given image (see Figure 7.2). In the first stage, the input image has been passed through a convolution layer to extract shallow features. The second stage is a residual network that has been used to reduce the thickness of smog to find the depth of the image. The dimensions of the computed image must be the same at both terminals of the residual connection for identity mapping. In ResNets, the identity mapping was performed by utilizing padded dimensions of zero entries and a stride of two [196]. In the proposed approach, three groups of layers have been designed such that each group consists of convolution, pooling, and unpooling layers. The pool size in the architecture is 2×2 for all groups of layers. The element-wise summation of the output of each group has been performed by using connectors.

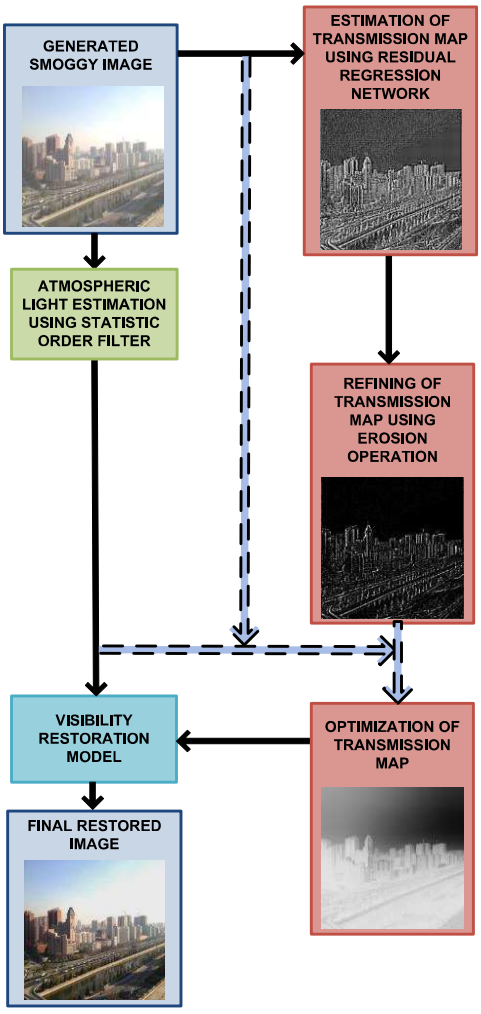


Fig. 7.1 Diagrammatic flow of proposed smog removal technique

The combination of pooling and unpooling layers has been used in the proposed approach such that the size of the computed image is constant at the start and end of the residual connection for identity mapping. It has been used to introduce non-linearity in the model to make it effective for real-world images. The third stage initializes with a convolution layer to find artifacts in transmission map. It has been followed by a combination of convolution-batch normalization-rectified linear unit to extract the shallow features. Batch normalization (batchnorm) layer standardizes the inputs by re-scaling and re-centring them. It makes the network stable and faster. Thus, it reduces the computation time and complexity of the network [50]. The normalized output has been passed through an activation function known as a rectified linear unit (ReLU).

The fourth stage consists of another residual network that comprises three groups of layers (i.e., convolution-batchnorm-pooling-unpooling) followed by a different group of layers (i.e., convolution-batch normalization) for final enhancement of features and

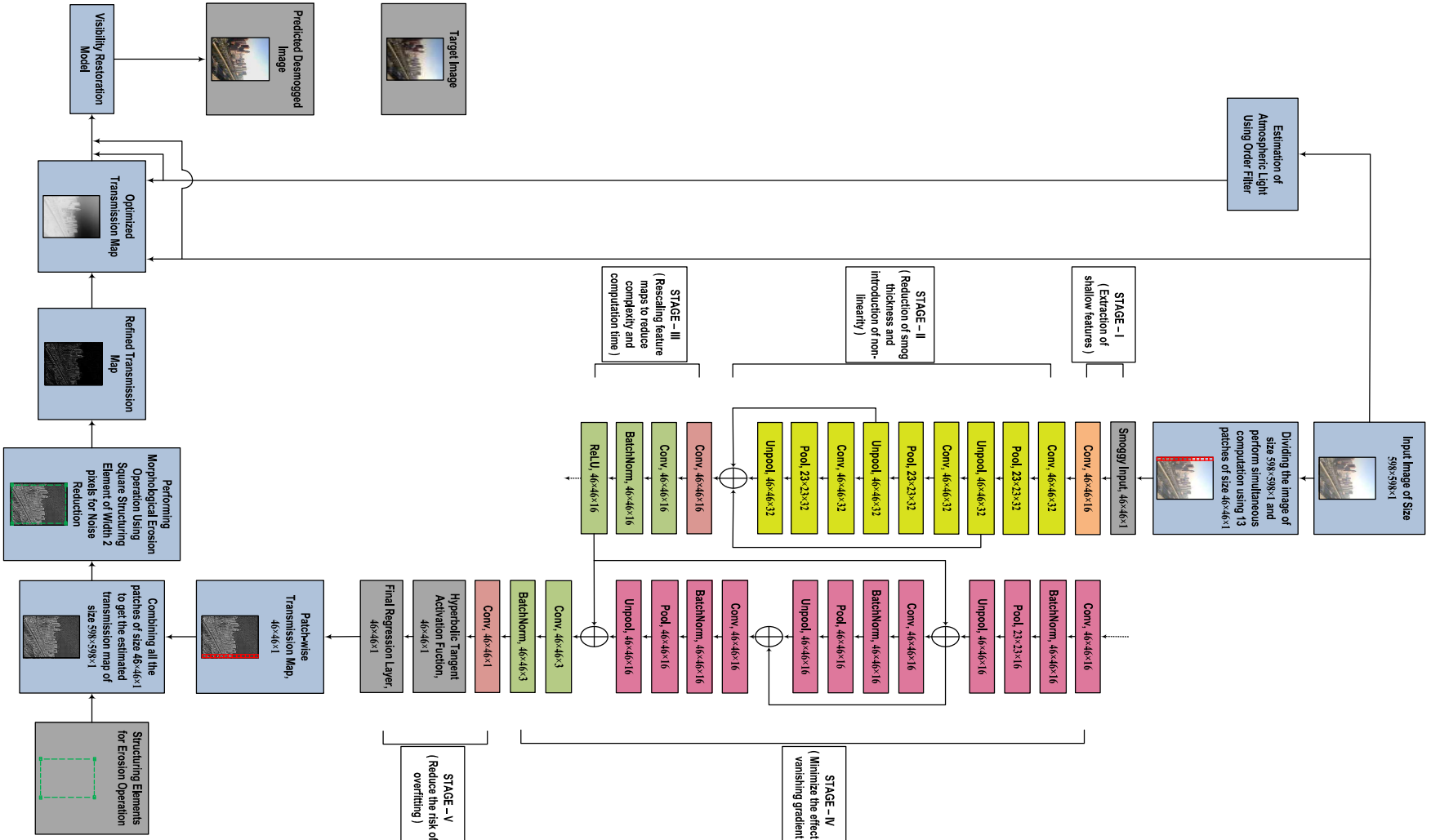


Fig. 7.2 Detailed illustration of residual regression and morphological erosion based desmogging technique

halo effect reduction. It also performs simultaneous operation of downsampling and upsampling in all three stages to reduce computational complexity. The last stage consists of a convolution layer and a hyperbolic tangent layer for early convergence of the loss and reduce the risk of overfitting. The final regression layer generates a refined transmission map of similar size to the input layer.

The complexity of the proposed deep learning architecture has been further discussed in terms of the model framework, its size, and the expressive capacity of the activation functions used to design the model [218]. The boundary of ReLU and tanh activation functions is given in Eqs. (7.1) and (7.2), respectively.

$$ReLU_{complexity} = O(M^{N_{in}}) \tag{7.1}$$

$$Tanh_{complexity} = 2^{\frac{M(M-1)}{2}} O((N_{in}l + N_{in})^{N_{in}+M}) \tag{7.2}$$

where N_{in} is the number of inputs. M is the number of hidden units. l is the number of hidden layers.

Table 7.1 depicts the layer-wise analysis of the proposed desmogging model. It presents the number and size of filters in each convolution layer. The stride and padding values of the convolution and pooling layers of the architecture have also been presented. Furthermore, the activations, learnable parameters, and total learnables for each layer have been discussed.

7.2.2 Morphological Erosion

The transmission map estimated by using RRNet has been noisy and presents extensive details that can be minimized via morphological operations. The mathematical morphology operations such as dilation and erosion preserve the significant details of the feature maps by adjusting their pixel values based on the neighbouring pixel values.

In this chapter, grayscale morphological erosion has been implemented to refine the transmission map. In this morphological operation, a square structuring element with a width of 2 pixels has been utilized so that the computed image contains only true values. The erosion of the estimated transmission map with respect to the structuring element in the window returns the pixel with the minimum value in that window's centre. It is mathematically represented as follows [219]:

$$T_1 = (T \exists W)(x,y) = \underbrace{min}_{(m,n)} \left(T(x+m, y+n) - W(m,n)\right) \tag{7.3}$$

where T_1 is the refined grayscale transmission map. T is the estimated grayscale transmission map of size 598×598. $W(x,y)$ are the structuring elements in the window of size 2×2. ∃ represents the erosion operation.

The erosion has been defined on the domain $D = \{(m,n) | m \epsilon \{1,2,\dots,i\}; n \epsilon \{1,2,\dots,j\}\}$. Sometimes, the entire domain is not covered by geometric probe. Thus, $W(x,y)$ is given as:

$$W(x,y) = \begin{Bmatrix} \geq 0, & (x,y)\epsilon G \\ -\infty, & otherwise \end{Bmatrix} \tag{7.4}$$

Table 7.1 Layer-wise analysis of the proposed desmogging model

Type	No. of Filters	Size of Filters	Stride	Padding	Activations	Learnable Parameters	Total Learnables
Input Layer	-	-	-	-	46×46×1	-	0
Convolution	16	3×3×1	[1,1]	Same	46×46×16	Weights: 3×3×1×16 Bias: 1×1×16	160
Convolution	32	3×3×16	[1,1]	Same	46×46×32	Weights: 3×3×16×32 Bias: 1×1×32	4640
Max Pooling	-	-	[2,2]	[0,0,0,0]	23×23×32	-	0
Max Unpooling	-	-	-	-	46×46×32	-	0
Convolution	32	3×3×32	[1,1]	Same	46×46×32	Weights: 3×3×32×32 Bias: 1×1×32	9248
Max Pooling	-	-	[2,2]	[0,0,0,0]	23×23×32	-	0
Max Unpooling	-	-	-	-	46×46×32	-	0
Convolution	32	3×3×32	[1,1]	Same	46×46×32	Weights: 3×3×32×32 Bias: 1×1×32	9248
Max Pooling	-	-	[2,2]	[0,0,0,0]	23×23×32	-	0
Max Unpooling	-	-	-	-	46×46×32	-	0
Addition	-	-	-	-	46×46×32	-	0
Convolution	16	3×3×32	[1,1]	Same	46×46×16	Weights: 3×3×32×16 Bias: 1×1×16	4624
Convolution	16	3×3×16	[1,1]	Same	46×46×16	Weights: 3×3×16×16 Bias: 1×1×16	2320

Table 7.1 (continue) Layer-wise analysis of the proposed desmogging model

Type	No. of Filters	Size of Filters	Stride	Padding	Activations	Learnable Parameters	Total Learnables
Batch Normalization	-	-	-	-	46×46×16	Offset: 1×1×16 Scale: 1×1×16	32
ReLU	-	-	-	-	46×46×16	-	0
Convolution	16	3×3×16	[1,1]	Same	46×46×16	Weights: 3×3×16×16 Bias: 1×1×16	2320
Batch Normalization	-	-	-	-	46×46×16	Offset: 1×1×16 Scale: 1×1×16	32
Max Pooling	-	-	[2,2]	[0,0,0,0]	23×23×16	-	0
Max Unpooling	-	-	-	-	46×46×16	-	0
Addition	-	-	-	-	46×46×16	-	0
Convolution	16	3×3×16	[1,1]	Same	46×46×16	Weights: 3×3×16×16 Bias: 1×1×16	2320
Batch Normalization	-	-	-	-	46×46×16	Offset: 1×1×16 Scale: 1×1×16	32
Max Pooling	-	-	[2,2]	[0,0,0,0]	23×23×16	-	0
Max Unpooling	-	-	-	-	46×46×16	-	0
Addition	-	-	-	-	46×46×16	-	0
Convolution	16	3×3×16	[1,1]	Same	46×46×16	Weights: 3×3×16×16 Bias: 1×1×16	2320
Batch Normalization	-	-	-	-	46×46×16	Offset: 1×1×16 Scale: 1×1×16	32
Max Pooling	-	-	[2,2]	[0,0,0,0]	23×23×16	-	0
Max Unpooling	-	-	-	-	46×46×16	-	0
Addition	-	-	-	-	46×46×16	-	0
Convolution	3	3×3×16	[1,1]	Same	46×46×3	Weights: 3×3×16×3 Bias: 1×1×3	435
Batch Normalization	-	-	-	-	46×46×3	Offset: 1×1×3 Scale: 1×1×3	6
Convolution	1	3×3×3	[1,1]	Same	46×46×1	Weights: 3×3×3 Bias: 1×1	28
Tanh	-	-	-	-	46×46×1	-	0
Regression Output	-	-	-	-	46×46×1	-	0

7.2.3 Estimation of Airlight

It uses a 2D order statistic filter in which the input image has been represented in form of 2D matrix [193]. The color channel with the fewest (non-zero) intensities is selected. If more than one color channel contains pixels with similar least non-zero numeric values of intensities, then the channel with the maximum number of such pixels has been considered. The pixels of that channel have been indexed in increasing order of their intensity, and each pixel has been replaced by the value of ordered index present in the domain. It is mathematically represented as follows:

$$WOS_o(n) = WOS_{i_{(w_o)}} = w_o'^{\,th}\,(largest[WOS_i(n) \times w]) \tag{7.5}$$

where $WOS_o(n)$ is output of weighted order statistic (WOS) filter. $WOS_i(n)$ is the input of the WOS filter. $WOS_{i_{(w_o)}}$ is order statistics for $w_o = 1, 2, \ldots, \mathrm{W}_N$. w is weight or strength of connection between two layers.

After the refinement of the transmission map and estimation of atmospheric airlight, the transmission map has been optimized using the input image, airlight, and refined transmission map. The optimized transmission map is the ratio of the sum of the square of the difference of grayscale input image and estimated airlight, and the pixels of the refined transmission map have been scaled down to 0.1 times their original values to the square of the difference of the grayscale input image and airlight. Finally, the edges have been smoothed and preserved by using a guided filter.

7.2.4 Image Reconstruction

The smoggy image has been reconstructed using a visibility restoration algorithm. Eq. (1.1) has been rewritten to obtain the desmogged output as shown in Eq. (7.6). It uses the input smoggy image, the obtained transmission map, and the atmospheric light. The exposure of a desmogged image to light is directly proportional to "gamma."

$$J(x,y) = \frac{I(x,y) - A}{[max(t(x,y), t_0(x,y))]^{gamma}} + A \tag{7.6}$$

In the proposed scheme, a minimum amount of smog in $t(x,y)$ has been preserved and assumed as $t_0(x,y)$ with a value of 0.0001. The value of gamma is assumed to be 0.02 for less smoggy images and 0.04 for dense smoggy images to minimise the texture distortion while implementing visibility restoration algorithm.

In Figure 7.2, Conv means the convolution operation performed by using a kernel of dimensions 3×3. ReLU means rectified linear unit, which is an activation function. Pool means a pooling operation with a pool size of 2×2. Unpool represents the unpooling operation, BatchNorm means batch normalization to standardize the inputs, ⊕ is the connector used for element-wise summation of layers. The dimensions of the image (height × width × number of filters) have been mentioned along each step.

7.3 Experimental Results and Discussion

The performance of the proposed regression-based model has been evaluated through qualitative and quantitative analysis. The results have been compared with five existing desmogging techniques namely, reconciling color (RC), information gain based bilateral filter (IGBF), illumination channel prior (ICP), variational minimization (VM) technique, and smoggy channel prior (SCP). The performance of the above-mentioned techniques has been evaluated on real-world smoggy images obtained from [118], [119], [122], [129], [130], [134], [135], [163] – [165].

7.3.1 Experimental Setup

The proposed model has been simulated on Intel (R) Core (TM) i7-6700 CPU @3.40GHZ with 8GB RAM and a 64-bit operating system on MATLAB R2021a. Clear image datasets namely, HUDRS, FRIDA, FRIDA-2, and RESIDE have been used to evaluate the performance of the proposed method. These images have been synthesized to obtain the smoggy images, and resized to 598×598 pixels. The training dataset consists of 3000 synthesized images. The trained architecture has been further tested on 600 smoggy images. The test images have been divided into six test sets.

IMSET-24, IMSET-25, IMSET-26, and IMSET-27 have synthetically generated smoggy images obtained from the clear images present in FRIDA, FRIDA2, HUDRS, and RESIDE respectively. IMSET-28 has real world smoggy images available online and comparative results have been obtained for the proposed technique, RRNet and five well-known desmogging techniques. IMSET-29 has real world images that have been obtained from [118], [119], [122], [129], [130], [134], [135], [163] – [165] to perform a comparative study of various performance metrics.

During training, 13 patches of the size of 46×46 pixels have been selected from the input smoggy image to extract the salient features and reduce the loss of information. All the patches have been processed simultaneously to reduce computation time. Adam

optimizer has been utilized with a gradient decay factor of 0.99 [220]. The L2 norm gradient threshold method has been used with a threshold value of 1. The initial learning rate and batch size are 0.009 and 5, respectively. L2 loss is more stable than L1 loss. L2 loss has been used to solve the problem of overfitting while training the model. For the target images and predicted images using the proposed model, it is given by Euclidean distance [221].

$$L_2\ loss = \sum_{i=1}^{n} \left(S_{0_{true}} - S_{0_{predicted}} \right)^2 \tag{7.7}$$

where $S_{0_{true}}$ and $S_{0_{predicted}}$ are pixels of ground truth image and predicted image, respectively.

7.3.2 Qualitative Analysis

The histogram of the images computed using the proposed architecture has been plotted and compared with the histogram plots of synthetically generated smoggy images and their respective ground truths. In Figures 7.3 and 7.4, the histogram plots of six smoggy images, and their respective desmogged images from IMSET-24 (i.e., IMGN108 to IMGN110) and IMSET-25 (i.e., IMGN111 to IMGN113) have been generated by the proposed architecture. The histogram plots of desmogged images have been similar to the plots generated for the respective ground truth images. It has been observed that the images computed using RRNet approaches towards uniform distribution of light.

The desmogging results of 33 sample images are presented in Figures 7.5 – 7.8. Figure 7.5 depicts the desmogging results of test images from HUDRS. It presents various steps for desmogging images, namely an estimated transmission map, a refined transmission map, an optimized transmission map, and a restored smoggy image, respectively. It has been observed that the result of IMGN114 generated by using the proposed architecture has more contrast than the ground truth image. Also, the computed results of IMGN115 and IMGN116 have more pixel intensity as compared to the ground truth images. Due to this reason, the color of sky region gets distorted. Furthermore, the noise is absent and the edges are smooth due to the implementation of morphological erosion and a guided filter.

The visual comparison of RRNet has been made with five well-known desmogging techniques in Figures 7.6 – 7.8. The computed results for IMSET-27 are

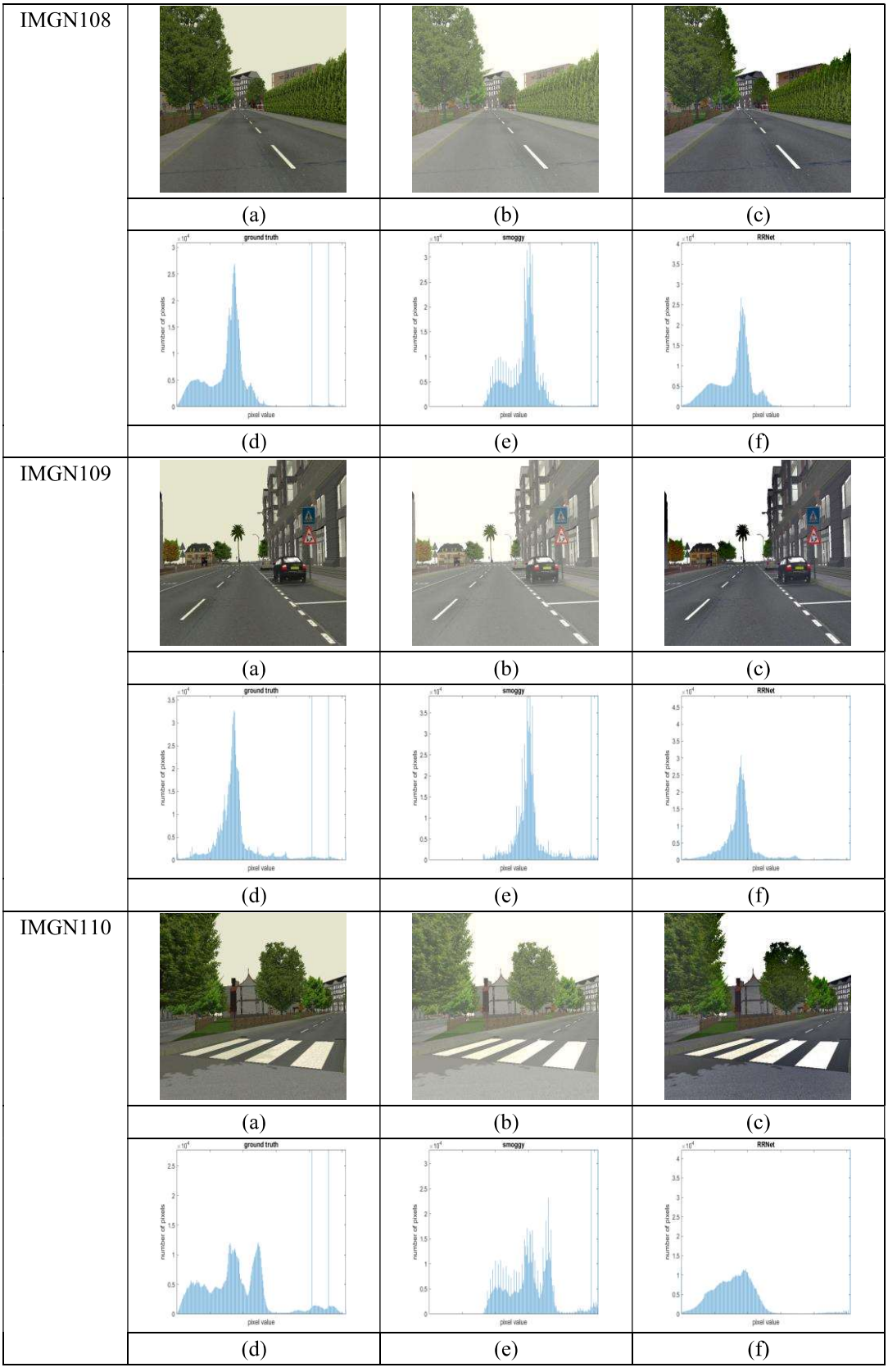


Fig. 7.3 Qualitative analysis of RRNet by using histogram plot. Results of sample images of IMSET-24 from FRIDA (a) ground truth; (b) smoggy image; (c) desmogged image using RRNet; (d) histogram plot of (a); (e) histogram plot of (b); (f) histogram plot of (c)

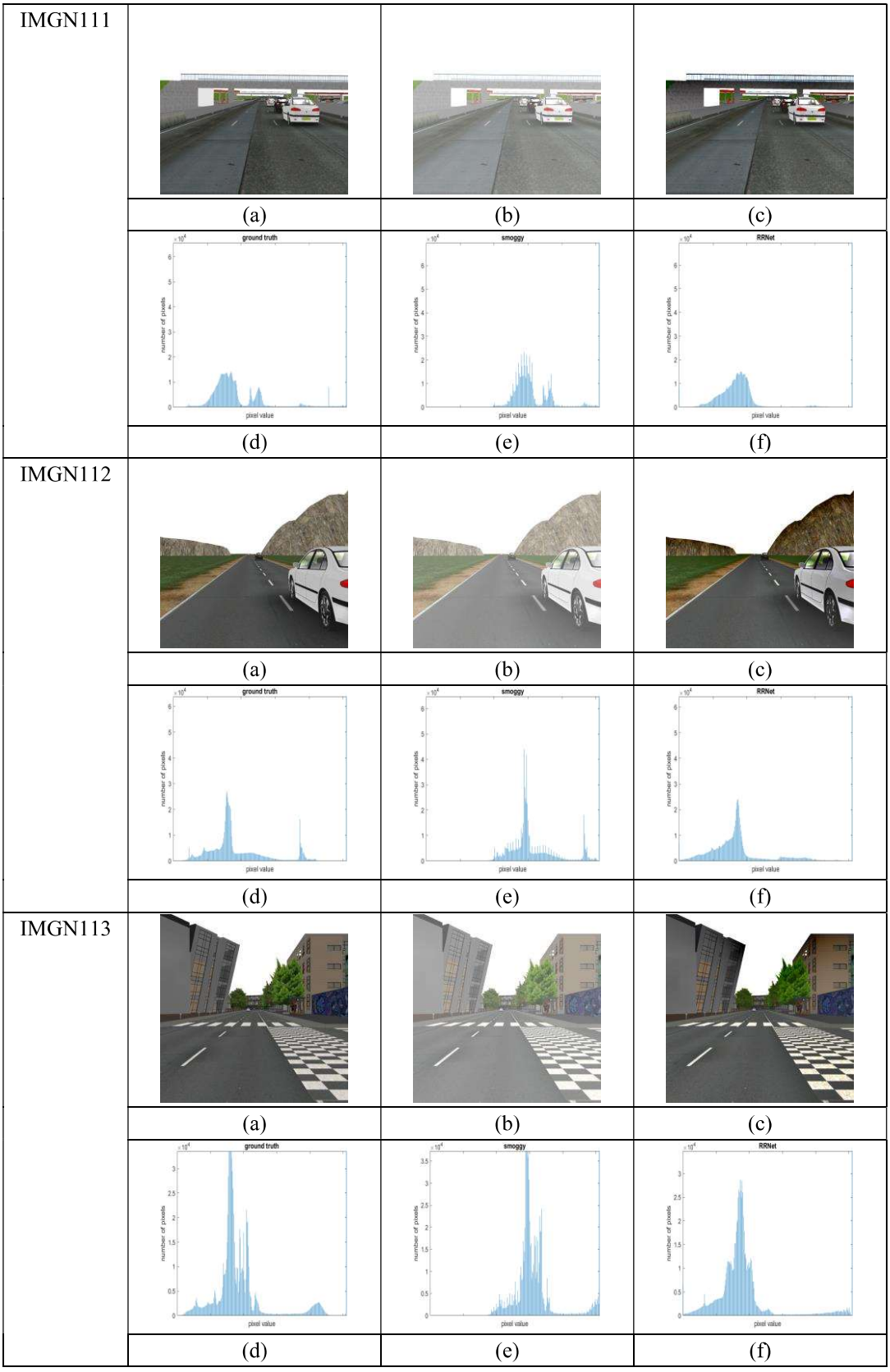


Fig. 7.4 Qualitative analysis of RRNet by using histogram plot. Results of sample images of IMSET-25 from FRIDA2 (a) ground truth; (b) smoggy image; (c) desmogged image using RRNet; (d) histogram plot of (a); (e) histogram plot of (b); (f) histogram plot of (c)

depicted in Figure 7.6. It has been observed that results obtained from RC do not have naturalness. The images computed using IGBF and ICP suffer from color distortion and a decrease in pixels' intensity. The smoggy images restored using VM are partially desmogged. Also, the sky region of computed images has color distortion. Thus, it has not been found suitable for dense smoggy images. The images obtained using SCP have distorted edges. The visual quality of images obtained using RRNet is better than those obtained using existing desmogging techniques. However, the computed image of IMGN117 suffers from the halo effect.

Figure 7.7 depicts a sample of real-world images (i.e., IMGN121 through IMGN128) desmogged using RRNet and other desmogging algorithms. It has been observed that images with less atmospheric smog namely, IMGN122, IMGN123, IMGN124, IMGN125, and IMGN127 give output images with no or minimum color distortion in the sky region. However, when the images with dense smog are processed using the proposed technique, the computed images suffer from blocking artifacts issue. It is because the dense layer of smog causes hinderance in the path from camera lens to the objects such as train in IMGN126 and buildings in IMGN128.

The comparative study of RRNet and other desmogging techniques for real-world smoggy images is depicted in Figure 7.8. IMSET-29 has been obtained from different literature works, such as IMGN129 and IMGN130 have been obtained from [118], IMGN131 through IMGN135 have been obtained from [122], IMGN136 and IMGN145 have been obtained [129], IMGN137 and IMGN146 have been obtained from [135], IMGN138 has been obtained from [134], IMGN139 through IMGN141 have been obtained from [164], and IMGN142 through IMGN144 have been obtained from [165].

It has been observed that the results generated by using IGBF and ICP have less pixel intensity and color distortion. Thus, they are not suitable for real-world smoggy images. Also, results obtained by SCP and RC suffer from edge distortion and texture distortion, respectively. Furthermore, VM has not been found suitable for images with large smog gradients and non-uniform distribution of smog. Eventually, the visual quality of results computed using RRNet was found to be better than those obtained from existing techniques. However, the computed image of IMGN137 using RRNet suffers from color distortion.

Fig. 7.5 Desmogging results of sample images of IMSET-26 from HUDRS (a) smoggy image, (b) estimated transmission map, (c) refined transmission map, (d) optimized transmission map, (e) desmogged image using RRNet, and (f) ground truth

Algorithm	IMGN117	IMGN118	IMGN119	IMGN120
Smoggy Image				
RC				
IGBF				
ICP				
VM				
SCP				
RRNet (Proposed Technique)				
Ground Truth				

Fig. 7.6 Desmogging results of sample images of IMSET-27 from RESIDE

Algorithm	IMGN121	IMGN122	IMGN123	IMGN124
Smoggy Image				
RC				
IGBF				
ICP				
VM				
SCP				
RRNet (Proposed Technique)				

Fig. 7.7 Comparative analysis of desmogging results of sample images of IMSET-28 (real-world images) available online

Algorithm	IMGN125	IMGN126	IMGN127	IMGN128
Smoggy Image				
RC				
IGBF				
ICP				
VM				
SCP				
RRNet (Proposed Technique)				

Fig. 7.7 (continue) Comparative analysis of desmogging results of sample images of IMSET-28 (real-world images) available online

Algorithm	IMGN129	IMGN130	IMGN131	IMGN132
Smoggy Image				
RC				
IGBF				
ICP				
VM				
SCP				
RRNet (Proposed Technique)				

Fig. 7.8 Desmogging results of sample images of IMSET-29 (real-world images). Smoggy images of IMGN129 and IMGN130 are obtained from [118]. Smoggy images of IMGN131 and IMGN132 are obtained from [122].

Algorithm	IMGN133	IMGN134	IMGN135	IMGN136
Smoggy Image				
RC				
IGBF				
ICP				
VM				
SCP				
RRNet (Proposed Technique)				

Fig. 7.8 (continue) Desmogging results of sample images of IMSET-29 (real-world images). Smoggy images of IMGN133 through IMGN135 are obtained from [122]. Smoggy image of IMGN136 is obtained from [129].

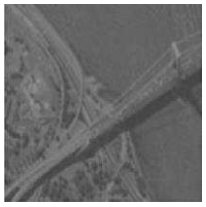
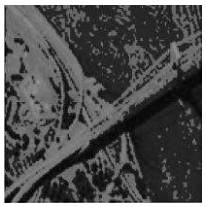
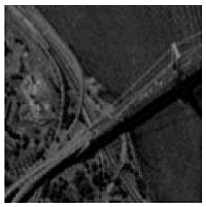
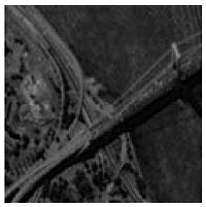
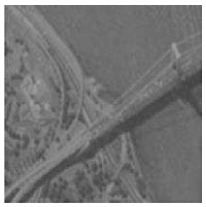
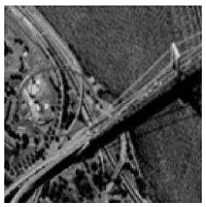
Algorithm	IMGN137	IMGN138	IMGN139	IMGN140
Smoggy Image				
RC				
IGBF				
ICP				
VM				
SCP				
RRNet (Proposed Technique)				

Fig. 7.8 (continue) Desmogging results of sample images of IMSET-29 (real-world images). Smoggy image of IMGN137 is obtained from [135]. Smoggy image of IMGN138 is obtained from [134]. Smoggy image of IMGN139 and IMGN140 is obtained from [164].

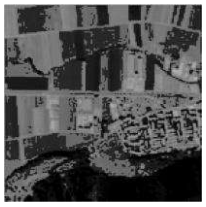
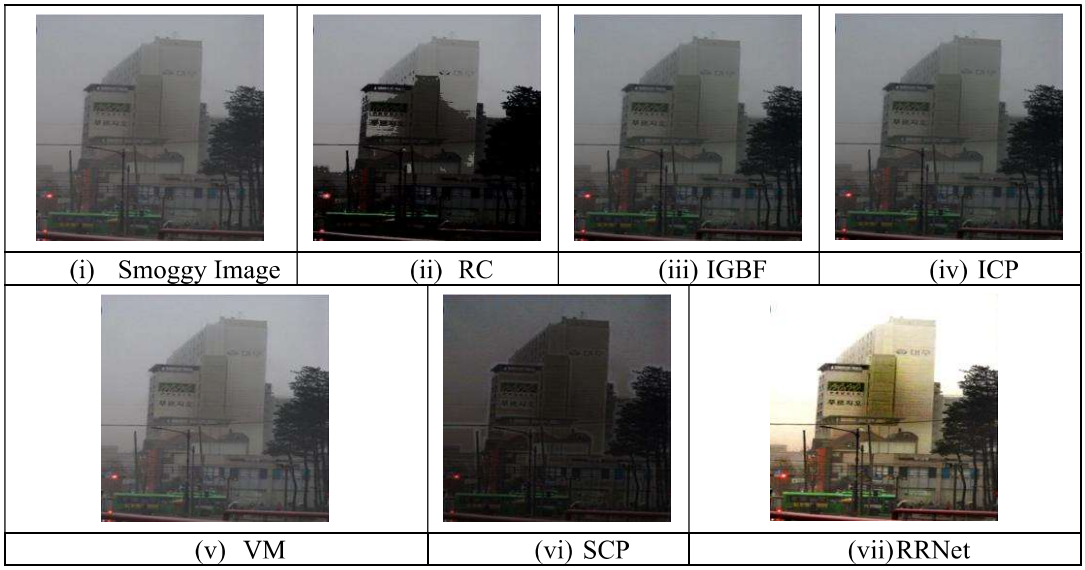
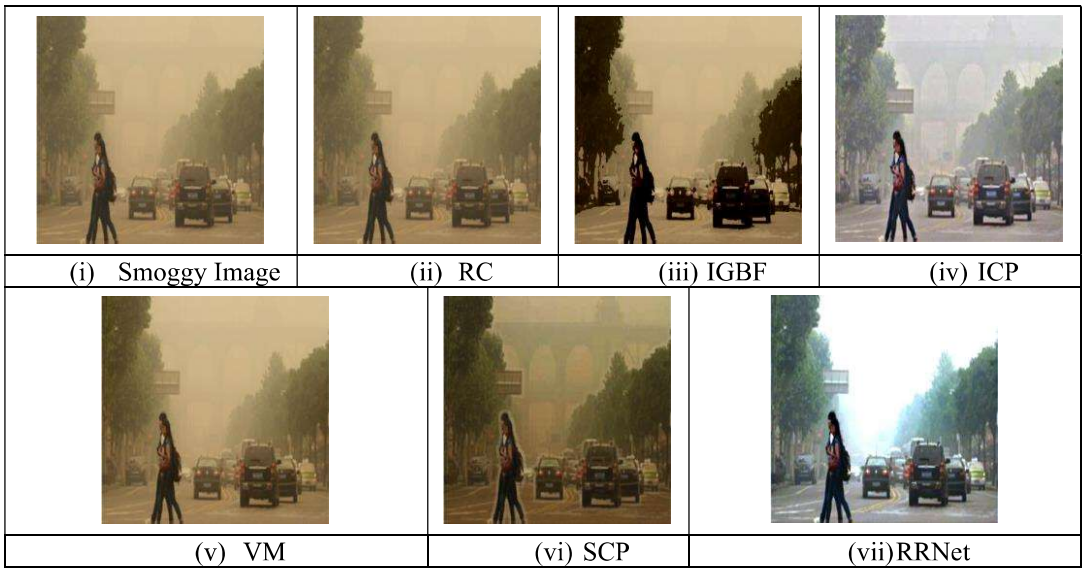
Algorithm	IMGN141	IMGN142	IMGN143	IMGN144
Smoggy Image				
RC				
IGBF				
ICP				
VM				
SCP				
RRNet (Proposed Technique)				

Fig. 7.8 (continue) Desmogging results of sample images of IMSET-29 (real-world images). Smoggy image of IMGN141 is obtained from [164]. Smoggy image of IMGN142 through IMGN144 are obtained from [165].



(a) Desmogging results of IMGN145



(b) Desmogging results of IMGN146

Fig. 7.8 (continue) Desmogging results of sample images of IMSET-29 (real-world images). Smoggy image of IMGN145 is obtained from [129]. Smoggy image of IMGN146 is obtained from [135].

Table 7.2 Estimated atmospheric light of synthetic smoggy images

IMGN 114	IMGN 115	IMGN 116	IMGN 117	IMGN 118	IMGN 119	IMGN 120	IMGN 121	IMGN 122
0.8952	0.9961	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

The estimation of atmospheric light from the given image has been performed by utilizing a 2D order statistic filter. Table 7.2 depicts the normalized values of atmospheric light used for obtaining an optimized transmission map from IMGN114 to IMGN122. It has been observed that most of the images have a unit value of atmospheric

light. Due to this reason, the sky region of the restored images suffers from color distortion and increased brightness.

The color deviation of images IMGN108 through IMGN122 computed using RRNet has been calculated by implementing the CIE (Commission Internationale de l'Éclairage) color-difference formula [103]. A low value of color deviation is desired. The color difference between some ground truth images and their respective desmogged images was computed by implementing the proposed technique as presented in Table 7.3. The obtained values of color deviation in the majority of the images are less than 20. The color deviation of IMGN111, IMGN112, and IMGN127 is zero when RRNet has been implemented.

The effectiveness of RRNet on real-world images has been calculated using fog aware density evaluator ($FADE$) [83], [111] – [113]. It depends on different factors such as color saturation, contrast energy, image entropy, and sharpness of the image instead of geographical position of camera, depth information, and human-rated judgements. A low value of fog density is desired. If the computed value is zero, then the image is fog-free. However, it only gives information about the fog component of the smoggy images. The fog densities of images from IMSET-24 through IMSET-26 are presented in Tables 7.4 – 7.6. It has been observed that $FADE$ values of IMGN108 through IMGN114 are less than 1. However, $FADE$ values of IMGN115 and IMGN116 are 1.07 and 1.11, respectively. Figure 7.9 depicts the fog densities of sample images of IMSET-26 namely, IMGN132, IMGN133, IMGN134, IMGN135, IMGN136, IMGN145, and IMGN146 obtained using RRNet.

Table 7.3 Qualitative analysis of RRNet based on color fidelity by using CIE

Datasets	**Images**		
FRIDA	**IMGN108**	**IMGN109**	**IMGN110**
	6.3643	12.6455	11.4605
FRIDA2	**IMGN111**	**IMGN112**	**IMGN113**
	0	0	1.9818
HUDRS	**IMGN114**	**IMGN115**	**IMGN116**
	30.6443	12.4319	7.078
RESIDE	**IMGN117**	**IMGN118**	**IMGN119**
	5.3058	0	18.0975
RESIDE (continue)	**IMGN120**	**IMGN121**	**IMGN122**
	19.3669	7.554	6.2839

These values have been compared with the results obtained from existing state-of-the-art desmogging techniques. It has been observed that IMGN134 has lowest fog density of 0.2181 when the proposed approach is implemented as compared to the existing desmogging techniques. Also, the best value of $FADE$ of IMGN145 obtained by using RC is 0.6115. It is followed by RRNet with value of 0.6365. Furthermore, RRNet has been ranked third best among other existing techniques, such that values are evaluated for IMGN132, IMGN133, IMGN135, and IMGN136 are 0.2386, 0.1939, 0.306, and 0.5395, respectively.

7.3.3 Quantitative Analysis

RRNet has been analyzed through various performance metrics namely, $PIQE$, $PSNR$, IE, $SSIM$, $NIQE$, and $BRISQUE$. The performance comparison of the proposed architecture has been made with RC, IGBF, ICP, VM, and SCP. The evaluated results are depicted in Figures 7.6 – 7.11 and Tables 7.7 – 7.9. $SSIM$ is computed only when ground truth or clear image is available. Tables 7.4 – 7.6 depict the $SSIM$ values obtained for the proposed technique for IMGN108 to IMGN116. The desired value of $SSIM$ is closer to 1. It has been observed that $SSIM$ is greater than 0.7 for all the images computed using the proposed approach except IMGN109 and IMGN110. $SSIM$ of IMGN111, IMGN112, IMGN113, and IMGN116 are greater than 0.9. It proves the effectiveness of RRNet. The $SSIM$ of IMGN108, IMGN114, and IMGN117 are in the range of 0.7-0.8.

Tables 7.4 – 7.6 depict the values of $PSNR$ obtained using the proposed approach. It has been observed that $PSNR$ obtained for IMSET-24 to IMSET-26 are in the range 15-25 except IMGN114. $PSNR$ of IMGN114 is very small, indicating that there has been a lot of noise in the computed image. $PSNR$ of IMGN108, IMGN109, IMGN110, IMGN111, IMGN112, IMGN113, IMGN116, and IMGN121 is found greater than 20.

If the quality score is estimated using block-wise distortion estimation, then the resultant parameter is $PIQE$. A smaller value of $PIQE$ is preferred. The value of $PIQE$ is considerably high for sample images because density of smog is high. Tables 7.4 – 7.6 show the values of $PIQE$ obtained from the proposed approach. It has been observed that the values of $PIQE$ obtained using RRNet are more than 50 for IMGN116 only, out of 9 images (IMGN108 to IMGN116). Figure 7.10 depicts a graphical representation of comparative analysis of the proposed approach with the existing techniques for sample

images of IMSET-27 namely, IMGN132, IMGN133, IMGN135, IMGN136, IMGN138, IMGN141, and IMGN144. It has been observed that $PIQE$ of IMGN138 obtained by implementing RRNet is smaller with a value of 59.1757 as compared to other desmogging techniques. Also, the least values of $PIQE$ for IMGN132, IMGN135, IMGN136, IMGN141, and IMGN144 obtained by using SCP are 57.1944, 46.9746, 59.5455, 62.3829, and 38.0363, respectively, followed by RRNet as 63.6176, 55.7962, 62.0418, 66.2811, and 42.2886, respectively.

Tables 7.4 – 7.6 show the values of $NIQE$ obtained from the proposed approach. It has been observed that the values of $NIQE$ obtained from using RRNet are more than 4.5 for IMGN111, IMGN112, and IMGN114 only, out of 9 images (IMGN108 to IMGN116). Table 7.7 compares sample images from IMSET-27 to IMSET-29 namely, IMGN117, IMGN118, IMGN120, IMGN123, IMGN129, IMGN130, IMGN133, IMGN134, IMGN135, IMGN136, IMGN137, IMGN138, IMGN139, IMGN142, IMGN145, and IMGN146, using RRNet and other desmogging techniques (values with '*' and '#' indicate first and second-best performance, respectively). The values of $NIQE$ for IMGN117, IMGN118, IMGN130, IMGN138, IMGN145, and IMGN146 are 2.62, 2.82, 3.3582, 4.0819, 3.6375, and 4.4314, respectively, which is least when RRNet is used. Also, the values of $NIQE$ for IMGN120, IMGN123, IMGN129, IMGN134, and IMGN136 are 3.25, 2.15, 4.0873, 7.0522, and 4.0841, respectively, which is second lowest when RRNet is used.

Table 7.4 Performance metrics of sample images from IMSET-24 by using RRNet

Images	*SSIM*	*PSNR*	*PIQE*	*NIQE*	*IE*	*BRISQUE*	*FADE*
IMGN108	0.7	21.39	33.54	3.76	5.97	40.19	0.45
IMGN109	0.64	21.03	38.54	3.3	5.55	43.97	0.69
IMGN110	0.66	20.17	34.8	3.13	6.11	39.84	0.48

Table 7.5 Performance metrics of sample images from IMSET-25 by using RRNet

Images	*SSIM*	*PSNR*	*PIQE*	*NIQE*	*IE*	*BRISQUE*	*FADE*
IMGN111	0.96	21.73	31.72	7.74	4.84	38.06	0.75
IMGN112	0.95	21.81	28.99	4.81	5.28	33.21	0.69
IMGN113	0.96	24.65	38.92	2.5	6.13	40.08	0.82

Table 7.6 Performance metrics of sample images from IMSET-26 by using RRNet

Images	*SSIM*	*PSNR*	*PIQE*	*NIQE*	*IE*	*BRISQUE*	*FADE*
IMGN114	0.71	11.95	42.83	6.16	4.93	38.25	0.99
IMGN115	0.82	19.99	46.85	3.98	7	40.42	1.07
IMGN116	0.93	20.16	55.69	4.02	7.04	41.48	1.11

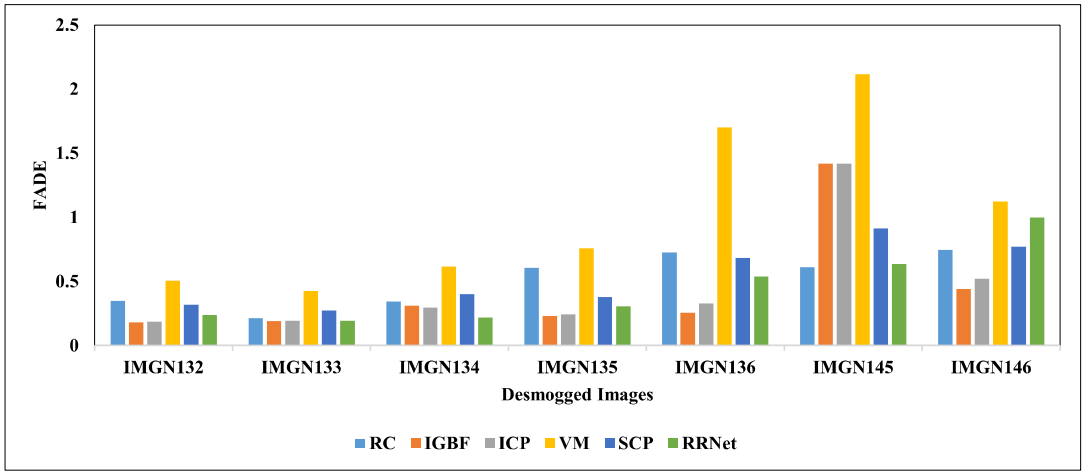


Fig. 7.9 Performance comparison of RRNet and the existing techniques based on $FADE$

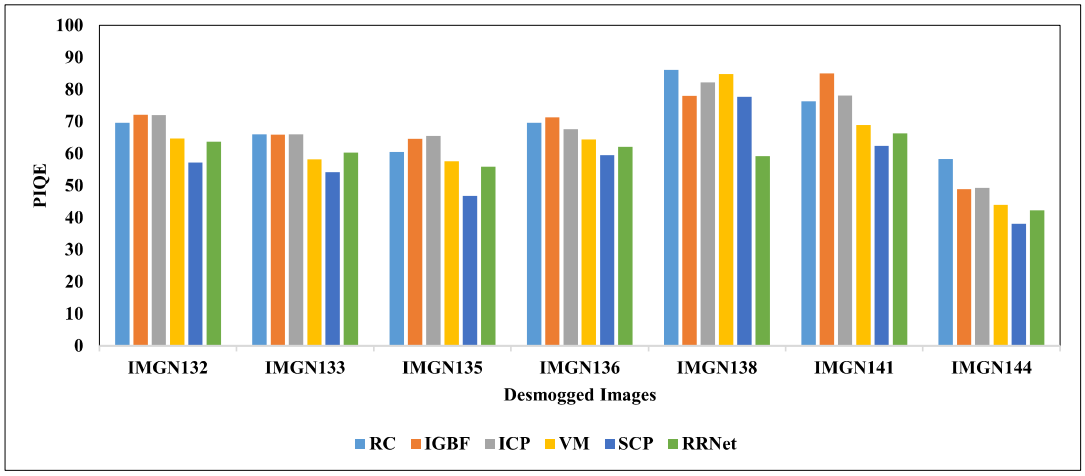


Fig. 7.10 Performance comparison of RRNet and the existing techniques based on $PIQE$

A higher value of image entropy is preferred. Tables 7.4 – 7.6 show the values of IE obtained from the proposed approach. It has been observed that the values of IE obtained using RRNet are less than 5 for IMGN111 and IMGN114 only, out of 9 images (IMGN108 to IMGN116).

Figure 7.11 depicts a graphical representation of the proposed approach compared to existing techniques for sample images of IMSET-29 namely, IMGN129, IMGN130, IMGN133, IMGN136, IMGN138, IMGN141, and IMGN143. It has been observed that IE obtained using RRNet is highest for IMGN129, IMGN130, IMGN141, and IMGN143, which is 7.5606, 7.5195, 7.3854, and 7.4464, respectively. Also, highest IE obtained for IMGN136 is 7.2774, followed by 6.9496, when RRNet is implemented.

Tables 7.4 – 7.6 show the values of $BRISQUE$ obtained from the proposed approach. It has been observed that the values of $BRISQUE$ obtained from RRNet for IMGN108 to IMGN116 are less than 44. Table 7.8 compares sample images from IMSET-27 to IMSET-29 namely, IMGN117, IMGN120, IMGN121, IMGN124, IMGN125, IMGN126, IMGN127, IMGN128, IMGN129, IMGN130, IMGN131, IMGN132, IMGN133, IMGN137, IMGN138, IMGN139, IMGN143, IMGN144, IMGN145, and IMGN146, using RRNet and other desmogging techniques (the values with '*' and '#' indicate first and second-best performance, respectively).

The values of $BRISQUE$ for IMGN121, IMGN130, IMGN133, and IMGN138 are 18.7, 30.2639, 37.7957, and 44.8793, respectively, which is least when RRNet is used. Also, the values of $NIQE$ for IMGN117, IMGN125, IMGN126, IMGN144, and IMGN146 are 29.23, 25.38, 33.06, 31.1, 40.7771, and 46.2775, respectively, which is second lowest when RRNet is used.

The computational complexity of a model is judged on two major factors namely, performance metrics and computation time. Time complexity depends on the experimental set-up, i.e., workstation and hyperparameters.

Table 7.7 Performance comparison of RRNet and existing techniques based on $NIQE$

Images	RC	IGBF	ICP	VM	SCP	RRNet
IMGN117	3.6923	4.2373	3.5541	**#2.8066**	2.8539	***2.62**
IMGN118	4.0508	3.8116	**#3.543**	3.6288	3.8755	***2.81**
IMGN120	3.3055	3.3465	3.3458	3.617	***3.1824**	**#3.25**
IMGN123	4.2621	3.0976	2.7248	***2.0229**	2.3245	**#2.15**
IMGN129	4.3181	4.2035	***4.054**	4.2328	4.6159	**#4.0873**
IMGN130	4.4942	3.9448	3.5423	3.9194	3.3803	***3.3582**
IMGN133	***6.6171**	7.4581	**#7.4336**	10.0136	8.1966	8.4356
IMGN134	7.326	***6.7832**	7.3895	9.5533	8.9986	**#7.0522**
IMGN135	8.4645	***7.8622**	**#8.001**	9.4959	8.2632	9.5258
IMGN136	***4.079**	4.4899	4.1837	4.5785	4.3461	**#4.0841**
IMGN137	**#5.3161**	***4.5516**	5.5079	5.6694	6.2668	5.9375
IMGN138	5.2292	5.486	**#5.1159**	5.4659	5.4251	***4.0819**
IMGN139	***6.6833**	9.0017	**#7.1721**	8.6784	7.6609	10.6955
IMGN142	4.8808	4.5318	***4.1731**	5.4801	5.8679	5.5554
IMGN145	**#3.9814**	4.4066	4.4066	4.3223	4.78	***3.6375**
IMGN146	**#4.6262**	5.3941	4.9876	4.9001	5.4322	***4.4314**

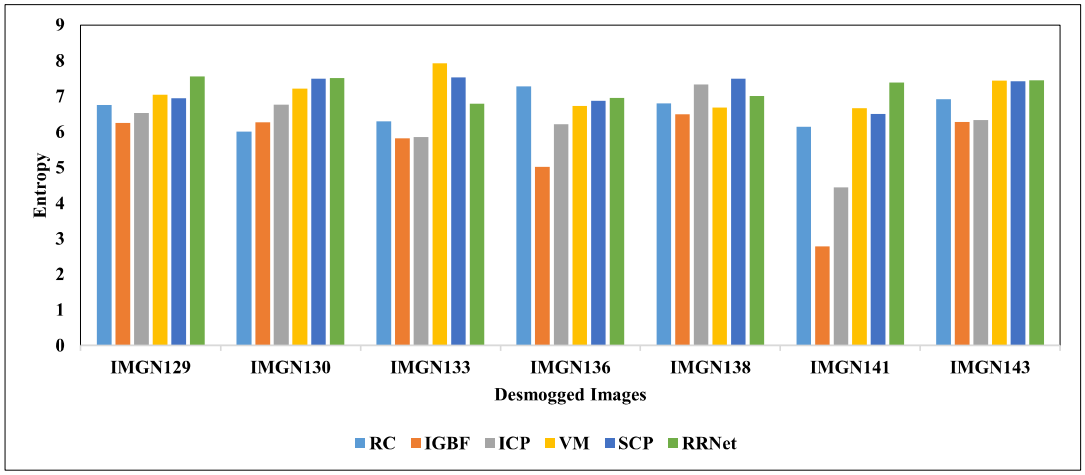


Fig. 7.11 Performance comparison of RRNet and the existing techniques based on *IE*

Table 7.8 Performance comparison of RRNet and existing techniques based on *BRISQUE*

Images	RC	IGBF	ICP	VM	SCP	RRNet
IMGN117	39.5687	36.0067	36.0635	37.9543	***19.1231**	**#29.23**
IMGN120	***11.3165**	30.5341	28.9523	**#14.2047**	21.2146	22.93
IMGN121	29.522	35.5125	**#22.7917**	36.3343	32.3393	***18.7**
IMGN124	26.7956	16.8511	**#16.8467**	26.8715	***15.5529**	33.13
IMGN125	32.5744	34.0026	31.8017	26.8715	***19.8341**	**#25.38**
IMGN126	48.2346	55.2812	52.2316	***25.8884**	34.0074	**#33.06**
IMGN127	32.9742	35.0907	33.5205	30.353	***21.6246**	**#31.1**
IMGN128	53.2551	38.5666	39.066	***28.314**	**#37.5571**	44.64
IMGN129	***35.389**	**#40.9632**	42.7825	45.1149	42.2516	42.0362
IMGN130	54.8201	51.4721	45.2049	43.5048	**#40.3044**	***30.2639**
IMGN131	***35.0358**	**#42.068**	44.3033	47.39	43.457	42.3116
IMGN132	***38.124**	45.5213	44.0738	**#43.4025**	43.4582	43.4582
IMGN133	41.4107	41.004	**#39.75**	43.4582	43.4582	***37.7957**
IMGN137	**#40.2055**	***38.9709**	42.0432	42.3709	42.553	42.4702
IMGN138	49.5275	**#49.182**	53.5611	50.6884	52.2108	***44.8793**
IMGN139	***37.2042**	48.7447	44.9089	**#43.4577**	43.4582	45.0789
IMGN140	43.4582	***34.384**	**#34.5041**	43.0814	42.5224	42.4735
IMGN144	***39.2597**	50.4171	42.1403	44.7112	41.1613	**#40.7771**
IMGN145	46.4208	**#43.2563**	**#43.2563**	***43.0912**	47.5414	51.2534
IMGN146	***44.0276**	51.0445	52.97	50.0412	49.1056	**#46.2775**

Since the learnable parameters are less in each layer of the proposed architecture, the computational time is less than one second for all images. Table 7.9 depicts the computation time of various sample images of IMSET-24 through IMSET-27 computed using RRNet. It has been observed that the proposed approach performed better than other existing visibility restoration techniques for 36 out of 57 times based on $PIQE$, $NIQE$, IE, and $BRISQUE$.

7.4 Ablation Studies

Several ablation experiments have been carried out to validate the contribution of each stage in the proposed model. In these experiments, different components were removed in each training. It mainly includes the following steps:

(1) The original image has been restored without extraction of shallow features in Stage-I.

(2) The original image has been restored without reducing the thickness of smog in Stage-II.

(3) The original image has been restored without rescaling feature maps in Stage-III.

(4) The original image has been restored without minimizing the effect of vanishing gradient in Stage-IV.

(5) The original image has been restored without reducing the risk of overfitting in Stage-V.

(6) The original image has been restored using all the stages.

The detailed ablation results have been obtained for IMSET-30 (i.e., IMGN108, IMGN145, and IMGN146). The qualitative analysis is presented in Figure 7.12 and Table 7.10. Also, $SSIM$, $PSNR$, $PIQE$, and IE of IMGN108 have been obtained for RRNet by excluding different stages. It has been observed that all algorithms are positive for visibility restoration in dense smog. Furthermore, the restored image suffers from color distortion and gradient artifact when Stage-III is included.

Table 7.9 Computation time of sample images (in seconds) computed by using proposed desmogging technique

IMGN 108	IMGN 109	IMGN 111	IMGN 112	IMGN 114	IMGN 115	IMGN 117	IMGN 119	IMGN 121
0.64153	0.64381	0.63110	0.64709	0.59661	0.55691	0.63910	0.60775	0.57873

Algorithm	IMGN108	IMGN145	IMGN146
Smoggy Image			
Stage-II + Stage-III + Stage-IV + Stage-V			
Stage-I + Stage-III + Stage-IV + Stage-V			
Stage-I + Stage-II + Stage-IV + Stage-V			
Stage-I + Stage-II + Stage-III + Stage-V			
Stage-I + Stage-II + Stage-III + Stage-IV			
Stage-I + Stage-II + Stage-III + Stage-IV + Stage-V			
Ground Truth		-	-

Fig. 7.12 Qualitative analysis of ablation results of RRNet

Table 7.10 Ablation studies on IMGN145 and IMGN146

Algorithm	Visibility Restoration in Dense Smog	Color Distortion	Gradient Artifact	Naturalness
Stage-II + Stage-III + Stage-IV + Stage-V	✓	✓	✓	✗
Stage-I + Stage-III + Stage-IV + Stage-V	✓	✓	✓	✗
Stage-I + Stage-II + Stage-IV + Stage-V	✓	✗	✗	✓
Stage-I + Stage-II + Stage-III + Stage-V	✓	✓	✓	✗
Stage-I + Stage-II + Stage-III + Stage-IV	✓	✓	✓	✗
Stage-I + Stage-II + Stage-III + Stage-IV + Stage-V	✓	✗	✓	✓

Table 7.11 Quantitative analysis of ablation results of IMGN108

Algorithm	SSIM	PSNR	$PIQE$	H
Stage-II + Stage-III + Stage-IV + Stage-V	0.6938	17.1533	34.0542	5.8308
Stage-I + Stage-III + Stage-IV + Stage-V	0.6981	18.1301	**33.0171**	5.865
Stage-I + Stage-II + Stage-IV + Stage-V	0.6756	15.2843	32.7879	5.8154
Stage-I + Stage-II + Stage-III + Stage-V	**0.703**	18.6577	33.0379	5.8875
Stage-I + Stage-II + Stage-III + Stage-IV	0.6987	17.9234	33.5559	5.8657
Stage-I + Stage-II + Stage-III + Stage-IV + Stage-V	0.7	**21.39**	33.54	**5.97**

As depicted in Table 7.11, the best value of $PSNR$ and IE is obtained when all the stages are utilized together. Also, when Stage-IV has not been used, $SSIM$ is nearest to 1. Although perception is least when thickness of smog has is reduced, naturalness is absent in the computed image.

7.5 Summary

In this chapter, a light residual regression network (RRNet) followed by morphological erosion has been proposed to obtain desmogged images. A residual regression network has been proposed to estimate the transmission map. The transmission map has been refined using a morphological operation called erosion. The atmospheric light has been estimated using a 2D order statistic filter for edge smoothing. RRNet has outperformed the existing desmogging techniques in terms of $PIQE$, $BRISQUE$, and $FADE$, by 8.05%, 2.64%, and 2.02%, respectively.

CHAPTER-8
CONCLUSION AND FUTURE SCOPE

The conclusion and future scope are presented in this chapter.

8.1 Conclusion

The various types of foggy image and video databases have been analyzed with their pros and cons. A new database Hazy Unpaired Dataset for Road Safety (HUDRS) has been developed in this work. It comprises 522 fog-free images and 1050 foggy images with different intensities. Foggy and fog-free images have been captured during the morning/evening and the day, respectively. The Canon Power Shot SX400 IS camera has been manually adjusted to capture real-time roadside images. The resolution of these images is 4608×3456. Each clear image has processed versions with synthetic haze of different intensities, producing 2088 synthetic hazy images. After the dataset acquisition process, the performance of existing prior-based and deep learning-based techniques have been studied. Moreover, one metaheuristic technique has also been evaluated. In addition to that, three deep learning-based and one prior-based algorithms have been proposed for visibility restoration. The hazy images have been restored by obtaining transmission map and airlight.

A metaheuristic technique called grey wolf optimization technique has been applied for image dehazing. In this algorithm, a random population of wolves and their respective positions have been initialized. The canny edge filter has been used to evaluate the fitness function and update the positions of the wolves, which resulted in the reduction of texture and edge distortions in the computed images. It has been found superior than existing dehazing techniques in terms of $PSNR$ and $SSIM$ by 2.9% and 0.43%, respectively.

The first deep learning-based model called hybrid CNN uses retinex, wavelet, and inverted wavelet filters to process hazy images. The wavelet and inverted filters have been used to enhance the features of non-sky and sky regions, respectively. The sky and non-sky features have been separately processed to obtain the coarse and fine transmission maps, respectively. The airlight has been estimated by using the statistic order filter. The optimized transmission map has been produced using the combined maps and estimated airlight. It has been found superior to the existing dehazing

techniques in terms of fog aware density evaluator ($FADE$) and naturalness image quality evaluator ($NIQE$), by 1.71%, and 0.89%, respectively.

A novel video dehazing model called Aethra-Net consisting of gush enhancer-based autoencoder and a vessel enhancement-based filter has been proposed. The transmission map has been estimated using a gush enhancement-based autoencoder. The low, medium, and high-level feature maps have been obtained using maximum pooling, average pooling, and global average pooling operations, respectively. Further, multiple blocks of ResNet-101 have been used to minimize vanishing gradient issue. The tubular structures have been emphasized using the vessel enhancement-based filter. The significant features have been combined from both feature maps. The airlight has been updated for each frame using mean filter. The performance analysis demonstrates that Aethra-Net outperforms the existing techniques in terms of $BRISQUE$, $SSIM$, and $PSNR$ by 15.29%, 0.95%, and 0.62%, respectively. It has achieved the average dehazing time of 0.11 seconds along with updated airlight for each frame.

A novel bounding function for gray-world kernel prior algorithm (BFGKP) has been proposed to obtain the clear images from real-time hazy images. The transmission map has been estimated using the gray-world kernel prior using the bounding function. The hazy image has been reconstructed by using the gamma factor to control the reflected light. The proposed algorithm minimizes the edge distortion and eliminates the flickering effect. BFGKP outperformed the existing algorithms in terms of $NIQE$ and $BRISQUE$ by 2.06% and 0.11%, respectively. Also, $FADE$ has been found better for 92.59% of test images. A hardware setup has been presented to accomplish real-time video dehazing using BFGKP. The framework consists of a GoPro Hero 7 camera to capture the image, a display screen or a system to view the output, a 4k Camlink to transfer the frames captured by GoPro in 4k resolution, and a bluerigger chord to set up connectivity. Further, BFGKP has been used to obtain results for real-time dehazing application using MATLAB Mobile application designed by MathWorks, Inc. The computation time of the proposed model is 7.68 times less than the existing techniques, i.e., dark channel prior, color attenuation prior, improved color attenuation prior, optimized transmission map estimation, non-linear bounding function, bounded channel difference prior, image dehazing and exposure, and color cast dependent technique.

An alternative approach uses a regression-based deep learning model followed by morphological erosion to obtain clear images from the smoggy (smoke + foggy) images.

The proposed architecture gave the estimated transmission map using a residual regression neural network (RRNet). It consists of five stages. Stage I extracts the shallow features which has been followed by reduction in the thickness of smog to find the depth of image in Stage II. Stage III makes the network stable and faster by using batch-normalization. The features enhancement and halo effect reduction has been performed in Stage IV. In Stage V, a hyperbolic tangent function has been used as activation function to reduce the risk of overfitting. The transmission map has been refined by using morphological erosion operations. The atmospheric light has been estimated using an order statistic filter for the smoggy image. The smoggy image has been reconstructed by using the gamma factor to control the exposure of light in the image. The performance analysis demonstrates that RRNet outperforms the existing smoggy techniques in terms of $PIQE$, $BRISQUE$, $FADE$, and IE, by 8.05%, 2.64%, 2.2%, and 0.69%, respectively.

8.2 Future Scope

1) In future, HUDRS can be analyzed by researchers to propose a novel dehazing model. Also, this dataset can be expanded in future by including real-time video frames to evaluate the accuracy of the developed dehazing architectures.
2) The performance of hybrid CNN, Aethra-Net, and RRNet may improve by considering further hyperparameter tuning. They can be improved by estimating atmospheric light for images with non-homogeneous and dense haze.
3) A fusion-based desmogging model may be considered by using transform-based desmogging approach. Also, an end-to-end architecture can be designed to eliminate the need of airlight estimation and reduce the computation time.
4) The developed prototype defogging model can help to reduce the frequency of accidents during the winter season. During foggy days, media can monitor the real-time traffic in the streets and alert people to avoid the accident-prone areas. It has the potential to be a game changer in the realm of road safety.

LIST OF PUBLICATIONS

A. Journal Articles:

- **Published**

1. A. Juneja, V. Kumar, S. K. Singla, "A Systematic Review on Foggy Datasets: Applications and Challenges," Archives of Computational Methods in Engineering, vol. 29, no. 3, pp. 1727-1752, 2022. **(Springer)**

   URL: https://doi.org/10.1007/s11831-021-09637-z. **(I.F. = 9.7)**

2. A. Juneja, S. K. Singla, V. Kumar, "HUDRS: hazy unpaired dataset for road safety," The Visual Computer, vol. 39, no. 9, pp. 3905-3922, 2023. **(Springer)**

   URL: https://doi.org/10.1007/s00371-022-02534-x. **(I.F. = 3.5)**

3. A. Juneja, V. Kumar, S. K. Singla, "A Comprehensive Review of Computational Desmogging Techniques," Archives of Computational Methods in Engineering, vol. 30, pp. 3723-3748, 2023. **(Springer)**

   URL: https://doi.org/10.1007/s11831-023-09918-9. **(I.F. = 9.7)**

4. A. Juneja, V. Kumar, S. K. Singla, "Desmogging of still images using residual regression network and morphological erosion," Multimedia Tools and Applications, 83(3), pp.7179-7214, 2023. **(Springer)**

   URL: https://doi.org/10.1007/s11042-023-15893-x. **(I.F. = 3.6)**

5. A. Juneja, V. Kumar, S. K. Singla, "Aethra-net: Single image and video dehazing using autoencoder," Journal of Visual Communication and Image Representation, vol. 94, p. 103855, 2023. **(Elsevier)**

   URL: https://doi.org/10.1016/j.jvcir.2023.103855. **(I.F. = 2.6)**

6. A. Juneja, V. Kumar, S. K. Singla, "Single Image Dehazing Using Hybrid Convolution Neural Network," Multimedia Tools and Applications, 83, pp.38355-38386, 2023. **(Springer)**

   URL: https://doi.org/10.1007/s11042-023-17132-9 **(I.F. = 3.6)**

- **Communicated**

1. A. Juneja, V. Kumar, S. K. Singla, "Visibility Enhancement Techniques for Hazy Images," Engineering Applications of Artificial Intelligence. **(Elsevier)**

2. A. Juneja, V. Kumar, S. K. Singla, "A Novel Dehazing Framework for Road Accidents Prevention," Signal, Image and Video Processing. **(Springer)**
3. A. Juneja, V. Kumar, S. K. Singla, "BFGKP: Fast Image Dehazing for Road Safety" Journal of Visual Communication and Image Representation. **(Elsevier)**

B. Book Chapter:

1. A. Juneja, V. Kumar, S. K. Singla, "Fast Image Desmogging for Road Safety Using Optimized Dark Channel Prior," AI for Big Data-Based Engineering Applications from Security Perspectives, pp. 107-128, CRC Press, 2023. **(Taylor and Francis)**

   URL: https://doi.org/10.1201/9781003230113-5.

C. Conference

1. A. Juneja, V. Kumar, S. K. Singla, "Single Image Dehazing Using Grey Wolf Optimization," Congress on Smart Computing Technologies (CSCT2023). (**Springer**) (**Accepted**)
2. A. Juneja, V. Kumar, S. K. Singla, "Partition Explainer: Next Generation in Single Image Dehazing," 2024 IEEE International Conference on Contemporary Computing and Communications (INC4). (**IEEE**) (**Accepted**)

REFERENCES

[1] R. D. Singh, A. Mittal, and R. K. Bhatia, "3D convolutional neural network for object recognition: a review," *Multimed Tools Appl*, vol. 78, no. 12, pp. 15951–15995, Jun. 2019, doi: 10.1007/s11042-018-6912-6.

[2] A. K. Dwivedi, A. K. Singh, D. Singh, and H. Kumar, "Development of an Adaptive Linear Mixture Model for Decomposition of Mixed Pixels to Improve Crop Area Estimation Using Artificial Neural Network," *IEEE Access*, vol. 11, pp. 5714–5723, 2023, doi: 10.1109/ACCESS.2023.3236665.

[3] S. Gupta and D. Singh, "Color Features Based Model for Land Cover Identification and Agriculture Monitoring with Satellite Images," in *Remote Sensing of Agriculture and Land Cover/Land Use Changes in South and Southeast Asian Countries*, Cham: Springer International Publishing, 2022, pp. 587–605. doi: 10.1007/978-3-030-92365-5_34.

[4] A. Kumar, S. K. Singh, S. Saxena, A. K. Singh, and S. Shrivastava, "Recent trends in histopathological image analysis," *Medical Information Processing and Security: Techniques and Applications*, p. 27, 2022.

[5] A. Kumar, V. Purohit, V. Bharti, R. Singh, and S. K. Singh, "MediSecFed: Private and Secure Medical Image Classification in the Presence of Malicious Clients," *IEEE Trans Industr Inform*, vol. 18, no. 8, pp. 5648–5657, Aug. 2022, doi: 10.1109/TII.2021.3138919.

[6] A. Sharma and S. K. Singh, "Early classification of multivariate data by learning optimal decision rules," *Multimed Tools Appl*, vol. 80, no. 28–29, pp. 35081–35104, Nov. 2021, doi: 10.1007/s11042-020-09366-8.

[7] P. He, K. Sun, and C. Zhu, "A Novel Image Encryption Algorithm Based on the Delayed Maps and Permutation-Confusion-Diffusion Architecture," *Security and Communication Networks*, vol. 2021, pp. 1–16, Mar. 2021, doi: 10.1155/2021/6679288.

[8] C. Chen, K. Sun, and Q. Xu, “A color image encryption algorithm based on 2D-CIMM chaotic map,” *China Communications*, vol. 17, no. 5, pp. 12–20, May 2020, doi: 10.23919/JCC.2020.05.002.

[9] G. J. Ansari, J. H. Shah, M. Yasmin, M. Sharif, and S. L. Fernandes, “A novel machine learning approach for scene text extraction,” *Future Generation Computer Systems*, vol. 87, pp. 328–340, Oct. 2018, doi: 10.1016/j.future.2018.04.074.

[10] J. Janai, F. Güney, A. Behl, and A. Geiger, “Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art,” *Foundations and Trends® in Computer Graphics and Vision*, vol. 12, no. 1–3, pp. 1–308, 2020, doi: 10.1561/0600000079.

[11] S.-C. Huang, B.-H. Chen, and Y.-J. Cheng, “An Efficient Visibility Enhancement Algorithm for Road Scenes Captured by Intelligent Transportation Systems,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2321–2332, Oct. 2014, doi: 10.1109/TITS.2014.2314696.

[12] H. Singh, A. Kumar, L. K. Balyan, and H. N. Lee, “Optimally sectioned and successively reconstructed histogram sub-equalization based gamma correction for satellite image enhancement,” *Multimed Tools Appl*, vol. 78, no. 14, pp. 20431–20463, Jul. 2019, doi: 10.1007/s11042-019-7383-0.

[13] Y. Zhou, W. Jing, J. Wang, G. Chen, R. Scherer, and R. Damaševičius, “MSAR-DefogNet: Lightweight cloud removal network for high resolution remote sensing images based on multi scale convolution,” *IET Image Process*, vol. 16, no. 3, pp. 659–668, Feb. 2022, doi: 10.1049/ipr2.12224.

[14] D. Singh, M. Kaur, M. Y. Jabarulla, V. Kumar, and H.-N. Lee, “Evolving Fusion-Based Visibility Restoration Model for Hazy Remote Sensing Images Using Dynamic Differential Evolution,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022, doi: 10.1109/TGRS.2022.3155765.

[15] R. Singh, A. K. Dubey, and R. Kapoor, “Improved Transmission Map for Dehazing of Natural Images,” 2021, pp. 339–347. doi: 10.1007/978-981-15-9927-9_34.

[16] Z. H. Arif *et al.*, “Comprehensive Review of Machine Learning (ML) in Image Defogging: Taxonomy of Concepts, Scenes, Feature Extraction, and Classification techniques,” *IET Image Process*, vol. 16, no. 2, pp. 289–310, Feb. 2022, doi: 10.1049/ipr2.12365.

[17] K. Athiappan *et al.*, “Identifying Influencing Factors of Road Accidents in Emerging Road Accident Blackspots,” *Advances in Civil Engineering*, vol. 2022, pp. 1–10, Oct. 2022, doi: 10.1155/2022/9474323.

[18] Anon., “ROAD ACCIDENTS IN INDIA,” Ministry of Road Transport and Highways. Accessed: Dec. 05, 2023. [Online]. Available: https://morth.nic.in/sites/default/files/RA_2021_Compressed.pdf

[19] D. Goel, R. Bhatia, and K. Bhatia, “Traffic Violations Prediction System on the Basis of Human Behaviour,” 2021, pp. 25–35. doi: 10.1007/978-981-16-1295-4_3.

[20] Anon., “Annual Report on Motor Vehicle Accident Fund & Compensation to Victims of Hit & Run Motor Accidents Scheme,” Ministry of Road Transport and Highways. Accessed: Dec. 05, 2023. [Online]. Available: https://morth.nic.in/sites/default/files/Annual-Report-FY-2022-23-MVL.pdf

[21] C. ’ ’Pro., “NHTSA: 8,730 People Died In The First Quarter Of 2021 In Vehicle Crashes,” CarProUSA. Accessed: Dec. 05, 2023. [Online]. Available: https://www.carprousa.com/blog/nhtsa-traffic-deaths-increase-in-first-quarter-of-2021

[22] F. Cozman and E. Krotkov, “Depth from scattering,” in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Comput. Soc, Jan. 1997, pp. 801–806. doi: 10.1109/CVPR.1997.609419.

[23] S. G. Narasimhan and S. K. Nayar, “Vision and the atmosphere,” in *ACM SIGGRAPH ASIA 2008 courses on - SIGGRAPH Asia ’08*, New York, New York, USA: ACM Press, 2008, pp. 1–22. doi: 10.1145/1508044.1508113.

[24] Useropwp2, “Ensure Clarity of Automotive Windshields With Haze Measurement,” Konica Minolta Color, Light, and Display Measuring Instruments. Accessed: Dec. 05, 2023. [Online]. Available:

https://sensing.konicaminolta.asia/ensure-clarity-of-automotivewindshields-with-haze-measurement/

[25] Anon., “Ecuador’s Sangay volcano erupts ,” DW Made for Minds. Accessed: Dec. 05, 2023. [Online]. Available: https://www.dw.com/en/ecuadors-sangay-volcano-erupts-spews-ash-8-kilometers/a-56845388

[26] Anon., “Driving Through Fog & Smoke: Safety Rules for Limited and Zero Visibility,” ePermitTest. Accessed: Dec. 05, 2023. [Online]. Available: https://www.epermittest.com/drivers-education/driving-through-fog-smoke

[27] L. Eliot, “Mighty Dust Storm Drifting Toward Southeastern U.S., Even Self-Driving Cars To Be Flummoxed,” Forbes. Accessed: Dec. 05, 2023. [Online]. Available: https://www.forbes.com/sites/lanceeliot/2020/06/24/mighty-dust-storm-drifting-toward-southeastern-us-even-self-driving-cars-to-be-flummoxed/?sh=5526410e1beb

[28] Anon., “About Us,” NewsroomPost. Accessed: Dec. 05, 2023. [Online]. Available: https://newsroompost.com/lifestyle/driving-in-smog-important-safety-tips-to-remember/355677.html

[29] W. GmbH, “Cars driving on road at morning mist stock photo,” Westend61. Accessed: Dec. 05, 2023. [Online]. Available: https://www.westend61.de/en/imageView/FRF00814/cars-driving-on-road-at-morning-mist

[30] Anon., “7 Expert Tips That Will Make You a Confident Driver Even in Foggy Weather,” Driving-Tests.org. Accessed: Dec. 05, 2023. [Online]. Available: https://driving-tests.org/beginner-drivers/what-to-do-when-driving-in-fog/

[31] F. Jacques, “Air pollution particles: what are they?,” Encyclopedia of the Environment. Accessed: Dec. 05, 2023. [Online]. Available: https://www.encyclopedie-environnement.org/en/air-en/air-pollution-particles-what-are-they/

[32] Anon., “What is the difference between smog and fog?” Accessed: Dec. 05, 2023. [Online]. Available: http://home.iitk.ac.in/~anubha/Smog.pdf

[33] N. V. Ogueke and N. Emekwuru, “Regulation of Nanorefrigerant Use: A Proactive Measure Against Possible Undesirable Health and Environmental Implications,” *European Journal of Sustainable Development Research*, vol. 1, no. 2, Jun. 2017, doi: 10.20897/ejosdr.201709.

[34] Anon., “Smog,” National Geographic Society. Accessed: Dec. 05, 2023. [Online]. Available: https://www.nationalgeographic.org/encyclopedia/smog/

[35] Anon., “Atmospheric Dust.” Accessed: Dec. 05, 2023. [Online]. Available: http://www.eniscuola.net/en/argomento/air- 29 pollution/pollutants-and-their-effects/atmospheric-dust/

[36] Anon., “Aerosols: Tiny Particulates in the Air.” Accessed: Dec. 05, 2023. [Online]. Available: https://scied.ucar.edu/learning-zone/air-quality/aerosols#:~:text=Aerosols%20are%20a%

[37] Anon., “About Smoke and Health.” Accessed: Dec. 05, 2023. [Online]. Available: https://www.ourair.org/about-smoke-and-health/#:~:text=Burning%20typically%20releases

[38] Anon., “Particle Sizes.” Accessed: Dec. 05, 2023. [Online]. Available: https://www.engineeringtoolbox.com/particle-sizes-d_934.html

[39] H. Israël and F. Kasten, “KOSCHMIEDERs Theorie der horizontalen Sichtweite,” in *Die Sichtweite im Nebel und die Möglichkeiten ihrer künstlichen Beeinflussung*, Wiesbaden: VS Verlag für Sozialwissenschaften, 1959, pp. 7–10. doi: 10.1007/978-3-663-04661-5_2.

[40] W. Wang, X. Yuan, X. Wu, and Y. Dong, “An airlight estimation method for image dehazing based on gray projection,” *Multimed Tools Appl*, vol. 79, no. 37–38, pp. 27185–27203, Oct. 2020, doi: 10.1007/s11042-020-09380-w.

[41] Y. Yang, C. Wang, R. Liu, L. Zhang, X. Guo, and D. Tao, “Self-augmented Unpaired Image Dehazing via Density and Depth Decomposition,” in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2022, pp. 2027–2036. doi: 10.1109/CVPR52688.2022.00208.

[42] W. Wang, X. Yuan, X. Wu, and Y. Liu, “Dehazing for images with large sky region,” *Neurocomputing*, vol. 238, pp. 365–376, May 2017, doi: 10.1016/j.neucom.2017.01.075.

[43] S. Chaturvedi, P. Khanna, and A. Ojha, “An Efficient Residual Convolutional Neural Network with Attention Mechanism for Smoke Detection in Outdoor Environment,” 2023, pp. 1–14. doi: 10.1007/978-3-031-31417-9_1.

[44] S. Chaturvedi, P. Khanna, and A. Ojha, “A survey on vision-based outdoor smoke detection techniques for environmental safety,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 185, pp. 158–187, Mar. 2022, doi: 10.1016/j.isprsjprs.2022.01.013.

[45] J. Gui *et al.*, “A Comprehensive Survey and Taxonomy on Single Image Dehazing Based on Deep Learning,” *ACM Comput Surv*, vol. 55, no. 13s, pp. 1–37, Dec. 2023, doi: 10.1145/3576918.

[46] C. C. Chatterjee, “Basics of the Classic CNN,” Medium. Accessed: Dec. 05, 2023. [Online]. Available: https://towardsdatascience.com/basics-of-the-classic-cnn-a3dce1225add

[47] J. Brownlee, “How Do Convolutional Layers Work in Deep Learning Neural Networks?,” Machine Learning Mastery. Accessed: Dec. 05, 2023. [Online]. Available: https://machinelearningmastery.com/convolutional-layers-for-deep-learning-neural-networks/

[48] J. Brownlee, “A Gentle Introduction to the Rectified Linear Unit (ReLU) for Deep Learning Neural Networks,” Machine Learning Mastery. Accessed: Dec. 05, 2023. [Online]. Available: https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/

[49] J. Yuan, B. Ni, and A. A. Kassim, “Half-CNN: A General Framework for Whole-Image Regression,” Dec. 2014, Accessed: Dec. 05, 2023. [Online]. Available: https://doi.org/10.48550/arxiv.1412.6885

[50] A. Bindal, “Normalization Techniques in Deep Neural Networks,” Techspace. Accessed: Dec. 05, 2023. [Online]. Available: https://medium.com/techspace-usict/normalization-techniques-in-deep-neural-networks-9121bf100d8

[51] X. Jiang, J. Sun, C. Li, and H. Ding, “Video Image Defogging Recognition Based on Recurrent Neural Network,” *IEEE Trans Industr Inform*, vol. 14, no. 7, pp. 3281–3288, Jul. 2018, doi: 10.1109/TII.2018.2810188.

[52] O. Kuybeda, “Scene-based nonuniformity correction using a convolutional recurrent neural network,” Patent 10,929,955, Feb. 23, 2021

[53] I. Goodfellow *et al.*, “Generative adversarial networks,” *Commun ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2020, doi: 10.1145/3422622.

[54] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-Image Translation with Conditional Adversarial Networks,” in *Computer Vision and Pattern Recognition*, 2017, pp. 1125–1134. Accessed: Dec. 05, 2023. [Online]. Available: https://arxiv.org/abs/1611.07004

[55] D. Singh and V. Kumar, “A Comprehensive Review of Computational Dehazing Techniques,” *Archives of Computational Methods in Engineering*, vol. 26, no. 5, pp. 1395–1413, Nov. 2019, doi: 10.1007/s11831-018-9294-z.

[56] A. Ansari, H. Danyali, and M. S. Helfroush, “HS remote sensing image restoration using fusion with MS images by EM algorithm,” *IET Signal Processing*, vol. 11, no. 1, pp. 95–103, Feb. 2017, doi: 10.1049/iet-spr.2016.0141.

[57] X. Tang and L. Jiao, “Fusion Similarity-Based Reranking for SAR Image Retrieval,” *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 2, pp. 242–246, Feb. 2017, doi: 10.1109/LGRS.2016.2636819.

[58] D. Singh and V. Kumar, “A novel dehazing model for remote sensing images,” *Computers & Electrical Engineering*, vol. 69, pp. 14–27, Jul. 2018, doi: 10.1016/j.compeleceng.2018.05.015.

[59] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J.-Y. Tourneret, M. Chen, and S. Godsill, “Multiband Image Fusion Based on Spectral Unmixing,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 7236–7249, Dec. 2016, doi: 10.1109/TGRS.2016.2598784.

[60] R. Balasubramanian, R. Kumar, and B. K. Kaushik, “FPGA implementation of image dehazing algorithm for real time applications,” in *Applications of Digital*

Image Processing XL, A. G. Tescher, Ed., SPIE, Sep. 2017, p. 111. doi: 10.1117/12.2274682.

[61] A. Conca, C. Ridella, and E. Sapori, "A Risk Assessment for Road Transportation of Dangerous Goods: A Routing Solution," *Transportation Research Procedia*, vol. 14, pp. 2890–2899, 2016, doi: 10.1016/j.trpro.2016.05.407.

[62] P. Pellegrini and J. Rodriguez, "Single European Sky and Single European Railway Area: A system level analysis of air and rail transportation," *Transp Res Part A Policy Pract*, vol. 57, pp. 64–86, Nov. 2013, doi: 10.1016/j.tra.2013.09.004.

[63] K. Fang, G. Y. Ke, and M. Verma, "A routing and scheduling approach to rail transportation of hazardous materials with demand due dates," *Eur J Oper Res*, vol. 261, no. 1, pp. 154–168, Aug. 2017, doi: 10.1016/j.ejor.2017.01.045.

[64] J. P. Kennedy and J. M. Wilson, "Liabilities and responsibilities: ocean transportation intermediaries (OTIs) and the distribution of counterfeit goods," *Maritime Economics & Logistics*, vol. 19, no. 1, pp. 182–187, Mar. 2017, doi: 10.1057/s41278-017-0064-x.

[65] A. Beck, J. Henneberger, S. Schöpfer, J. Fugal, and U. Lohmann, "HoloGondel: in situ cloud observations on a cable car in the Swiss Alps using a holographic imager," *Atmos Meas Tech*, vol. 10, no. 2, pp. 459–476, Feb. 2017, doi: 10.5194/amt-10-459-2017.

[66] C. Qing, F. Yu, X. Xu, W. Huang, and J. Jin, "Underwater video dehazing based on spatial–temporal information fusion," *Multidimens Syst Signal Process*, vol. 27, no. 4, pp. 909–924, Oct. 2016, doi: 10.1007/s11045-016-0407-2.

[67] W. Liu, X. Hou, J. Duan, and G. Qiu, "End-to-End Single Image Fog Removal Using Enhanced Cycle Consistent Adversarial Networks," *IEEE Transactions on Image Processing*, vol. 29, pp. 7819–7833, 2020, doi: 10.1109/TIP.2020.3007844.

[68] J.-P. Tarel, N. Hautiere, A. Cord, D. Gruyer, and H. Halmaoui, "Improved visibility of road scene images under heterogeneous fog," in *2010 IEEE*

Intelligent Vehicles Symposium, IEEE, Jun. 2010, pp. 478–485. doi: 10.1109/IVS.2010.5548128.

[69] Y. Zhang, L. Ding, and G. Sharma, "HazeRD: An outdoor scene dataset and benchmark for single image dehazing," in *2017 IEEE International Conference on Image Processing (ICIP)*, IEEE, Sep. 2017, pp. 3205–3209. doi: 10.1109/ICIP.2017.8296874.

[70] C. Ancuti, C. O. Ancuti, and C. De Vleeschouwer, "D-HAZY: A dataset to evaluate quantitatively dehazing algorithms," in *2016 IEEE International Conference on Image Processing (ICIP)*, IEEE, Sep. 2016, pp. 2226–2230. doi: 10.1109/ICIP.2016.7532754.

[71] D. Scharstein *et al.*, "High-Resolution Stereo Datasets with Subpixel-Accurate Ground Truth," 2014, pp. 31–42. doi: 10.1007/978-3-319-11752-2_3.

[72] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor Segmentation and Support Inference from RGBD Images," 2012, pp. 746–760. doi: 10.1007/978-3-642-33715-4_54.

[73] C. O. Ancuti, C. Ancuti, M. Sbert, and R. Timofte, "Dense-Haze: A Benchmark for Image Dehazing with Dense-Haze and Haze-Free Images," in *2019 IEEE International Conference on Image Processing (ICIP)*, IEEE, Sep. 2019, pp. 1014–1018. doi: 10.1109/ICIP.2019.8803046.

[74] C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer, "O-HAZE: A Dehazing Benchmark with Real Hazy and Haze-Free Outdoor Images," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, Jun. 2018, pp. 867–8678. doi: 10.1109/CVPRW.2018.00119.

[75] B. Li *et al.*, "Benchmarking Single-Image Dehazing and Beyond," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 492–505, Jan. 2019, doi: 10.1109/TIP.2018.2867951.

[76] C. O. Ancuti, C. Ancuti, and R. Timofte, "NH-HAZE: An Image Dehazing Benchmark with Non-Homogeneous Hazy and Haze-Free Images," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*

(CVPRW), IEEE, Jun. 2020, pp. 1798–1805. doi: 10.1109/CVPRW50498.2020.00230.

[77] D. Berman, T. Treibitz, and S. Avidan, "Non-local Image Dehazing," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 1674–1682. doi: 10.1109/CVPR.2016.185.

[78] Anon., "Dehazing using color-lines: Results and Comparisons." Accessed: Dec. 06, 2023. [Online]. Available: https://www.cs.huji.ac.il/w~raananf/projects/dehaze_cl/results/

[79] K. Ma, W. Liu, and Z. Wang, "Perceptual evaluation of single image dehazing algorithms," in *2015 IEEE International Conference on Image Processing (ICIP)*, IEEE, Sep. 2015, pp. 3600–3604. doi: 10.1109/ICIP.2015.7351475.

[80] J.-P. Tarel, N. Hautiere, L. Caraffa, A. Cord, H. Halmaoui, and D. Gruyer, "Vision Enhancement in Homogeneous and Heterogeneous Fog," *IEEE Intelligent Transportation Systems Magazine*, vol. 4, no. 2, pp. 6–20, 2012, doi: 10.1109/MITS.2012.2189969.

[81] C. Ancuti, C. O. Ancuti, R. Timofte, and C. De Vleeschouwer, "I-HAZE: A Dehazing Benchmark with Real Hazy and Haze-Free Indoor Images," 2018, pp. 620–631. doi: 10.1007/978-3-030-01449-0_52.

[82] K. Ma, H. Li, H. Yong, Z. Wang, D. Meng, and L. Zhang, "Robust Multi-Exposure Image Fusion: A Structural Patch Decomposition Approach," *IEEE Transactions on Image Processing*, vol. 26, no. 5, pp. 2519–2532, May 2017, doi: 10.1109/TIP.2017.2671921.

[83] Lark Kwon Choi, Jaehee You, and A. C. Bovik, "Referenceless Prediction of Perceptual Fog Density and Perceptual Image Defogging," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3888–3901, Nov. 2015, doi: 10.1109/TIP.2015.2456502.

[84] M. Everingham, L. VanGool, C. K. I. Williams, J. Winn, and A. Zisserman, "The PASCAL Visual Object Classes Challenge 2007 (VOC2007) Results." Accessed: Dec. 06, 2023. [Online]. Available: http://ww38.pascalnetwork.org/challenges/VOC/voc2007/workshop/index.html

[85] A. Saxena, Min Sun, and A. Y. Ng, “Make3D: Learning 3D Scene Structure from a Single Still Image,” *IEEE Trans Pattern Anal Mach Intell*, vol. 31, no. 5, pp. 824–840, May 2009, doi: 10.1109/TPAMI.2008.132.

[86] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2009, pp. 248–255. doi: 10.1109/CVPR.2009.5206848.

[87] A. Geiger, P. Lenz, and R. Urtasun, “Are we ready for autonomous driving? The KITTI vision benchmark suite,” in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2012, pp. 3354–3361. doi: 10.1109/CVPR.2012.6248074.

[88] C. O. Ancuti *et al.*, “NTIRE 2021 NonHomogeneous Dehazing Challenge Report,” in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, Jun. 2021, pp. 627–646. doi: 10.1109/CVPRW53098.2021.00074.

[89] C. O. Ancuti and et al., “NTIRE 2023 HR NonHomogeneous Dehazing Challenge Report,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, Jun. 2023, pp. 1808–1825.

[90] M. Cordts *et al.*, “The Cityscapes Dataset for Semantic Urban Scene Understanding,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 3213–3223. doi: 10.1109/CVPR.2016.350.

[91] C. Sakaridis, D. Dai, and L. Van Gool, “Semantic Foggy Scene Understanding with Synthetic Data,” *Int J Comput Vis*, vol. 126, no. 9, pp. 973–992, Sep. 2018, doi: 10.1007/s11263-018-1072-8.

[92] Z. Li and N. Snavely, “MegaDepth: Learning Single-View Depth Prediction from Internet Photos,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2018, pp. 2041–2050. doi: 10.1109/CVPR.2018.00218.

[93] S. Emberton, L. Chittka, and A. Cavallaro, “Underwater image and video dehazing with pure haze region segmentation,” *Computer Vision and Image*

Understanding, vol. 168, pp. 145–156, Mar. 2018, doi: 10.1016/j.cviu.2017.08.003.

[94] A. S. Abdul Ghani and N. A. Mat Isa, "Automatic system for improving underwater image contrast and color through recursive adaptive histogram modification," *Comput Electron Agric*, vol. 141, pp. 181–195, Sep. 2017, doi: 10.1016/j.compag.2017.07.021.

[95] Y.-T. Peng and P. C. Cosman, "Underwater Image Restoration Based on Image Blurriness and Light Absorption," *IEEE Transactions on Image Processing*, vol. 26, no. 4, pp. 1579–1594, Apr. 2017, doi: 10.1109/TIP.2017.2663846.

[96] B.-H. Chen, S.-C. Huang, and J. H. Ye, "Hazy Image Restoration by Bi-Histogram Modification," *ACM Trans Intell Syst Technol*, vol. 6, no. 4, pp. 1–17, Aug. 2015, doi: 10.1145/2710024.

[97] U. A. Nnolim, "Improved partial differential equation-based enhancement for underwater images using local–global contrast operators and fuzzy homomorphic processes," *IET Image Process*, vol. 11, no. 11, pp. 1059–1067, Nov. 2017, doi: 10.1049/iet-ipr.2017.0259.

[98] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004, doi: 10.1109/TIP.2003.819861.

[99] N. Sharma, V. Kumar, and S. K. Singla, "Single Image Defogging using Deep Learning Techniques: Past, Present and Future," *Archives of Computational Methods in Engineering*, vol. 28, no. 7, pp. 4449–4469, Dec. 2021, doi: 10.1007/s11831-021-09541-6.

[100] Y. He, S. Chen, B. Zhang, and K. Chen, "Multimedia Tilt Photography-Assisted Remote Sensing Technology in Mine Ecological Restoration," *Comput Intell Neurosci*, vol. 2022, pp. 1–5, Jun. 2022, doi: 10.1155/2022/1442738.

[101] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," *Geosci*

Model Dev, vol. 7, no. 3, pp. 1247–1250, Jun. 2014, doi: 10.5194/gmd-7-1247-2014.

[102] M. R. Luo, G. Cui, and B. Rigg, "The development of the CIE 2000 colour-difference formula: CIEDE2000," *Color Res Appl*, vol. 26, no. 5, pp. 340–350, Oct. 2001, doi: 10.1002/col.1049.

[103] G. Sharma, W. Wu, and E. N. Dalal, "The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations," *Color Res Appl*, vol. 30, no. 1, pp. 21–30, Feb. 2005, doi: 10.1002/col.20070.

[104] D. Singh and V. Kumar, "Dehazing of remote sensing images using improved restoration model based dark channel prior," *The Imaging Science Journal*, vol. 65, no. 5, pp. 282–292, Jul. 2017, doi: 10.1080/13682199.2017.1329792.

[105] C. Li and J. Guo, "Underwater image enhancement by dehazing and color correction," *J Electron Imaging*, vol. 24, no. 3, p. 033023, Jun. 2015, doi: 10.1117/1.JEI.24.3.033023.

[106] N. Hautière, J.-P. Tarel, D. Aubert, and É. Dumont, "BLIND CONTRAST ENHANCEMENT ASSESSMENT BY GRADIENT RATIOING AT VISIBLE EDGES," *Image Analysis & Stereology*, vol. 27, no. 2, p. 87, May 2011, doi: 10.5566/ias.v27.p87-95.

[107] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'Completely Blind' Image Quality Analyzer," *IEEE Signal Process Lett*, vol. 20, no. 3, pp. 209–212, Mar. 2013, doi: 10.1109/LSP.2012.2227726.

[108] Venkatanath N, Praneeth D, Maruthi Chandrasekhar Bh, S. S. Channappayya, and S. S. Medasani, "Blind image quality evaluation using perception based features," in *2015 Twenty First National Conference on Communications (NCC)*, IEEE, Feb. 2015, pp. 1–6. doi: 10.1109/NCC.2015.7084843.

[109] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012, doi: 10.1109/TIP.2012.2214050.

[110] C. Wang and H.-W. Shen, "Information Theory in Scientific Visualization," *Entropy*, vol. 13, no. 1, pp. 254–273, Jan. 2011, doi: 10.3390/e13010254.

[111] L. K. Choi, J. You, and A. C. Bovik, “Referenceless perceptual fog density prediction model,” *Human Vision and Electronic Imaging XIX*, vol. 9014, pp. 90–101, Feb. 2014.

[112] L. K. Choi, J. You, and A. C. Bovik, “Referenceless perceptual image defogging,” in *2014 Southwest Symposium on Image Analysis and Interpretation*, IEEE, Apr. 2014, pp. 165–168. doi: 10.1109/SSIAI.2014.6806055.

[113] L. K. Choi, J. You, and A. C. Bovik, “LIVE Image Defogging Database.” Accessed: Dec. 06, 2023. [Online]. Available: http://live.ece.utexas.edu/research/fog/fade_defade.html

[114] S. Bezryadin, P. Bourov, and D. Ilinih, “Brightness Calculation in Digital Image Processing,” *International Symposium on Technologies for Digital Photo Fulfillment*, vol. 1, no. 1, pp. 10–15, Jan. 2007, doi: 10.2352/ISSN.2169-4672.2007.1.0.10.

[115] L. Kratz and K. Nishino, “Factorizing Scene Albedo and Depth from a Single Foggy Image,” in *2009 IEEE 12th International Conference on Computer Vision*, IEEE, Sep. 2009, pp. 1701–1708. doi: 10.1109/ICCV.2009.5459382.

[116] Kaiming He, Jian Sun, and Xiaoou Tang, “Single Image Haze Removal Using Dark Channel Prior,” *IEEE Trans Pattern Anal Mach Intell*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011, doi: 10.1109/TPAMI.2010.168.

[117] Qingsong Zhu, Jiaming Mai, and Ling Shao, “A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior,” *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3522–3533, Nov. 2015, doi: 10.1109/TIP.2015.2446191.

[118] L. Zhang, X. Li, B. Hu, and X. Ren, “Research on Fast Smog Free Algorithm on Single Image,” in *2015 First International Conference on Computational Intelligence Theory, Systems and Applications (CCITSA)*, IEEE, Dec. 2015, pp. 177–182. doi: 10.1109/CCITSA.2015.55.

[119] R. Wang and G. Wang, “Single Smog Image Dehazing Method,” in *2016 3rd International Conference on Information Science and Control Engineering (ICISCE)*, IEEE, Jul. 2016, pp. 621–625. doi: 10.1109/ICISCE.2016.138.

[120] J. Zhang, Y. Cao, S. Fang, Y. Kang, and C. W. Chen, “Fast Haze Removal for Nighttime Image Using Maximum Reflectance Prior,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 7016–7024. doi: 10.1109/CVPR.2017.742.

[121] T. M. Bui and W. Kim, “Single Image Dehazing Using Color Ellipsoid Prior,” *IEEE Transactions on Image Processing*, vol. 27, no. 2, pp. 999–1009, Feb. 2018, doi: 10.1109/TIP.2017.2771158.

[122] W.-T. Chen, S.-Y. Yuan, G.-C. Tsai, H.-C. Wang, and S.-Y. Kuo, “Color Channel-Based Smoke Removal Algorithm Using Machine Learning for Static Images,” in *2018 25th IEEE International Conference on Image Processing (ICIP)*, IEEE, Oct. 2018, pp. 2855–2859. doi: 10.1109/ICIP.2018.8451581.

[123] D. Singh and V. Kumar, “Defogging of road images using gain coefficient-based trilateral filter,” *J Electron Imaging*, vol. 27, no. 01, p. 1, Jan. 2018, doi: 10.1117/1.JEI.27.1.013004.

[124] D. Singh and V. Kumar, “Single image haze removal using integrated dark and bright channel prior,” *Modern Physics Letters B*, vol. 32, no. 04, p. 1850051, Feb. 2018, doi: 10.1142/S0217984918500513.

[125] D. Singh and V. Kumar, “Single image defogging by gain gradient image filter,” *Science China Information Sciences*, vol. 62, no. 7, p. 79101, Jul. 2019, doi: 10.1007/s11432-017-9433-4.

[126] D. Ngo, G.-D. Lee, and B. Kang, “Improved Color Attenuation Prior for Single-Image Haze Removal,” *Applied Sciences*, vol. 9, no. 19, p. 4011, Sep. 2019, doi: 10.3390/app9194011.

[127] D. Singh, V. Kumar, and M. Kaur, “Single image dehazing using gradient channel prior,” *Applied Intelligence*, vol. 49, no. 12, pp. 4276–4293, Dec. 2019, doi: 10.1007/s10489-019-01504-6.

[128] D. Singh, V. Kumar, and M. Kaur, “Image dehazing using window-based integrated means filter,” *Multimed Tools Appl*, vol. 79, no. 47–48, pp. 34771–34793, Dec. 2020, doi: 10.1007/s11042-019-08286-6.

[129] J. Bala and K. Lakhwani, “Desmogging of Smog Affected Images Using Illumination Channel Prior,” 2020, pp. 417–425. doi: 10.1007/978-981-15-0324-5_36.

[130] J. Li, Q. Hu, and M. Ai, “Haze and Thin Cloud Removal via Sphere Model Improved Dark Channel Prior,” *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 3, pp. 472–476, Mar. 2019, doi: 10.1109/LGRS.2018.2874084.

[131] M. Kaur, D. Singh, V. Kumar, and K. Sun, “Color image dehazing using gradient channel prior and guided L0 filter,” *Inf Sci (N Y)*, vol. 521, pp. 326–342, Jun. 2020, doi: 10.1016/j.ins.2020.02.048.

[132] D. Ngo, S. Lee, and B. Kang, “Robust Single-Image Haze Removal Using Optimal Transmission Map and Adaptive Atmospheric Light,” *Remote Sens (Basel)*, vol. 12, no. 14, p. 2233, Jul. 2020, doi: 10.3390/rs12142233.

[133] S. C. Raikwar and S. Tapaswi, “Lower Bound on Transmission Using Non-Linear Bounding Function in Single Image Dehazing,” *IEEE Transactions on Image Processing*, vol. 29, pp. 4832–4847, 2020, doi: 10.1109/TIP.2020.2975909.

[134] J. Bala and K. Lakhwani, “Single image desmogging using oblique gradient profile prior and variational minimization,” *Multidimens Syst Signal Process*, vol. 31, no. 4, pp. 1259–1275, Oct. 2020, doi: 10.1007/s11045-020-00707-2.

[135] J. Bala and K. Lakhwani, “Single image desmogging using Gradient channel prior and Information gain based bilateral,” in *2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE)*, IEEE, Feb. 2020, pp. 1–6. doi: 10.1109/ICETCE48199.2020.9091768.

[136] Y. Gao, H. Li, Y. Su, and J. Li, “Structural Patch Decomposition Fusion for Single Image Dehazing,” 2020, pp. 304–314. doi: 10.1007/978-981-15-8135-9_22.

[137] S. Kanti Dhara, M. Roy, D. Sen, and P. Kumar Biswas, “Color Cast Dependent Image Dehazing via Adaptive Airlight Refinement and Non-Linear Color

Balancing," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 5, pp. 2076–2081, May 2021, doi: 10.1109/TCSVT.2020.3007850.

[138] X. Zhao, "Single Image Dehazing Using Bounded Channel Difference Prior," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, Jun. 2021, pp. 727–735. doi: 10.1109/CVPRW53098.2021.00082.

[139] M. Ju, C. Ding, W. Ren, Y. Yang, D. Zhang, and Y. J. Guo, "IDE: Image Dehazing and Exposure Using an Enhanced Atmospheric Scattering Model," *IEEE Transactions on Image Processing*, vol. 30, pp. 2180–2192, 2021, doi: 10.1109/TIP.2021.3050643.

[140] Q. Hu, Y. Zhang, Y. Zhu, Y. Jiang, and M. Song, "Single image dehazing algorithm based on sky segmentation and optimal transmission maps," *Vis Comput*, vol. 39, no. 3, pp. 997–1013, Mar. 2023, doi: 10.1007/s00371-021-02380-3.

[141] F. Hussain and J. Jeong, "Visibility Enhancement of Scene Images Degraded by Foggy Weather Conditions with Deep Neural Networks," *J Sens*, vol. 2016, pp. 1–9, 2016, doi: 10.1155/2016/3894832.

[142] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: An End-to-End System for Single Image Haze Removal," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187–5198, Nov. 2016, doi: 10.1109/TIP.2016.2598681.

[143] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, "Single Image Dehazing via Multi-scale Convolutional Neural Networks," 2016, pp. 154–169. doi: 10.1007/978-3-319-46475-6_10.

[144] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "AOD-Net: All-in-One Dehazing Network," in *2017 IEEE International Conference on Computer Vision (ICCV)*, IEEE, Oct. 2017, pp. 4780–4788. doi: 10.1109/ICCV.2017.511.

[145] X. Zhao, K. Wang, Y. Li, and J. Li, "Deep fully convolutional regression networks for single image haze removal," in *2017 IEEE Visual Communications and Image Processing (VCIP)*, IEEE, Dec. 2017, pp. 1–4. doi: 10.1109/VCIP.2017.8305035.

[146] C. Li, J. Guo, F. Porikli, H. Fu, and Y. Pang, “A Cascaded Convolutional Neural Network for Single Image Dehazing,” *IEEE Access*, vol. 6, pp. 24877–24887, 2018, doi: 10.1109/ACCESS.2018.2818882.

[147] K. Liu, L. He, S. Ma, S. Gao, and D. Bi, “A Sensor Image Dehazing Algorithm Based on Feature Learning,” *Sensors*, vol. 18, no. 8, p. 2606, Aug. 2018, doi: 10.3390/s18082606.

[148] W. Ren *et al.*, “Gated Fusion Network for Single Image Dehazing,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2018, pp. 3253–3261. doi: 10.1109/CVPR.2018.00343.

[149] G. Kim, S. Ha, and J. Kwon, “Adaptive Patch Based Convolutional Neural Network for Robust Dehazing,” in *2018 25th IEEE International Conference on Image Processing (ICIP)*, IEEE, Oct. 2018, pp. 2845–2849. doi: 10.1109/ICIP.2018.8451252.

[150] R. Wang, Y. You, Y. Zhang, W. Zhou, and J. Liu, “Ship detection in foggy remote sensing image via scene classification R-CNN,” in *2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC)*, IEEE, Aug. 2018, pp. 81–85. doi: 10.1109/ICNIDC.2018.8525532.

[151] A. Dudhane and S. Murala, “C^2MSNet: A Novel Approach for Single Image Haze Removal,” in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, IEEE, Mar. 2018, pp. 1397–1404. doi: 10.1109/WACV.2018.00157.

[152] R. Li, J. Pan, Z. Li, and J. Tang, “Single Image Dehazing via Conditional Generative Adversarial Network,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2018, pp. 8202–8211. doi: 10.1109/CVPR.2018.00856.

[153] D. Chen *et al.*, “Gated context aggregation network for image dehazing and deraining,” in *IEEE winter conference on applications of computer vision (WACV)*, IEEE, Jan. 2019, pp. 1375–1383.

[154] L.-Y. Huang, J.-L. Yin, B.-H. Chen, and S.-Z. Ye, “Towards Unsupervised Single Image Dehazing With Deep Learning,” in *2019 IEEE International Conference*

on Image Processing (ICIP), IEEE, Sep. 2019, pp. 2741–2745. doi: 10.1109/ICIP.2019.8803316.

[155] A. Wang, W. Wang, J. Liu, and N. Gu, "AIPNet: Image-to-Image Single Image Dehazing With Atmospheric Illumination Prior," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 381–393, Jan. 2019, doi: 10.1109/TIP.2018.2868567.

[156] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia, "FFA-Net: Feature Fusion Attention Network for Single Image Dehazing," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 07, pp. 11908–11915, Apr. 2020, doi: 10.1609/aaai.v34i07.6865.

[157] A. Singh, A. Bhave, and D. K. Prasad, "Single Image Dehazing for a Variety of Haze Scenarios Using Back Projected Pyramid Network," 2020, pp. 166–181. doi: 10.1007/978-3-030-66823-5_10.

[158] A. Bennur and M. Gaggar, "LCA-Net: Light Convolutional Autoencoder for Image Dehazing," *Computer Vision and Pattern Recognition*, Aug. 2020.

[159] H. Dong and et al., "Multi-scale boosted dehazing network with dense feature fusion," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 2157–2167.

[160] W. Ren, J. Pan, H. Zhang, X. Cao, and M.-H. Yang, "Single Image Dehazing via Multi-scale Convolutional Neural Networks with Holistic Edges," *Int J Comput Vis*, vol. 128, no. 1, pp. 240–259, Jan. 2020, doi: 10.1007/s11263-019-01235-8.

[161] H. Wu and et al., "Contrastive Learning for Compact Single Image Dehazing," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE/CVF, Jun. 2021, pp. 10551–10560.

[162] M. Fu, H. Liu, Y. Yu, J. Chen, and K. Wang, "DW-GAN: A Discrete Wavelet Transform GAN for NonHomogeneous Dehazing," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, Jun. 2021, pp. 203–212. doi: 10.1109/CVPRW53098.2021.00029.

[163] A. Kumar and A. Jain, “Image smog restoration using oblique gradient profile prior and energy minimization,” *Front Comput Sci*, vol. 15, no. 6, p. 156706, Dec. 2021, doi: 10.1007/s11704-020-9305-8.

[164] A. Jain and A. Kumar, “Desmogging of still smoggy images using a novel channel prior,” *J Ambient Intell Humaniz Comput*, vol. 12, no. 1, pp. 1161–1177, Jan. 2021, doi: 10.1007/s12652-020-02161-1.

[165] J. Zhang, X. Zhang, T. Li, Y. Zeng, G. Lv, and F. Nian, “Visible light polarization image desmogging via Cycle Convolutional Neural Network,” *Multimed Syst*, vol. 28, no. 1, pp. 45–55, Feb. 2022, doi: 10.1007/s00530-021-00802-9.

[166] X. Li, Z. Hua, and J. Li, “Attention-based adaptive feature selection for multi-stage image dehazing,” *Vis Comput*, vol. 39, no. 2, pp. 663–678, Feb. 2023, doi: 10.1007/s00371-021-02365-2.

[167] S. Zhang, J. Zhang, F. He, and N. Hou, “DRDDN: dense residual and dilated dehazing network,” *Vis Comput*, vol. 39, no. 3, pp. 953–969, Mar. 2023, doi: 10.1007/s00371-021-02377-y.

[168] Z. Tu *et al.*, “MAXIM: Multi-Axis MLP for Image Processing,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE/CVF, 2022, pp. 5769–5780.

[169] S. Shit, D. K. Das, D. N. Ray, and B. Roy, “An encoder-decoder based CNN architecture using end to end dehaze and detection network for proper image visualization and detection,” *Comput Animat Virtual Worlds*, Feb. 2023, doi: 10.1002/cav.2147.

[170] W. Li, G. Fan, and M. Gan, “Progressive encoding-decoding image dehazing network,” *Multimed Tools Appl*, Jun. 2023, doi: 10.1007/s11042-023-15638-w.

[171] T. Hoang, H. Zhang, A. Yazdani, and V. Monga, “TransER: Hybrid Model and Ensemble-based Sequential Learning for Non-homogenous Dehazing,” in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, Jun. 2023, pp. 1670–1679. doi: 10.1109/CVPRW59228.2023.00168.

[172] Y. Song, J. Li, X. Wang, and X. Chen, “Single image dehazing using ranking convolutional neural network,” *IEEE Trans Multimedia*, vol. 20, no. 6, pp. 1548–1560, 2017.

[173] J. Li, G. Li, and H. Fan, “Image Dehazing Using Residual-Based Deep CNN,” *IEEE Access*, vol. 6, pp. 26831–26842, 2018, doi: 10.1109/ACCESS.2018.2833888.

[174] J.-H. Kim, W.-D. Jang, J.-Y. Sim, and C.-S. Kim, “Optimized contrast enhancement for real-time image and video dehazing,” *J Vis Commun Image Represent*, vol. 24, no. 3, pp. 410–425, Apr. 2013, doi: 10.1016/j.jvcir.2013.02.004.

[175] W. Ren *et al.*, “Deep Video Dehazing With Semantic Segmentation,” *IEEE Transactions on Image Processing*, vol. 28, no. 4, pp. 1895–1908, Apr. 2019, doi: 10.1109/TIP.2018.2876178.

[176] G. Mandal, P. De, and D. Bhattacharya, “A real-time fast defogging system to clear the vision of driver in foggy highway using minimum filter and gamma correction,” *Sādhanā*, vol. 45, no. 1, p. 40, Dec. 2020, doi: 10.1007/s12046-020-1282-y.

[177] P. Soma and R. K. Jatoth, “An efficient and contrast-enhanced video de-hazing based on transmission estimation using HSL color model,” *Vis Comput*, vol. 38, no. 7, pp. 2569–2580, Jul. 2022, doi: 10.1007/s00371-021-02132-3.

[178] R. Kumar, R. Balasubramanian, and B. K. Kaushik, “Efficient Method and Architecture for Real-Time Video Defogging,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 10, pp. 6536–6546, Oct. 2021, doi: 10.1109/TITS.2020.2993906.

[179] S.-J. Peng, H. Zhang, X. Liu, W. Fan, B. Zhong, and J.-X. Du, “Real-time video dehazing via incremental transmission learning and spatial-temporally coherent regularization,” *Neurocomputing*, vol. 458, pp. 602–614, Oct. 2021, doi: 10.1016/j.neucom.2020.02.134.

[180] S. Adidela, S. Singh, T. Sahu, and A. Mishra, “Single Image And Video Dehazing: A Dark Channel Prior (DCP)-based Approach,” in *2021 IEEE 18th*

India Council International Conference (INDICON), IEEE, Dec. 2021, pp. 1–6. doi: 10.1109/INDICON52576.2021.9691546.

[181] X. Zhang and et al., "Learning To Restore Hazy Video: A New Real-World Dataset and a New Method," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE/CVF, 2021, pp. 9239–9248.

[182] Anon., "Preparing Your Dataset for Machine Learning: 8 Basic Techniques That Make Your Data Better," AltexSoft. Accessed: Dec. 06, 2023. [Online]. Available: https://www.altexsoft.com/blog/datascience/preparing-your-dataset-for-machine-learning-8-basic-techniques-that-make-your-data-better/

[183] Y. Chen *et al.*, "The image annotation algorithm using convolutional features from intermediate layer of deep learning," *Multimed Tools Appl*, vol. 80, no. 3, pp. 4237–4261, Jan. 2021, doi: 10.1007/s11042-020-09887-2.

[184] L. D. Mech, "Alpha status, dominance, and division of labor in wolf packs," *Can J Zool*, vol. 77, no. 8, pp. 1196–1203, Nov. 1999, doi: 10.1139/z99-099.

[185] C. Muro, R. Escobedo, L. Spector, and R. P. Coppinger, "Wolf-pack (Canis lupus) hunting strategies emerge from simple rules in computational simulations," *Behavioural Processes*, vol. 88, no. 3, pp. 192–197, Nov. 2011, doi: 10.1016/j.beproc.2011.09.006.

[186] L. Ding and A. Goshtasby, "On the Canny edge detector," *Pattern Recognit*, vol. 34, no. 3, pp. 721–725, Mar. 2001, doi: 10.1016/S0031-3203(00)00023-6.

[187] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, Mar. 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[188] H. Zhang, V. Sindagi, and V. M. Patel, "Multi-scale Single Image Dehazing Using Perceptual Pyramid Deep Network," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, Jun. 2018, pp. 1015–101509. doi: 10.1109/CVPRW.2018.00135.

[189] Z. Xizhi, “The Application of Wavelet Transform in Digital Image Processing,” in *2008 International Conference on MultiMedia and Information Technology*, IEEE, Dec. 2008, pp. 326–329. doi: 10.1109/MMIT.2008.134.

[190] W. Zhang, W. Chen, Z. Wang, and H. Li, “An improved fog-removing method for the traffic monitoring image,” in *2014 12th International Conference on Signal Processing (ICSP)*, IEEE, Oct. 2014, pp. 938–941. doi: 10.1109/ICOSP.2014.7015142.

[191] Q. Fu, C. Jung, and K. Xu, “Retinex-Based Perceptual Contrast Enhancement in Images Using Luminance Adaptation,” *IEEE Access*, vol. 6, pp. 61277–61286, 2018, doi: 10.1109/ACCESS.2018.2870638.

[192] F. A. Dharejo *et al.*, “A deep hybrid neural network for single image dehazing via wavelet transform,” *Optik (Stuttg)*, vol. 231, p. 166462, Apr. 2021, doi: 10.1016/j.ijleo.2021.166462.

[193] K. E. Barner and G. R. Arce, “Order-statistic filtering and smoothing of time series: Part 2,” in *Handbook of Statistics-16 Order Statistics and their Applications*, 1996.

[194] Anon., “Home - MatConvNet.” Accessed: Dec. 06, 2023. [Online]. Available: https://www.vlfeat.org/matconvnet/

[195] A. Vedaldi and K. Lenc, “MatConvNet,” in *Proceedings of the 23rd ACM international conference on Multimedia*, New York, NY, USA: ACM, Oct. 2015, pp. 689–692. doi: 10.1145/2733373.2807412.

[196] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[197] S. Hochreiter, “The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 06, no. 02, pp. 107–116, Apr. 1998, doi: 10.1142/S0218488598000094.

[198] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, “Multiscale vessel enhancement filtering,” 1998, pp. 130–137. doi: 10.1007/BFb0056195.

[199] H. Wang, J. Peng, D. Chen, G. Jiang, T. Zhao, and X. Fu, “Attribute-Guided Feature Learning Network for Vehicle Reidentification,” *IEEE MultiMedia*, vol. 27, no. 4, pp. 112–121, Oct. 2020, doi: 10.1109/MMUL.2020.2999464.

[200] Anon., “Angiography,” National Health Service, United Kingdom. Accessed: Dec. 06, 2023. [Online]. Available: https://www.nhs.uk/conditions/angiography/#:~:text=Angiography%20is%20a%20type%20of,during%20angiography%20are%20called%20angiograms

[201] D. Ziou and S. Tabbone, “Edge Detection Techniques-An Overview,” *Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications*, vol. 8, no. 4, pp. 537–559, 1998.

[202] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever, “Scale and the differential structure of images,” *Image Vis Comput*, vol. 10, no. 6, pp. 376–388, Jul. 1992, doi: 10.1016/0262-8856(92)90024-W.

[203] J. J. Koenderink, “The structure of images,” *Biol Cybern*, vol. 50, no. 5, pp. 363–370, Aug. 1984, doi: 10.1007/BF00336961.

[204] T. Lindeberg, “Edge detection and ridge detection with automatic scale selection,” in *Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, 1996, pp. 465–470. doi: 10.1109/CVPR.1996.517113.

[205] A. Mehra, M. Mandal, P. Narang, and V. Chamola, “ReViewNet: A Fast and Resource Optimized Network for Enabling Safe Autonomous Driving in Hazy Weather Conditions,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4256–4266, Jul. 2021, doi: 10.1109/TITS.2020.3013099.

[206] J. Kumar and A. K. Singh, “Security and Privacy-Preservation of IoT Data in Cloud-Fog Computing Environment,” *Cryptography and Security*, Dec. 2022.

[207] R. Gupta and A. K. Singh, “A privacy-preserving model for cloud data storage through fog computing,” *International Journal of Computer Aided Engineering and Technology*, vol. 17, no. 3, p. 348, 2022, doi: 10.1504/IJCAET.2022.125716.

[208] G. Buchsbaum, “A spatial processor model for object colour perception,” *J Franklin Inst*, vol. 310, no. 1, pp. 1–26, Jul. 1980, doi: 10.1016/0016-0032(80)90058-7.

[209] S. Chincholkar and M. Rajapandy, “Fog Image Classification and Visibility Detection Using CNN,” 2020, pp. 249–257. doi: 10.1007/978-3-030-30465-2_28.

[210] K. G. Lore, A. Akintayo, and S. Sarkar, “LLNet: A deep autoencoder approach to natural low-light image enhancement,” *Pattern Recognit*, vol. 61, pp. 650–662, Jan. 2017, doi: 10.1016/j.patcog.2016.06.008.

[211] Anon., “GoPro Hero7 Black Specs,” Digital Photography Review. Accessed: Dec. 06, 2023. [Online]. Available: https://www.dpreview.com/products/gopro/actioncams/gopro_hero7_black/specifications

[212] Anon., “BlueRigger 4K HDMI Cable (Packs, 4K 60Hz, High Speed),” Bluerigger. Accessed: Dec. 06, 2023. [Online]. Available: https://bluerigger.com/products/bluerigger-4k-hdmi-cable-packs-4k-60hz-high-speed

[213] Anon., “Cam Link 4K – Technical Specifications,” Elgato. Accessed: Dec. 06, 2023. [Online]. Available: https://help.elgato.com/hc/en-us/articles/360027963272-Cam-Link-4K-Technical-Specifications

[214] Y. Cimtay, “Smart and real-time image dehazing on mobile devices,” *J Real Time Image Process*, vol. 18, no. 6, pp. 2063–2072, Dec. 2021, doi: 10.1007/s11554-021-01085-z.

[215] H. G. Abreha, M. Hayajneh, and M. A. Serhani, “Federated Learning in Edge Computing: A Systematic Survey,” *Sensors*, vol. 22, no. 2, p. 450, Jan. 2022, doi: 10.3390/s22020450.

[216] P. Soille, *Morphological Image Analysis*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999. doi: 10.1007/978-3-662-03939-7.

[217] B. Ganguly, A. Bhattacharya, A. Srivastava, D. Dey, and S. Munshi, “Single Image Haze Removal With Haze Map Optimization for Various Haze

Concentrations," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 1, pp. 286–301, Jan. 2022, doi: 10.1109/TCSVT.2021.3059573.

[218] X. Hu, L. Chu, J. Pei, W. Liu, and J. Bian, "Model complexity of deep learning: a survey," *Knowl Inf Syst*, vol. 63, no. 10, pp. 2585–2619, Oct. 2021, doi: 10.1007/s10115-021-01605-0.

[219] R. Mondal, M. S. Dey, and B. Chanda, "Image Restoration by Learning Morphological Opening-Closing Network," *Mathematical Morphology - Theory and Applications*, vol. 4, no. 1, pp. 87–107, Jan. 2020, doi: 10.1515/mathm-2020-0103.

[220] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *Mach Learn*, Jan. 2017.

[221] K. Janocha and W. M. Czarnecki, "On Loss Functions for Deep Neural Networks in Classification," *Schedae Informaticae*, vol. 1/2016, 2017, doi: 10.4467/20838476SI.16.004.6185.