

# AWDMC-Net: Classification of Adversarial Weather Degraded Multiclass scenes using a Convolution Neural Network

Sourav Dey Roy, Mrinal Kanti Bhowmik\*

Department of Computer Science and Engineering, Tripura University (A Central University), Suryamaninagar 799022, Tripura (W), India

## ARTICLE INFO

Communicated by Nikos Paragios

### Keywords:

Atmospheric/weather conditions  
E-TUVD dataset  
Convolutional Neural Network  
Skip connections  
Pruning  
Classification  
Performance evaluation

## ABSTRACT

Computer vision systems in outdoor environments are strongly affected by different atmospheric/weather conditions. Therefore, understanding the actual behavior of outdoor scenes is necessary for effective removal and improvement of the overall performance of computer vision systems. Although the classification of atmospheric/weather conditions has been well explored, reporting on the same in multiclass problem using Convolutional Neural Networks (CNNs) has received very little attention. In response to address this disparity, we propose a new CNN architecture named the “Adversarial Weather Degraded Multi-class scenes Classification Network (AWDMC-Net)” for outdoor scene classification degraded by different atmospheric/weather conditions. The proposed network is based on adopting different combinations of skip connections in building blocks of CNN there after adaptively pruning the least important convolutional kernels from the network. For effective pruning, we proposed a new pruning criterion named “Entropy Guided Mean- $l_1$  Norm” that can adaptively evaluate the importance of convolutional kernels by considering the filters and their corresponding output feature maps. The prediction performance of our proposed model was evaluated on our newly designed E-TUVD (Extended Tripura University Video Dataset) and on publicly available benchmark datasets. Our newly created video dataset, E-TUVD, consists of 147 video clips (approximately 793800 frames) that represent six atmospheric/weather conditions, namely, fog, dust, rain, haze, poor illumination, and clear day conditions. Our proposed model achieves an accuracy of 93.85%, a specificity of 93.79%, and a sensitivity of 94.18% on our dataset, which outperforms the prevailing standard CNN models and recent state-of-the-art methods for atmospheric/weather classification tasks. Furthermore, our network also reduces the time consumption for atmospheric/weather classification tasks, and therefore mostly meets the requirements of practical applications in real-world scenarios.

## 1. Introduction

During the last few decades, human perceptibility to the appearance of color and contrast of outdoor scenes through the atmosphere has been well studied. While atmospheric/weather conditions have a strong impact on our everyday lives, they also alter the appearance of outdoor scenes (Narasimhan and Nayar, 2003). Generally, camera provides rich visual information regarding the presence of various candidate objects in the outdoor scenes and becomes one of the essential conventional sensors in the computer vision community (Shirmohammadi and Ferrero, 2014). Most computer vision systems for surveillance, traffic monitoring, augmented reality, vehicle navigation, and other assistant driving systems are currently designed to perform under clear visibility (Kroemer et al., 2010; Liu et al., 2016; Martinez-Martin and Pobil, 2017; Bore et al., 2018; Cai et al., 2021; Guo et al., 2018; Chen et al., 2012), but unfortunately this is not always the case. In general, the appearance of outdoor scenes is mainly altered for two reasons:

- **Atmospheric/Weather Effects:** The atmospheric/weather effect is caused by fog, dust, rain, and other factors that reduce visibility and cause images to lose contrast and color information. Such conditions basically alter the key characteristics (i.e., intensity, color, polarization, and coherence) of sunlight due to scattering by atmospheric particles.
- **Natural Light/Illumination Effects:** Conversely, in an outdoor environment, the illumination effect is mostly caused by variations in the intensity of sunlight at different times of the day.

Thus, the applicability of various computer vision systems, including outdoor video surveillance and vehicle aided driving systems, is strongly affected by various atmospheric/weather phenomenon (i.e., fog, haze, dust, rain, etc.) (Chen et al., 2012; Narasimhan and Nayar, 2002). This is because the several candidate objects (such as vehicles, pedestrian, animals, etc.) present in the scenes (so as acquired by conventional sensors (i.e., camera)) to be detected/recognized (which

\* Corresponding author.

E-mail address: [mrinalkantibhowmik@tripurauniv.ac.in](mailto:mrinalkantibhowmik@tripurauniv.ac.in) (M.K. Bhowmik).

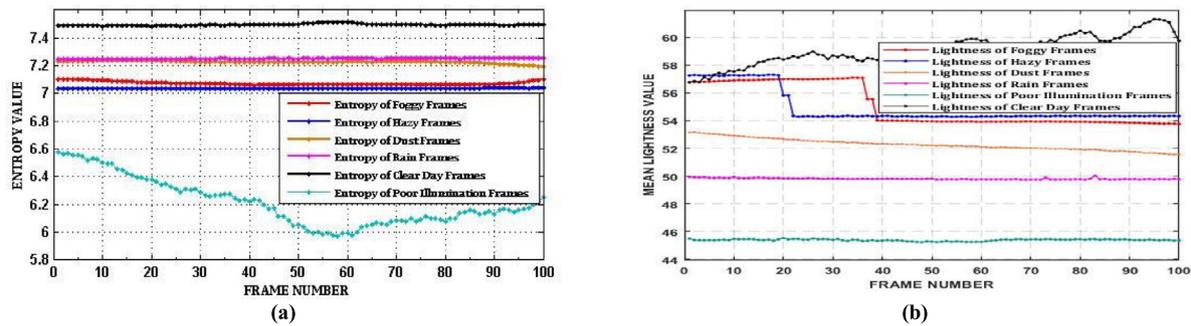


Fig. 1. (a) Entropy values of image frames at different atmospheric/Weather conditions; (b) Mean lightness values of image frames at different atmospheric/Weather conditions (All the frames in different atmospheric/weather conditions are selected from our E-TUVD Dataset (Roy and Bhowmik, 2020)).

is an integral part of many computer vision systems) have insufficient detailed appearance information due to the high loss in contrast in such adverse atmospheric/weather conditions. Therefore, this is a fundamental problem for many vision systems to be dealing with and has attracted a significant attention from the academic, research, and industrial communities. As these degradation conditions strongly affect the normal functioning of many computer vision systems, understanding their actual behavior on scene appearances provides valuable information for decreasing their influence on assistance systems. It is obvious that due to the presence of extreme atmospheric/weather conditions, light passing through the atmosphere is attenuated due to absorption and scattering (Narasimhan and Nayar, 2002). As a consequence, only a small portion of the light is reflected and reaches the observer, which in turn decreases the contrast and visibility in the image. However, the actual characteristics of absorption and scattering are complex and significantly depend on the types, orientations, sizes, and distributions of particles of atmospheric/weather conditions (Narasimhan and Nayar, 2002). For example, in the case of static atmospheric/weather conditions (i.e., fog, haze, and dust), constituent particles are too small (1–10  $\mu\text{m}$ ) (Garg and Nayar, 2004) to be visible to a camera, whereas the associated particles of dynamic atmospheric/weather conditions (i.e., rain and snow) are larger (0.1–3.5 mm) (Garg and Nayar, 2004), and are therefore visible in the image. Hence, the characteristics of atmospheric scattering models (i.e., composed of attenuation and air light) can be used to characterize the effects of static conditions but are not appropriate to describe the effect of dynamic conditions. Conversely, the study of dynamic weather conditions involves a stochastic model that acquires the effects of a large particle moving at high speeds (Garg and Nayar, 2004). This aspect indicates that the weather affecting an image must be recognized before any restoration processing is performed.

On the other hand, understanding the atmospheric/weather effects pertaining to outdoor scenes and their effective classification has promising prospects in a variety of applications, such as:

- **Border Security and Surveillance:** Essentially, North-Eastern states and other states of India share international borders. Commercialized cameras are installed by governments in border areas for continuous monitoring, but these cameras appear fragile under challenging atmospheric/weather degradation conditions. In such extreme situations, there is a chance for suspected intruders to enter foreign lands and perform illegal or suspicious activities that may go undetected by an unassisted vision system because of the substantial contrast loss.
- **Automatic Driving Systems:** Many industries are attempting to cope with challenges due to illumination/weather degradation conditions; however, such degraded conditions will still impair the vision capability of installed cameras and create confusing reflections that may result in problems for self-driving cars.

Additionally, atmospheric/weather conditions have a strong impact on air transport. A number of flight paths are often rerouted or canceled every year during the winter season or rainy seasons due to the poor visibility conditions experienced at different national/international airports in India and other countries across the globe. Other application areas include anomaly detection, route monitoring and obstacle identification in paths for robots (i.e., robotic vision) (Lins et al., 2015; Bodenhausen et al., 2014; Anghinolfi et al., 2013), and forest inventories for wild life monitoring and tracking (Zhang et al., 2016a; Kellenberger et al., 2019). Therefore, there is a need for the development of a vision based technique to study to what extent and in what sense such challenging visual conditions can be dealt with, in order to enable benefits for a broader range of computer vision applications. In real-world scenarios, understanding the effects of atmospheric/weather conditions on scene appearance appears to be a simple problem for human visual systems, but it is a difficult challenge for computer vision systems to decide whether a given image is foggy or clear day conditions. For example, Fig. 1 shows the characterization of image frames using entropy and mean lightness under different atmospheric/weather conditions (i.e., degradation due to fog, haze, dust, rain, poor illumination, and clear day conditions). The image frames are selected from our own created E-TUVD dataset (Roy and Bhowmik, 2020). It can be observed that the statistical feature values (i.e., entropy and mean lightness values) are higher under clear day conditions and lower under poor illumination conditions. However, these features are usually restricted to certain classes of atmospheric/weather conditions and cannot be applied to a diverse set of atmospheric/weather applications (i.e., in the case of multiclass atmospheric/ weather conditions). The same can be observed in Fig. 1, where the difference between these two feature values for the asymmetric characterization of fog and haze conditions is smaller. Thus, understanding the atmospheric/weather conditions pertaining to the scenes for effective restoration of the image quality in poor visibility conditions is a significant task in computer vision and increases the road safety, thereby monitoring the vehicle speed in response to real-time atmospheric/weather situations.

In recent years, significant contributions have been made as an effort to investigate the effectiveness of atmospheric/weather classification problems. Most of these approaches are either based on designing conventional handcrafted features or based on the application of Convolution Neural Networks (CNNs), as shown in Table 1. Although researchers have designed many handcrafted features (Chen et al., 2012; Roser and Moosmann, 2008; Lu et al., 2014; Kurihata et al., 2005; Pavlic et al., 2013; Zhang and Ma, 2015) for weather classification, these features are usually restricted to certain classes of atmospheric/weather conditions and cannot be applied to diverse ranges of atmospheric/weather applications. Meanwhile, most of these handcrafted feature based approaches as reported in Chen et al. (2012), Roser and Moosmann (2008) and Zhang and Ma (2015), may provide noticeable results for images acquired in controlled environments but miss the mark for atmospheric/ weather classification of images taken in uncontrolled environments, i.e., without any sky features. Recently,

Convolution Neural Networks (CNNs) have gained a significant impact in the area of feature extraction and have shown a noticeable performance in comparison to conventional approaches. Inspired by the outstanding performance of CNNs, several works have been conducted to adopt CNNs in atmospheric/weather classification problems (Lu et al., 2014; Elhoseiny et al., 2015; Guerra et al., 2018; Zhu et al., 2016; Lin et al., 2017). Most of these aforementioned techniques perform two class weather classification (Chen et al., 2012; Roser and Moosmann, 2008; Lu et al., 2014; Kurihata et al., 2005; Pavlic et al., 2013; Zhang and Ma, 2015; Elhoseiny et al., 2015). Although some researchers have reported the performance of CNNs for multiclass weather classification problems (Guerra et al., 2018; Zhu et al., 2016; Lin et al., 2017), these methods either used basic CNNs as a feature extraction method from pre-processed segmented images or combined hand-crafted weather features with basic CNN extracted features and classified using conventional classifiers.

Focusing on the importance of automatic understanding and classification of the atmospheric/weather effects using CNNs on scene appearance, our primary contributions include the following:

1. We proposed a new end-to-end network named AWDMC-Net (Adversarial Weather Degraded Multi-class scenes Classification Network) for the classification of different atmospheric/weather degraded outdoor scenes. The proposed network is based on the multiple skip connections on basic building blocks of CNN and adaptive pruning of convolutional kernels/filters. To the best of our knowledge, **this is the first application of a deep learning framework for designing a new architecture based on holistic image sequences for multiclass atmospheric/weather classification on real-time outdoor scenes.**
2. For efficient and effective pruning, we **proposed a new pruning criterion called “Entropy Guided Mean- $l_1$  Norm”**. While the current pruning criteria only consider only the filter weights for calculating the scores of the filters, the proposed pruning criterion adaptively considers the convolutional kernels/filters and their corresponding output feature maps for evaluating the filter/kernel importance in the network.
3. We conduct extensive experiments on available benchmarks and **our newly designed video dataset that contains 147 video clips (i.e., approximately 793 800 frames), each spanning 1–5 min of duration, named the “Extended Tripura University Video Dataset (E-TUVD)”**. The dataset contains six classes of atmospheric/weather degradation conditions. The dataset has more weather categories and a wider application range than those corresponding to other similar available datasets.
4. We also explored the **prediction performance of fifteen well-known CNN models pre-trained on the ImageNet Challenge Dataset (Deng et al., 2009) as a fixed feature extraction module and a fine tuning module** for the classification of atmospheric/weather degraded image sequences and compared them with our proposed AWDMC-Net model. The advantages and perception capability of CNNs to automatically extract image features are therefore fully explored in this manuscript for the classification of atmospheric/weather conditions.

**Paper Outline.** Section 2 provides a brief survey on different atmospheric/weather classification algorithms and the previous datasets used for atmospheric/weather classification tasks. Section 3 describes the E-TUVD dataset under different atmospheric/weather conditions. In Section 4, a detailed description of our proposed atmospheric/weather classification architecture (AWDMC-Net) is provided. Section 5 presents a thorough evaluation of the proposed architecture on the created dataset and other benchmark datasets, followed by a discussion on the performance comparison with respect to the state-of-the-art methods. Finally, the conclusion of the paper is provided in Section 6.

## 2. Related work

Numerous atmospheric/weather conditions have been classified depending on a vast range of computer vision algorithms. This section provides an overview of the most related work on atmospheric/weather classification, which can be considered a kind of scene recognition. Additionally, an overview of various public and private datasets used by researchers for atmospheric/weather classification tasks is explored in this section.

### 2.1. Overview of existing methodology

There are several works in the literature to construct atmospheric/weather recognition and classification algorithms for driving assistance. In this subsection, we categorized the atmospheric/weather classification algorithms into two major approaches: conventional feature/learning based approaches, and deep learning based approaches. An overview of various representative methods in each of the aforementioned approaches is shown in Table 1.

**Conventional Feature/Learning Based Approach.** The atmospheric/weather classification algorithms based on this approach use visual characteristics and features (i.e., local and global features) to recognize/ classify weather conditions, thereby dealing with variations in view, lighting, and of objects in scenes. In Kurihata et al. (2005), H. Kurihata et al. recognized raindrops using image features from principal component analysis (PCA) and further adopted template matching to detect raindrops. On the other hand, Roser and Moosmann (2008) used single monoscopic color images for the extraction of histogram based features from blocks defining regions of interests (ROIs). Here, five sets of histogram features, including local contrast, minimum brightness, sharpness, hue, and saturation, were extracted from ROI, and the weather conditions were classified using a support vector machine (SVM). In Yan et al. (2009), X. Yan et al. also used three sets of features, including the histogram of gradient amplitude (HGA), HSV color histogram (HHSV), and road information (RI), which are extracted from images captured by an in-vehicle vision system, and employed the Real AdaBoost classifier to achieve the weather classification task. Likewise, Zhao et al. (2011) developed a general framework to extract features signifying both the autocorrelation of pixel-wise intensities over time and the maximum directional length of rain streaks or snowflakes. Thereafter, the classification of different weather conditions using the C-SVM classifier designate the efficiency of the extracted features. In Chen et al. (2012), Z. Chen et al. developed a method for the classification of weather conditions using a support vector machine (SVM) classifier trained on the sky region features of panorama images. To avoid interference from non-sky regions, thresholding based segmentation was used. Then, a set of seven features are extracted, including Scaling Invariant Feature Transform (SIFT), Local Binary Pattern (LBP), Hue (H), Saturation (SA), Brightness (BR), Gradient Magnitude Computed from Sobel Operator (GM), and residual computed from reference image (M). For a further increase in the classification accuracy, multi-pass active learning was adopted to choose the training set. For instance, Li et al. (2014) used a set of four features, such as the power spectrum slope (PPS), contrast, noise, and saturation, to classify weather phenomena where every SVM classifier on the non-leaf node of the decision tree was constructed. In Song et al. (2014), H. Song et al. used five sets of features, including inflection point information (IIP), Power Spectrum Slope (PPS), Edge Gradient Energy (EGE), contrast, noise, and saturation, to discriminate multiple weather conditions. Then, weather classification was performed using the K-Nearest Neighbor (KNN) classifier based on these extracted feature sets. In Zhang and Ma (2015), Z. Zhang et al. used a learning based method for the classification of multiclass weather images. The method uses multiple local and global features to represent various weather conditions. For instance, channels a and b in the LAB color space of the sky region are used to specify sunny weather, Histogram of Gradient

(HOG) is used to specify rainy weather, Histogram of Intensity (HOI) is used to specify snowy weather, and the Dark Channel Prior (DCP) method (He et al., 2011) is used to specify haze weather. Additionally, some global features (i.e., contrast and saturation) are used to discriminate multiclass weathers. Then, multiple kernel learning (MKL) is used to learn an adaptive classifier and fuse these local and global features, thereby classifying the weather conditions. Likewise, Zhang et al. (2016b) further extended their previous weather classification method (Zhang and Ma, 2015), where selection and classification of multiple weather features were performed via learned sparse code followed by the MKL algorithm. In Chu et al. (2017), W.T. Chu et al. observed the correlations between the properties of various weather conditions and developed computational models based on a random forest classifier for weather recognition. Furthermore, in Wang et al. (2018), S. Wang et al. fused the real-time meteorological weather data (i.e., wind speed (WS), temperature (TEMP), humidity (HUM), and time (TM)) with the image features (i.e., Dark Channel (DC)) to recognize different weather conditions using the Adaboost classifier.

**Deep Learning Based Approach.** Deep learning is often considered a neural network with deep structures. The notion behind the concept of neural networks originated in the 1940s (Pitts and McCulloch, 1947) and became widespread with the advancement of the back-propagation algorithm by Rumelhart et al. (1986). Similar to many other topics, several methods have been established in the literature for atmospheric/weather image recognition/classification tasks, thereby utilizing the convolution structure of deep models. Elhoseiny et al. (2015) adopted CNN to resolve the task of weather classification where transfer learning of the AlexNet model pre-trained on an ImageNet Challenge Dataset (Deng et al., 2009) was used, thereby replacing the output layer with two neurons, i.e., one representing cloudy conditions and the other representing sunny conditions. In addition, Zhu et al. (2016) used three existing network structures, AlexNet, modified AlexNet, and GoogLeNet, as fine-tuning modules and without fine-tuning modules for weather classification. In Lu et al. (2014), C. Lu et al. constructed 4717 dimension feature vectors consisting of CNN features and five handcrafted weather features. The CNN features corresponding to each weather degraded image were computed using AlexNet, and five handcrafted features were computed defining sky and shadow regions (Lalonde et al., 2010), contrast, reflection, and haze in an image. The feature vectors concatenating six feature components corresponding to each weather degraded image were used for training the latent SVM framework to classify the weather conditions. For instance, Li et al. (2017) developed a multi-task CNN framework to concurrently address the weather classification task. The framework consists of two parts: weather cue semantic segmentation using Shelhamer et al. (2017) and weather classification using fine tuning of the VGG-16 layers. In Lin et al. (2017), D. Lin et al. highlights the significance of regional cues for the construction of features Table 2. Overview of Previous Datasets Used for Atmospheric/Weather Classification Tasks in Comparison to our E-TUVD Dataset related to the classification of weather conditions. A region selection and concurrency model (RSCM) was proposed based on a deep CNN to overcome these tasks. First, the images were segmented into different regions using Joulin et al. (2012), Rubio et al. (2012) and Joulin et al. (2010), and then for each region a latent variable was used to specify the region containing discriminating weather related information. Second, weather assessment/recognition was performed using the Siamese architecture of neural networks (Chopra et al., 2005; Hadsell et al., 2006), thereby independently training the RSCM for each considered weather class. Similarly, Guerra et al. (2018) used state-of-the-art pre-trained CNNs (i.e., CaffeNet, PlacesCNN, ResNet-50/101/152, and VGG-16/19) to extract deep features from super pixel bounding masks of weather images. Then, the extracted features from each individual pre-trained CNN were classified independently using the SVM classifier. In Shi et al. (2018), Y. Shi et al. used Mask R-CNN (He et al., 2017) to extract the foreground regions and foreground

edges in the visual weather images and further superimposed them into the same-scale matrix to classify weather conditions using the VGG-16 model. In addition, An et al. (2018) used AlexNet and ResNet for deep feature extraction and combined a multiclass SVM to perform weather classification. In Zhao et al. (2018), B. Zhao et al. presented a CNN-RNN architecture consisting of a CNN (convolutional/pooling layers of VGG-16 Russakovsky et al., 2015) to extract discriminative features, a channel-wise attention module to recalibrate feature responses, and a convolutional Long Short-Term Memory (LSTM) to estimate the relationships and classify different atmospheric/weather conditions based on their labels. Ibrahim et al. (2019) proposed a deep CNN framework named 'WeatherNet'. The framework is based on transfer learning of the ResNet-50 model and comprises of four deep CNN models, i.e., NightNet, GlareNet, Precipitation Net, and FogNet, to recognize dawn/dusk, day, night-time, glare, rain, snow, and fog, respectively. In Zhao et al. (2019), B. Zhao et al. proposed a multi-task framework for classification of five weather conditions (i.e., sunny, cloud, fog, rain, and snow). The proposed framework is divided into four modules i.e., the shared layers, the classification branch, the segmentation branch and the fusion of classification feature and segmentation feature. In this framework, the share layers are similar to the five consecutive convolutional/pooling layers of VGG-16 (Russakovsky et al., 2015) with incorporated average pooling and are further used for both weather cue segmentation and classification of weather conditions. Then, the share layers are divided into weather classification and segmentation tasks. The weather-cue segmentation module comprises shared layers followed by convolutional and deconvolutional layers where the features from the preceding layers are taken into consideration as proposed by Long et al. (2016). Finally, the weather-cues the segmentation features are fused with the classification features for weather classification tasks. Xia et al. (2020) proposed a CNN model entitled 'ResNet15' for classifying and recognizing weather images based on fine-tuning of the ResNet-50 CNN architecture. Also, Al-Haija et al. (2020) used self-reliant framework based on transfer learning of ResNet-18 CNN architecture pre-trained on ImageNet Challenge Dataset (Deng et al., 2009) for four-class atmospheric/weather classification task including sunrise, shine, rain, and cloudy.

## 2.2. Overview of atmospheric/weather classification dataset

The rapid design of algorithms for atmospheric/weather classification tasks originates from the accessibility of benchmark datasets defining various representative challenges under such degraded conditions. Several benchmark datasets are designed to evaluate the performance of atmospheric/weather classification methods. In this subsection, a review of most of the related datasets used by the research communities for atmospheric/weather classification is presented. Each of the available datasets are image based datasets. Among these datasets, the RESIDE- $\beta$ /RESIDE-Standard dataset (Li et al., 2018) contains 90 967 images in two classes of weather conditions i.e., haze and clear day conditions. Conversely, the WeatherNet dataset (Ibrahim et al., 2019) contains a diverse set of 23 865 images in seven different weather situations, including day, night, glare, fog, rain, snow, and clear conditions. In addition, MWD (Lin et al., 2017) provides a set of 65 000 images under six weather conditions: sunny, clear, rain, snow, haze, and thunder. To form a large-scale weather dataset, Image2Weather (Chu et al., 2016) contains 255 837 images in sunny, cloudy, snow, rain, and fog weather conditions. Weather2Dataset (Zhu et al., 2016) contains 16 635 images in four different atmospheric/weather conditions, including sunny, rain, blizzard, and fog situations. Additionally, the MWI dataset (Zhang and Ma, 2015) contains 20 387 images in four different atmospheric/weather conditions (i.e., sunny, rain, snow, and haze situations). On the other hand, RFS dataset (Guerra et al., 2018) provided 3300 images under three weather conditions (i.e., rain, fog, and snow conditions), and WeatherDataset-4 (Xia et al., 2020) provided 7983 weather degraded images under fog, rain, snow, and sunny conditions. Rather than elaborately describing the existing datasets, the key highlights of the datasets are summarized in Table 2.

**Table 1**  
Overview of previous works for atmospheric/weather classification/Recognition.

Approach	Authors/Year	Method used	Category/ Conditions	Dataset	Performance evaluation	
Conventional feature/ Learning based approach	H. Kurihata et al./2005	Kurihata et al. (2005)	Principal Component Analysis (PCA)	BC/R	Private Dataset	Re-0.24; Pr-0.87
	M. Roser et al./2008	Roser and Moosmann (2008)	Histogram Features (BR,C,SH,SA,H), and SVM	MC/CD;HR;LR	Private Dataset	CE- 5.2%
	X. Yan et al./2009	Yan et al. (2009)	Features (HGA,HHSV,RI), and Adaboost	MC/S;CL;R	Private Dataset	Acc- 90.62%
	X. Zhao et al./2011	Zhao et al. (2011)	Features (AC,MB,SA), and C-SVM	MC/F;R;SN;IV	Private Dataset	Acc- 92.37%
	Z. Chen et al./2012	Chen et al. (2012)	Sky Region Extraction, Features (SIFT, LBP, H,SA,V,GM,M), MKL, and SVM	MC/S;CL;OV	Private Dataset	Acc- 95.00%
	Q. Li et al./2014	Li et al. (2014)	Features (PPS,C,N,SA), DT, and SVM	MC/CD;OV;F;R	WILD and Private Dataset	CE- 10.64%
	H. Song et al./2014	Song et al. (2014)	Features (IIP,PPS,EGE,C,SA,N), and KNN	MC/S;SN;F;R	Private Dataset	Acc- 90.00%
	Z. Zhang et al./2015	Zhang and Ma (2015)	Local (LAB,HOG,HOI,DCP) and Global Features (C,SA), MKL, and Adaptive Classifier	MC/S;SN;R;H	MWI Dataset	Acc- 0.5944
	Z. Zhang et al./2016	Zhang et al. (2016b)	Local (LAB,HOG,HOI,DCP) and Global Features (C,SA), DL, MKL, and SVM	MC/S;SN;R;H	MWI Dataset	Pr- 0.73; Re- 0.71; FM- 0.72; Acc- 0.71
	W.T. Chu et al./2017	Chu et al. (2017)	Correlation, and Random Forest Classifier	MC/S;CL;SN;R;F	Image2Weather	Acc- 80.00%
S. Wang et al./2018	Wang et al. (2018)	Weather Features (WS, TEMP, HUM, TM), Image Features (DCP), and Adaboost	MC/S;R;H;SN	Private Dataset	Acc- 83.09%	
Deep learning based approach	M. Elhoseiny et al./2015	Elhoseiny et al. (2015)	AlexNet where the output layer replaced with two nodes	BC/CL;S	Private Dataset	Acc- 82.20%
	Z. Zhu et al./2016	Zhu et al. (2016)	AlexNet, Modified AlexNet, and GoogleNet (With and Without Fine Tuning)	MC/S;RS;B;F	Weather Dataset	Acc- 94.50%
	C. Lu et al./2014	Lu et al. (2014)	Handcrafted Features (SRD (SIFT and Mean HSV), SD-R (Wang et al., 2018), IM, C, DCP), CNN Feature (AlexNet), SVM	BC/CL;S	Sun, Labelme, and Flickr Dataset	Acc- 91.4%
	X. Li et al./2017	Li et al. (2017)	Weather-Cues Semantic Segmentation using (Rumelhart et al., 1986) and Classification using Fine Tuning of VGG-16	BC/CL;S	Dataset (Lu et al., 2014)	Acc- 90.70%
	D. Lin et al./2017	Lin et al. (2017)	Region Selection and Concurrency Model (RSCM)	BC/S;CL;R;S;N;H;T	MWD Dataset and Dataset (Lu et al., 2014)	Acc- 90.07%
	J.C.V. Guerra et al./2018	Guerra et al. (2018)	Super Pixel Mask, CNN Features (ResNet50/101/152, Places CNN, VGG-16/19, CaffeNet), and SVM	MC/R;F;SN	RFS Dataset	mAP- 68% to 81%
	Y. Shi et al./2018	Shi et al. (2018)	Mask R-CNN, and VGG-16	MC/S;F;SN;R	Private Dataset	Acc- 94.71%
	J. An et al./2018	An et al. (2018)	CNN Features (AlexNet, and ResNet), and SVM	MC/S;C;H;SN	Dataset (Lu et al., 2014) (Liu et al., 2018), and D-Hazy (Ancuti et al., 2016)	Acc- 95.00%
	B. Zhao et al./2019	Zhao et al. (2018)	CNN-RNN (Recurrent Neural Network) Architecture	MC/S;CL;F;SN;MO;R	TA Dataset and MLWC Dataset	Pr- 0.92; Re- 0.82; F1-0.87;
	M.R. Ibrahim et al./2019	Ibrahim et al. (2019)	Parallel Deep CNN models via Transfer Learning of ResNet-50 Model	MC/ D;N;G;F;R;SN;CD	WeatherNet Dataset	Pr- 0.90; Re- 0.87; F1-0.88;Acc- 93.8%
B. Zhao et al./2019	Zhao et al. (2019)	Weather Cue Segmentation using (Long et al., 2016) and Classification (VGG-16)	MC/S;CL;F;R;SN	Dataset (Lu et al., 2014) and Private Dataset	Acc- 95.09%	
J. Xia et al./2020	Xia et al. (2020)	ResNet15 (Fine Tuning of ResNet50)	MC/F;R;SN;S	WeatherDataset-4	Acc- 88.30%	
Q.A. Al-Haija et al./2020	Al-Haija et al. (2020)	Transfer learning of ResNet-18	MC/SR;SH;R;CL	Multi-class Weather Dataset	Acc- 98.22%	

BC—Binary Classification; MC—Multi-Class Classification; F—Fog Condition; H—Haze Condition; D—Dust Condition; R—Rain Condition; HR—Heavy rain; LR—Light Rain; PI—Poor Illumination Condition; LL—Low Light Condition; CD—Clear Day Condition; S—Sunny Day Condition, SN—Snow Condition; CL—Cloudy Condition; T—Thunder; B—Blizzard; OV—Overcast; IV—Illumination Variation; D—Day; N—Night; RS—Rain Storm; G—Glare; MO—Moist; SR—Sunrise; SH—Shine; SIFT—Scaling Invariant Feature Transform; LBP—Local Binary Pattern; H—Hue; SA—Saturation; V/BR—Brightness; GM—Gradient Magnitude Computed from Sobel Operator; M—Motion (residual computed from reference image); LAB—a and b channels in the LAB color space; HOG—Histogram of Gradient; HOI—Histogram of Intensity; DC/ DCP—Dark Channel/ Dark Channel Prior; HGA—Histogram of Gradient Amplitude; HHSV—Histogram of HSV Color Space; PPS—Power Spectrum Slope; N—Noise; DL—Dictionary Learning; MKL—Multiple Kernel Learning; DT—Decision Tree; SVM—Support Vector Machine; IIP—Image Inflection Point; EGE—Edge Gradient Energy; KNN—K-Nearest Neighbor; C—Contrast; SH—Sharpness; WS—Wind Speed; TEMP—Temperature; HUM—Humidity; TM—Time; SRD—Sky Region Detection; HSV—HSV Color Space; SD-R—Shadow Detection and Ranking; IM—Image Matting; TA Dataset—Transient Attribute Dataset; MLWC—Multi-Label Weather Classification Dataset; Acc—Accuracy; Pr—Precision; Re—Recall; F1/FM—F-Measure; CE—Classification Error; mAP—mean Average Precision.

**Table 2**  
Overview of previous datasets used for atmospheric/weather classification tasks in comparison to our E-TUVD dataset.

Dataset		Weather conditions	Type of dataset	No. of Images/ Frames/Videos	Resolution (in Pixels)	Format	Availability
RESIDE- $\beta$ / RESIDE-Standard	Li et al. (2018)	H/CD	IBD	90 967	[178 $\times$ 178] to [620 $\times$ 460]	.jpg/ .png	A
Weather2Dataset	Zhu et al. (2016)	S/R/B/F	IBD	16 635	NP	NP	NA
RFS	Guerra et al. (2018)	R/F/SN	IBD	3300	[560 $\times$ 420]	.jpg	A
MWD	Lin et al. (2017)	S/CL/R/SN/H/T	IBD	65 000	[3456 $\times$ 1801] to [300 $\times$ 121]	.jpg	A
MWI	Zhang and Ma (2015)	S/R/SN/H	IBD	20 387	NP	NP	A
WeatherNet	Ibrahim et al. (2019)	DA/N/G/F/R/SN/CD	IBD	23 865	NP	.jpg/ .png	NA
Image2Weather	Chu et al. (2016)	S/CL/SN/R/F	IBD	255 837	NP	NP	NA
WeatherDataset-4	Xia et al. (2020)	F/R/SN/S	IBD	7983	[256 $\times$ 256]	.jpg	A
<b>E-TUVD [Our Dataset]</b>		F/H/D/R/PI/CD	FBD	793 800 (approx.)	[1920 $\times$ 1080]	.jpg	A

F—Fog Condition; H—Haze Condition; D—Dust Condition; R—Rain Condition; PI—Poor Illumination Condition; LL—Low Light Condition; CD—Clear Day Condition; S—Sunny Day Condition, SN—Snow Condition; CL—Cloudy Condition; T—Thunder; B—Blizzard; DA—Day; N—Night; G—Glare; FBD—Frame Based Dataset; IBD—Image Based Dataset; A—Available; NA—Not Available; NP—Not Provided.

### 2.3. Significant gaps and our contributions to these aspects

Although the algorithms and datasets reviewed in Tables 1–2 have advanced the research for atmospheric/weather classification, they have several limitations:

- As found in the literature, most of the research works for atmospheric/weather classification and recognition were performed on public/private datasets (as shown in Table 2). However, these datasets are restricted in either scale or quantity for training traditional deep learning models. Although the Image2Weather dataset (Al-Hajja et al., 2020) contains 255 837 images for weather classification tasks, this dataset is not publicly accessible to the research community. Moreover, to date, no weather dataset is available for the research community containing dust conditions as one of the pre-defined challenges, which is also one of the major causes of atmospheric/weather phenomena in outdoor settings. In contrast, our E-TUVD (Roy and Bhowmik, 2020) dataset contains a diverse set of atmospheric/weather challenges, including dust conditions, for evaluating standard baseline models in such complex situations. Additionally, compared to previous datasets, our dataset contains more frames depicting urban scenes in each of the considered weather classes.
- A comprehensive survey of the literature has revealed the usefulness of atmospheric/weather classification tasks in a vast range of computer vision applications. Although different national and international researchers proposed and published various research papers (as shown in Table 1) for atmospheric/weather classification tasks using CNNs, most of these research papers either used the existing CNNs pre-trained on the ImageNet Challenge Dataset (Deng et al., 2009) as a fine-tuning/transfer learning module (Elhoseiny et al., 2015; Ibrahim et al., 2019; Zhao et al., 2019) or used them as a feature extraction method or fused conventional features with basic CNN based features, and classified them using conventional classifiers (Guerra et al., 2018; An et al., 2018). Although very few works have been proposed in the literature for establishing new CNN based frameworks/algorithms for atmospheric/weather classification (Lin et al., 2017; Li et al., 2017; Shi et al., 2018; Zhao et al., 2019), these algorithms involve various pre-processing techniques for region detection (i.e., sky and boundary region detection), which makes the framework highly reliant on the success of the pre-processing steps.

To advance this area, in our present work different skip connections and pruning criteria on building blocks of CNNs are adopted to design a novel architecture named AWDMC-Net, for the classification of atmospheric/weather degraded image sequences.

### 3. Dataset description

The existence of atmospheric/weather challenges in outdoor scenes has remained a significant research gap for both computational photography and high-level computer vision tasks. Advancement in this area, thereby focusing on such degraded conditions, will directly benefit various blooming computer vision applications, including video surveillance and autonomous driving. To develop and test complex computer vision algorithms in extreme atmospheric or weather degraded conditions, a benchmark dataset is required to define such complex real-world scenarios. In the literature, various image based datasets are proposed to meet the specific necessities of atmospheric/weather classification tasks (as reviewed in Table 2). However, as atmospheric/weather challenges have a strong impact on various high-level computer vision tasks, our motive is to provide a video based dataset comprised of several salient objects in adversarial atmospheric/weather conditions.

The E-TUVD dataset is a pioneering annotated video dataset for detecting moving objects in degraded atmospheric/weather conditions (i.e., Fog, Haze, Dust, Rain and Poor Illumination). Other than the abovementioned five atmospheric or weather degraded conditions, E-TUVD also contains video clips of clear days. The current dataset consists of 147 video sequences (approximately 793 800 frames) that were captured in a variety of atmospheric/weather situations. A detailed description of the dataset video recording conditions, dataset information, and ground-truth annotation related to E-TUVD were elaborately discussed in our previous work (Roy and Bhowmik, 2020). Each frame of the E-TUVD contains various types of moving objects; in addition, the scenes were acquired mostly in urban areas, which are subjected to larger surface variations because of the existence of objects such as trees, houses, warehouses, office buildings, streets, and residents. All the characteristics of the E-TUVD dataset point to its importance in the realm of moving object identification and high-level vision tasks in outdoor situations with a variety of atmospheric/weather-related problems. The dataset is available for the research community in E-TUVD (2020). In our present work, we used the frames from E-TUVD for atmospheric/weather classification.

The overall sample sizes of the E-TUVD dataset in different atmospheric/weather conditions used for training, validation, and testing of our proposed AWDMC-Net model are shown in Table 3. From the E-TUVD dataset, we have selected 60 000 frames randomly consisting of balanced classes of atmospheric/weather conditions (i.e., Fog, Haze, Dust, Rain, Poor Illumination, and Clear Day). For effective training and testing of the proposed AWDMC-Net model with respect to the atmospheric/weather degraded conditions, frames from the video clips in each of the considered atmospheric/weather conditions of the E-TUVD dataset are selected randomly thereby maintaining certain timestamps between the consecutive frames so as to reduce the redundancy between the consecutive frames. Moreover, the consecutive

**Table 3**  
Sample size of E-TUVD used for Training, Validation, and Testing of the proposed AWDMC-Net model.

Category	Atmospheric/Weather conditions						Total
	Foggy condition	Haze condition	Dust condition	Rain condition	Poor illumination	Clear day	
Training and validation set	8000	8000	8000	8000	8000	8000	48 000
Testing set	2000	2000	2000	2000	2000	2000	12 000
<b>Total number of frames</b>	<b>10 000</b>	<b>10 000</b>	<b>10 000</b>	<b>10 000</b>	<b>10 000</b>	<b>10 000</b>	<b>60 000</b>

frames used for our proposed AWDMC-Net model are different with respect to the presence of several salient and candidate objects in the atmospheric/weather degraded scenes. For effective labeling (i.e., annotation) of the captured video clips of the E-TUVD dataset with respect to the atmospheric/weather categories, five research members of the research laboratory were selected and trained to ensure uniformity in labeling of atmospheric/weather conditions. During the annotation, certain guidelines are provided to the annotators related to what to label, what are the class labels, and how to assign class scores. Each of the annotators are asked to freely view the video clips assigned to them and are asked to assign the class scores [i.e., either 0 to 1] for each video clip belonging to each of the atmospheric/weather conditions. Also, the annotators were instructed that depending on their perception of the atmospheric/weather conditions pertaining to the scenes, they have to assign the class label '1' to that perceived atmospheric/weather condition and the remaining atmospheric/weather conditions need to assign the class label '0'. Moreover, the annotators were also instructed to assign class label '1' to only one of the considered atmospheric/weather conditions of the E-TUVD dataset. Depending upon this protocol, five different atmospheric/weather class labels will be received from five different annotators for each of the video clips of the E-TUVD dataset. Once the five reference class labels for each of the video clips of E-TUVD dataset were obtained, the final class labels for each of the video clips of the E-TUVD dataset is generalized depending upon the maximum voting policy scheme. This is generally performed using a threshold value,  $T$ . Thus the final class label of a particular video clip of the E-TUVD dataset is considered as a fog condition, if at least  $T$  (in our study we have considered the value of  $T$  as 3) number of research members included it as a fog condition (i.e., thereby assigning class label 1 for the fog condition) and so on for the other conditions of the E-TUVD dataset. As a result, the final class label is the result of the most comprehensive agreement among all research members (i.e., annotators). In general, atmospheric/weather phenomena are described and measured by several variables of the earth's atmosphere, such as temperature, humidity, dew point, wind speed, wind direction, atmospheric pressure, and visibility, according to the World Meteorological Organization (WMO) network. The variations and interactions of those metrological variables/parameters, and how they change over time can provide valuable information regarding atmospheric information and analysis of the acquired data. Each video clip of E-TUVD also holds the convenient ground truth information related to these atmospheric variables/parameters and so on obtained from the Regional Meteorological Department of Tripura. Finally, the atmospheric/weather labels assigned to each of the video clips of E-TUVD dataset based on maximum voting policy scheme are validated with the associated meteorological weather information (i.e., atmospheric variables/parameters) of the E-TUVD dataset.

Some of the sample frames of the E-TUVD dataset representing diverse sets of atmospheric/weather degradation conditions are displayed in Fig. 2.

#### 4. Network architecture and implementation

In this section, we propose our atmospheric/weather classification network, AWDMC-Net, which stands for the "Adversarial Weather Degraded Multi-class scenes Classification Network". The sequence of steps we adopted to achieve the atmospheric/weather classification task is described in the upcoming subsections.

##### 4.1. Ablation study

Depending upon the number of learnable parameters, different CNN architectures vary from each other in their characteristics (Chen et al., 2017; Yamashita et al., 2018). This can be observed by tuning the kernels and their sizes and the number of neurons in fully connected layers. In our proposed study, we have studied the characteristics of various CNN architectures on E-TUVD based on tuning the network parameters (i.e., the number of learnable parameters, network depth, number of kernels in each layer, etc.). The comparative analysis of different prototype networks developed and the training accuracy acquired in each specific case on the E-TUVD dataset are shown in Table 4. In each of the specific cases, neurons in each of the fully connected layers are kept fixed in order to avoid the abrupt increase in the number of learnable parameters. In Table 4, it has been reported that there is high risk for model overfitting if the number of learnable parameters is greater than or equal to the number of considered samples for training the prototype architectures. Additionally, it can be observed from Table 4 that the maximum accuracy is obtained for model 4. This is because the number of kernels in each layer of our proposed architecture was selected in a way so that the total number of training samples (i.e., atmospheric/weather degraded image sequences/frames) was comparatively greater than the total number of learnable parameters. Therefore, model 4 comprised of five sequentially stacked convolution blocks is used in our proposed architecture. The elaborate representation of our proposed architecture is described in the next subsection.

##### 4.2. Our proposed AWDMC-Net architecture

Based on the characteristics of our dataset, the block diagram of our proposed CNN architecture for the classification of outdoor scenes degraded by different atmospheric/weather degradation conditions is shown in Fig. 3. The multilayer structure of our proposed CNN can abstract the atmospheric/weather degraded image frames/sequences layer by layer in order to acquire a higher level of discriminating feature expression, which significantly affects the efficacy of the atmospheric/weather classification. Each of the network components is elaborated next.

##### Network Overview.

Let  $\mathbf{I} : \Omega \rightarrow \mathfrak{R}^{d \times n_T}$  represent input RGB atmospheric/weather degraded image sequences each having a size of  $112 \times 112$  pixels, and let  $\Omega^+$  be the open bounded subset such that  $\Omega \subset \mathfrak{R}^{d \times n_T}$ . Additionally,  $\mathbf{O}_T \rightarrow \mathfrak{R}^{N_c \times n_T}$  represents the corresponding label matrix for each image sequence ( $\mathbf{I}$ ), where  $n_T$  represents the number of training samples,  $d$  denotes the dimension, and  $N_c$  denotes the number of classes. Each of the network components is elaborated next.

##### Convolutional Layer.

The convolutional layer is the fundamental building component of CNN architectures. In this layer, the number of kernels (also sometimes termed as filters) with adjustable weights and biases are used to extract the discriminating features and represent them as an output feature map. Let us consider  $\mathbf{I}_{\text{FM}}^{n-1}$  as an input feature map (i.e., output feature map from the  $n-1$  layer),  $\mathbf{C}_k^n$  as the convolutional kernel, and  $b_k^n$  as the corresponding bias at the  $n$ th layer. Then, for a  $k$ th output feature map ( $\mathbf{I}_{\text{FM}}^n$ ) at  $n$ th layer, the  $r$ th receptive field inputted from the  $n-1$  layer is convolved with the  $k$ th kernel of the  $n$ th layer, and consequently

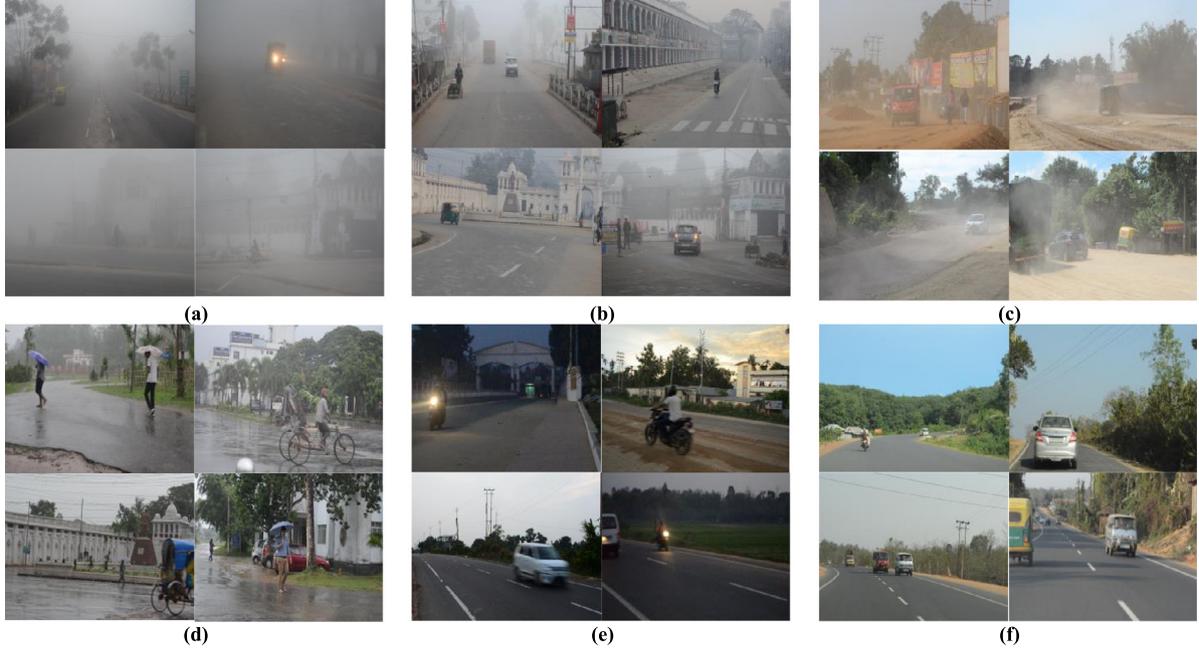


Fig. 2. Sample frames of the E-TUVD dataset in different atmospheric/Weather degradation conditions (a) Foggy conditions; (b) Haze conditions; (c) Dust conditions; (d) Rain conditions; (e) Poor illumination conditions; (f) Clear day.

Table 4  
Different CNN architectures (Prototypes) studied and their performance on E-TUVD for atmospheric/weather classification.

Model No	No. of layers	No. of convolution layer	No. of convolution filters in each layer	Kernel size of convolution layers	No. of pooling layers	Kernel size of the pooling layers	No. of fully connected layers	Neurons in each layer	No. of learnable parameters	Learning accuracy (%)
1	27	4	64,32,16,16	(3,3)	2	(3,3),(3,3)	3	100,50,6	263 612	68.97
2	30	5	16,32,64,32,16	(3,3)	2	(3,3),(2,2)	3	100,50,6	576 598	75.13
3	32	5	64,128,128,64,32	(3,3)	3	(3,3),(2,2),(2,2)	3	100,50,6	586 726	78.35
4	32	5	16,32,64,32,16	(3,3)	3	(2,2),(2,2),(2,2)	3	100,50,6	372 318	90.40
5	33	6	32,64,128,64,32,16	(3,3)	2	(3,3),(2,2)	3	100,50,6	720 406	78.97
6	33	6	16,32,64,64,128,128	(3,3)	2	(3,3),(2,2)	3	100,50,6	4 441 222	32.06
7	32	5	16,32,64,32,16	(3,3)	3	(3,3),(2,2),(2,2)	3	100,50,6	182 348	54.37
8	35	6	8,16,16,32,32,64	(3,3)	3	(3,3),(2,2),(2,2)	3	100,50,6	1 302 038	63.88
9	37	6	128,64,32,32,64,128	(3,3)	4	(3,3),(2,2),(2,2),(2,2)	3	100,50,6	408 892	81.02
10	40	7	8,16,32,32,64,64,128	(3,3)	4	(2,2),(2,2),(2,2),(2,2)	3	100,50,6	783 654	77.05

added the bias. The process is repeated in each convolution layer of the proposed architecture, and with the same shown in Eq. (1). Then, the final output is passed through batch normalization  $\mathcal{B}(\bullet)$  followed by a non-linear activation function  $\xi(\bullet)$ .

$$\mathbf{I}_{\text{FM}_k}^n = \xi \left( \mathcal{B} \left( \sum_{r=1}^m \mathbf{I}_{\text{FM}_r}^{n-1} * \mathbf{C}_{kr}^n + b_k^n \right) \right) \quad (1)$$

In our proposed network, we have used a Rectified Linear Unit (ReLU) (Nair and Hinton, 2010) as the activation function  $\xi(\bullet)$  stacked with each convolution layer, where the negative values in the output of the convolutional layers are replaced by zeros and the positive values remain unchanged, which may be expressed as shown in Eq. (2):

$$\xi(\mathbf{I}_{\text{FM}_k}^n) = \max(0, \mathbf{I}_{\text{FM}_k}^n) \quad (2)$$

In our proposed network, stacks of  $\text{CB}_N$  convolution layers are used sequentially, where  $N \in \{1, 2, \dots, 5\}$ . Each of the  $\text{CB}_N$  layers has  $C_k$  kernels, where  $k \in \{16, 32, 64, 32, 16\}$  and the size of each kernel ( $C_k$ ) is set to  $3 \times 3$  respectively.

#### Batch Normalization Layer.

To mitigate the stability of the network and greatly quicken the training process, batch normalization layers  $\mathcal{B}(\bullet)$  are used in our proposed network after each convolution layer, and also after first and second fully connected layers as illustrated in Fig. 3. This layer normalizes

the output feature maps of the convolutional layer ( $\mathbf{I}_{\text{FM}_k}^n$ ) by subtracting the mean ( $\text{MU}_\psi$ ) of mini-batch  $\psi$  where  $\psi \in \{1, 2, 3, \dots, 32\}$ , and dividing by the batch standard deviation  $\text{SD}_\psi^2$ . The same is shown in Eq. (3).

$$\mathcal{B}(\mathbf{I}_{\text{FM}_k}^n) = \mathbf{I}_{\text{FM}_k}^n - \text{MU}_\psi / \sqrt{\text{SD}_\psi^2 + \epsilon} \quad (3)$$

Here,  $\epsilon = e^{-05}$  is a constant value used to avoid complex values.

#### Pooling Layer.

This layer independently performs a down sampling operation over each activated feature map, there by reducing the dimensionality of the feature map (i.e., the output of the activation layer), and decreasing the number of subsequent learnable parameters. In our proposed network, we have used the maximum pooling operation (Chen et al., 2017) after the first, third and fifth convolution blocks (i.e.,  $\text{CB}_1$ ,  $\text{CB}_3$ , and  $\text{CB}_5$ ), as shown in Fig. 3. Here, the maximum activation value from each non-overlapping rectangular window ( $W_x \times W_y$ ) of the input feature map ( $\mathbf{I}_{\text{FM}_k}^n$ ) is kept by excluding all others, where  $\{W_x, W_y\} \in 2$ . Thus, the layer quickly converges the training accuracy by choosing the discriminative and position invariant features.

#### Fully Connected Layer.

After manifold blocks of convolutional, activation and pooling layers, the final feature maps corresponding to each image sequence (I) are

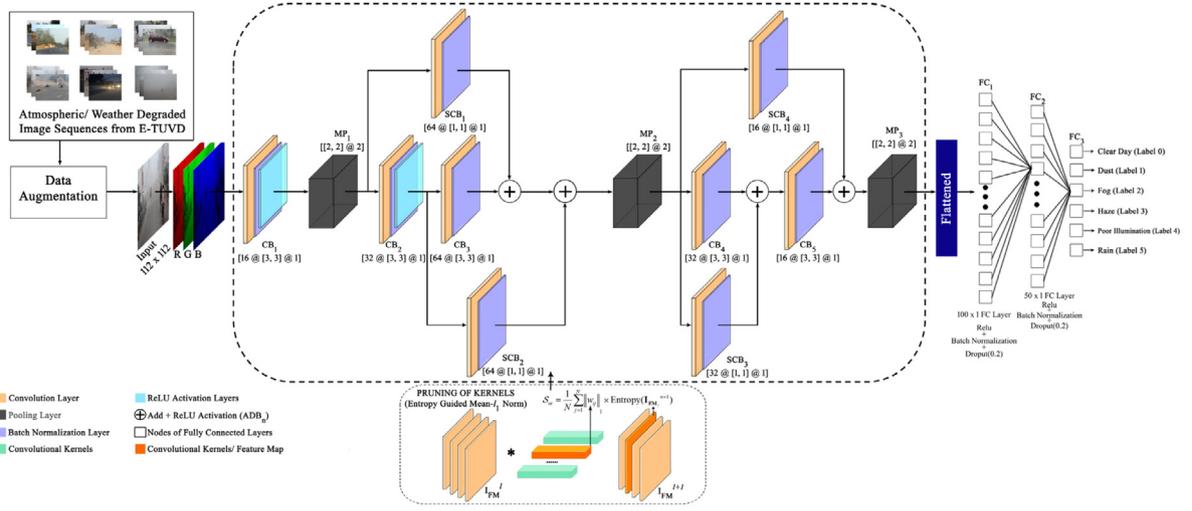


Fig. 3. Pipeline of our Proposed CNN Architecture (AWDMC-Net) for Classification of Atmospheric/Weather Conditions.

flattened and converted into a one-dimensional feature vector  $\mathbf{I}_{\text{FM}}^{n-1}$  of size  $m$  and linked into the  $L$  number of fully connected layers ( $\text{FC}_L$ ), where  $L = \{1, 2\}$ . Each  $L$  consists of 100 and 50 neurons respectively, which are densely connected, i.e., all the neurons of the previous layers are interconnected with one another. The output of the fully connected layer is shown in Eq. (4).

$$\text{FC}_L^n = \xi \left( B \left( \sum_{r=1}^m \mathbf{I}_{\text{FM}_r}^{n-1} \times C_{kr}^n + b_k^n \right) \right) \quad (4)$$

The output of each of these layers is batch normalized,  $B(\bullet)$ , and activated,  $\xi(\bullet)$ , as shown in Eqs. (2) and (3), respectively.

#### Dropout Layer.

This layer overcomes the overfitting problem of the model by disabling the activity of certain neurons for performance improvement (Srivastava et al., 2014). In our proposed architecture, we have used two dropout layers, as shown in Fig. 3. The dropout layers are used after the first and second fully connected layers, with a probability of 0.2.

#### Output Layer.

This layer is also a part of the fully connected layer (represented as  $\text{FC}_3$  in Fig. 3), which is densely connected to the preceding fully connected layers and is liable for class label prediction. In our proposed network, this layer consists of six neurons, as shown in Fig. 3, each representing a particular class of atmospheric/weather conditions (i.e., Clear Condition (Label 0), Dust Condition (Label 1), Foggy Condition (Label 2), Haze Condition (Label 3), Poor Illumination Condition (Label 4), and Rain Condition (Label 5)). As the objective of the proposed network is multiclass classification of atmospheric/weather degraded outdoor scenes, we have used the softmax activation function (Gao and Pavel, 2017) in this layer where the output real values from the last fully connected layer, are normalized to target class probabilities. Let  $\mathcal{H}(\bullet)$  be the mapping function of the entire process from input to output. The softmax activated output function  $\mathbb{S}(\bullet)$  then be represented as shown in Eq. (5).

$$P_L = \mathbb{S}(\mathbf{O}_L = c | \mathcal{H}) = e^{\mathcal{H}_L} / \sum_{c=1}^{N_c} e^{\mathcal{H}_c} \quad (5)$$

Here,  $\mathbf{O}_L$  is the actual class probability belonging to class  $c$ , which signifies real values ranging from 0 to 1, and  $c \in \{1, 2, \dots, N_c\}$ , where  $N_c = 6$  represents the number of classes. Depending on the highest probability score of the output layer (comprised of six neurons), the proposed AWDMC-Net classifies  $112 \times 112$  sized input images into either of the six different atmospheric/weather classes (i.e., Clear

Condition, Dust Condition, Foggy Condition, Haze Condition, Poor Illumination, and Rain Condition).

#### Skip Connections.

In deep networks, more complex, discriminating, and fine detailed features are generally learned as the network grows because of which the features learned by the network in the previous layers are not able to sustained and may result in a decrease in the prediction performance of the network. Therefore, utilizing the concept of skip connections in the network, the subsequent layers can learn the features of the preceding layers directly. According to Huang et al. (2017), there are two types of skip connections, i.e., residual connectivity (addition shortcut used in ResNet (Huang et al., 2017)), and dense connectivity (concatenation shortcut used in DenseNet (Huang et al., 2017)). Usually, the selection of kernels and their corresponding sizes controls the number of learnable parameters which further affects the network performance. In Prateek Joshi (2021), it was reported that there is a high risk for model overfitting if the number of learnable parameters is greater than or equal to the number of considered samples for training the model.

Depending upon this phenomenon, we have used addition based skip connections in our proposed network, as shown in Fig. 3. The main reason for using the addition based skip connections compared to the concatenation based skip connections is that the numbers of feature maps and parameters remain consistent throughout the network. Let us consider  $\text{ADB}_n$  as the output of the  $n$ th addition layer (as shown in Fig. 3). This means that the output feature map of the  $n$ th addition layer is a non-linear transformation  $\xi(\bullet)$  to the convoluted, followed by a batch normalized feature map of the output of the  $(n-i)$ th layer plus the output of the  $(n-1)$ th layer (i.e., a convoluted feature map of the preceding layer). The same is shown in Eq. (6).

$$\text{ADB}_n = \xi \left( B \left( \sum_{r=1}^m \mathbf{I}_{\text{FM}_r}^{n-i} * C_k^n + b_k^n \right) + \mathbf{I}_{\text{FM}}^{n-1} \right) \quad (6)$$

Here,  $i$  is the number of layers skipped from the addition layer. The resultant feature map  $\text{ADB}_n$  is then fed to the next layer. In our proposed network (described in Section 4.1), there are five sequentially stacked convolutional blocks ( $\text{CB}_n$ ), where each block considers a different number of convolutional kernels. This means that the preceding and subsequent feature maps are not of the same size. Therefore, we have used skip connections ( $\text{SCB}_N$ ) in our proposed network, each of which is passed by a  $1 \times 1$  convolutional block and added to the subsequent layer. Moreover using skip connection in the proposed network, the local feature maps more relevant to the local semantic information will be assigned a higher weight via this  $1 \times 1$  convolution layer. In particular, we have investigated whether increasing the

**Table 5**  
Effect of skip connections on the learning accuracy.

Skip connection	Input from the output of	Connected to the output of	Other associated skip connections	Number of skip connections	Training accuracy	Validation accuracy	EX/IN
No	–	–	–	–	92.97%	90.40%	–
SCB <sub>1</sub>	Input	CB <sub>1</sub>	–	1	92.08%	89.65%	EX
SCB <sub>2</sub>	MP1	CB2	–	1	92.71%	89.89%	EX
SCB <sub>3</sub>	MP1	CB3	–	1	95.35%	91.03%	IN
SCB <sub>4</sub>	CB2	CB3,ADB <sub>1</sub>	SCB3	2	95.76%	92.45%	IN
SCB <sub>5</sub>	MP2	CB4	SCB3, SCB4	3	95.84%	92.89%	IN
SCB <sub>6</sub>	MP2	CB5	SCB3, SCB4, SCB5	4	97.45%	94.45%	IN
SCB <sub>7</sub>	CB3	ADB <sub>4</sub>	SCB3, SCB4, SCB5, SCB6	5	94.34%	90.88%	EX

E—Excluded from the Network; IN—Included in the Network; SCB<sub>N</sub>—Nth Skip Connection Block; CB<sub>N</sub>—Nth Convolutional Block; MP<sub>N</sub>—Nth Max Pooling Block, ADB<sub>N</sub>—Nth Addition Layer.

number of skip connections (SCB<sub>N</sub>) in the network can improve the accuracy, and we have determined the number of skip connections that have the highest learning accuracy. The effect of the skip connections (SCB<sub>N</sub>) on the learning accuracy of our proposed network is shown in Table 5. The various associated hyper parameters of the network training are consistent, as described in Section 4.3. Table 5 show that a maximum accuracy of 94.45% is obtained with four skip connection blocks (i.e., SCB<sub>3</sub>, SCB<sub>4</sub>, SCB<sub>5</sub>, and SCB<sub>6</sub>). These findings indicate that the subsequent layer of our proposed network, i.e., AWDMC-Net, can learn the features of the preceding layers, and that the accuracy is also increased. Therefore, we have included these four skip connections in our proposed network, thereby excluding the others as shown in Fig. 3. These four skip connections are further epitomized as SCB<sub>1</sub>, SCB<sub>2</sub>, SCB<sub>3</sub>, and SCB<sub>4</sub>, respectively in Fig. 3 for better visualization. Each of the SCB<sub>N</sub> layers has C<sub>k</sub> kernels, where k ∈ {64, 64, 32, 16} and the size of each kernel (C<sub>k</sub>) is set to 1 × 1 respectively.

The description of each layer associated with our proposed AWDMC-Net for the classification of atmospheric/weather degradation conditions in outdoor scenes is summarized in Table 6.

#### 4.3. Training of our proposed AWDMC-Net architecture

In this subsection, we discussed the loss function used to optimize the proposed network, thereby tuning the learnable kernels. As our main task is a multiclass problem, we have incorporated categorical cross entropy (Rusiecki, 2019) as a loss function  $\mathcal{L}(\bullet)$  to quantify the loss of our proposed network, which indicates the error incurred in the learnable parameters and is expressed in Eq. (7).

$$\mathcal{L}(O_L, P_L) = - \sum_{c=1}^{N_c} O_L^c \times \log(P_L^c) \quad (7)$$

Here,  $O_L^c$  and  $P_L^c$  represent the original probability (i.e., ground truth label) and network predicted probability for the  $c$ th class, respectively, and  $N_c$  represents the total number of atmospheric/weather classes considered in our study. The main motive is to update the learnable parameters in such a way that the loss function  $\mathcal{L}(\bullet)$  is minimized. Therefore, the gradient for the loss function regarding the mapping function  $\frac{\partial(\mathcal{L}(O_L, P_L))}{\partial H_i}$  is used to update the learnable parameters in the next iteration. Here,  $i$  represents a correct class. The process will be iteratively repeated until the loss function is minimized to a certain small value. To continue the training process, the Adam optimizer (Kingma and Ba, 2014) was used in our proposed architecture. After successive iterations, it will provide probabilities or scores signifying that the image sequences belong to one of the six pre-defined classes, as shown in Fig. 3. The image sequences under consideration will be allocated with the corresponding class label for which the predicted probability score is a maximum, which is shown in Eq. (8).

$$\mathfrak{S}(\mathbf{I}) = \operatorname{argmax}(P_L^c) \quad (8)$$

Here, the class label function  $\mathfrak{S}(\bullet)$  predicts the label of each input image sequence  $\mathbf{I}$ , as per Eq. (8). Thus, after successive training the class label of each of the test image sequences is predicted by the network

**Table 6**  
Layer wise description of AWDMC-NET architecture for classification of atmospheric/weather conditions.

Layer type	Input	Kernel size	Filters/Stride	Connected to	No. of LPs in each layer
INPUT	112 × 112 × 3	–	–	–	0
CONV <sub>1</sub>	112 × 112 × 3	3 × 3	16/1	INPUT	448
BN <sub>1</sub>	112 × 112 × 16	–	–	CONV <sub>1</sub>	64
RELU <sub>1</sub>	112 × 112 × 16	–	–	BN <sub>1</sub>	0
POOL <sub>1</sub>	112 × 112 × 16	2 × 2	–/2	RELU <sub>1</sub>	0
CONV <sub>2</sub>	56 × 56 × 16	3 × 3	32/1	POOL <sub>1</sub>	4640
BN <sub>2</sub>	56 × 56 × 32	–	–	CONV <sub>2</sub>	128
RELU <sub>2</sub>	56 × 56 × 32	–	–	BN <sub>2</sub>	0
CONV <sub>3</sub>	56 × 56 × 32	3 × 3	64/1	RELU <sub>2</sub>	18 496
BN <sub>3</sub>	56 × 56 × 64	–	–	CONV <sub>3</sub>	256
CONV <sub>6</sub>	56 × 56 × 32	1 × 1	64/1	POOL <sub>1</sub>	1088
BN <sub>6</sub>	56 × 56 × 64	–	–	CONV <sub>6</sub>	256
ADB <sub>1</sub>	56 × 56 × 64	–	–	BN <sub>3</sub> , BN <sub>6</sub>	0
RELU <sub>3</sub>	56 × 56 × 64	–	–	ADB <sub>1</sub>	0
CONV <sub>7</sub>	56 × 56 × 64	1 × 1	64/1	RELU <sub>2</sub>	2112
BN <sub>7</sub>	56 × 56 × 64	–	–	CONV <sub>7</sub>	256
ADB <sub>2</sub>	56 × 56 × 64	–	–	ADB <sub>1</sub> , BN <sub>7</sub>	0
RELU <sub>4</sub>	56 × 56 × 64	–	–	ADB <sub>2</sub>	0
POOL <sub>2</sub>	56 × 56 × 64	2 × 2	–/2	RELU <sub>4</sub>	0
CONV <sub>4</sub>	28 × 28 × 64	3 × 3	32/1	POOL <sub>2</sub>	18 464
BN <sub>4</sub>	28 × 28 × 32	–	–	CONV <sub>4</sub>	128
CONV <sub>8</sub>	28 × 28 × 64	1 × 1	32/1	POOL <sub>2</sub>	2080
BN <sub>8</sub>	28 × 28 × 32	–	–	CONV <sub>8</sub>	128
ADB <sub>3</sub>	28 × 28 × 32	–	–	BN <sub>4</sub> , BN <sub>8</sub>	0
RELU <sub>5</sub>	28 × 28 × 32	–	–	ADB <sub>3</sub>	0
CONV <sub>5</sub>	28 × 28 × 32	3 × 3	16/1	BN <sub>4</sub>	4624
BN <sub>5</sub>	28 × 28 × 16	–	–	CONV <sub>5</sub>	64
CONV <sub>9</sub>	28 × 28 × 64	1 × 1	16/1	POOL <sub>2</sub>	1040
BN <sub>9</sub>	56 × 56 × 16	–	–	CONV <sub>9</sub>	64
ADB <sub>4</sub>	28 × 28 × 32	–	–	BN <sub>5</sub> , BN <sub>9</sub>	0
RELU <sub>6</sub>	28 × 28 × 16	–	–	ADB <sub>4</sub>	0
POOL <sub>3</sub>	28 × 28 × 16	2 × 2	–/2	RELU <sub>6</sub>	0
FLAT <sub>1</sub>	14 × 14 × 16	–	–	POOL <sub>3</sub>	0
FC <sub>1</sub>	3136 × 1	–	–	FLAT <sub>1</sub>	313 700
BN <sub>12</sub>	100 × 1	–	–	FC <sub>1</sub>	400
RELU <sub>12</sub>	100 × 1	–	–	–	0
DO <sub>1</sub> (20%)	–	–	–	–	0
FC <sub>2</sub>	100 × 1	–	–	DO <sub>1</sub>	5050
BN <sub>13</sub>	50 × 1	–	–	FC <sub>2</sub>	200
RELU <sub>13</sub>	50 × 1	–	–	BN <sub>13</sub>	0
DO <sub>2</sub> (20%)	–	–	–	–	0
FC <sub>3</sub>	50 × 1	–	–	DO <sub>2</sub>	306

Total number of learnable parameters: 3,73,992 Parameters

Conv<sub>n</sub>—n-th Convolution Layer; RELU<sub>n</sub>: n-th Activation Layer; Pool<sub>n</sub>: n-th Pooling Layers, FC<sub>n</sub>—n-th Fully Connected Layer; BN<sub>n</sub>—n-th Batch Normalization, ADB<sub>n</sub>—n-th Addition Layer; DO<sub>n</sub>—n-th Dropout Layer; LPs: Learnable Parameters.

passed to it, depending on the discriminative features it had learned during the training process.

#### Setting of the Training Parameters:

To train out the proposed network, 48 000 image sequences (as shown in Table 3) are arbitrarily selected from E-TUVD and further augmented, as described in Section 5.1, providing balanced coverage

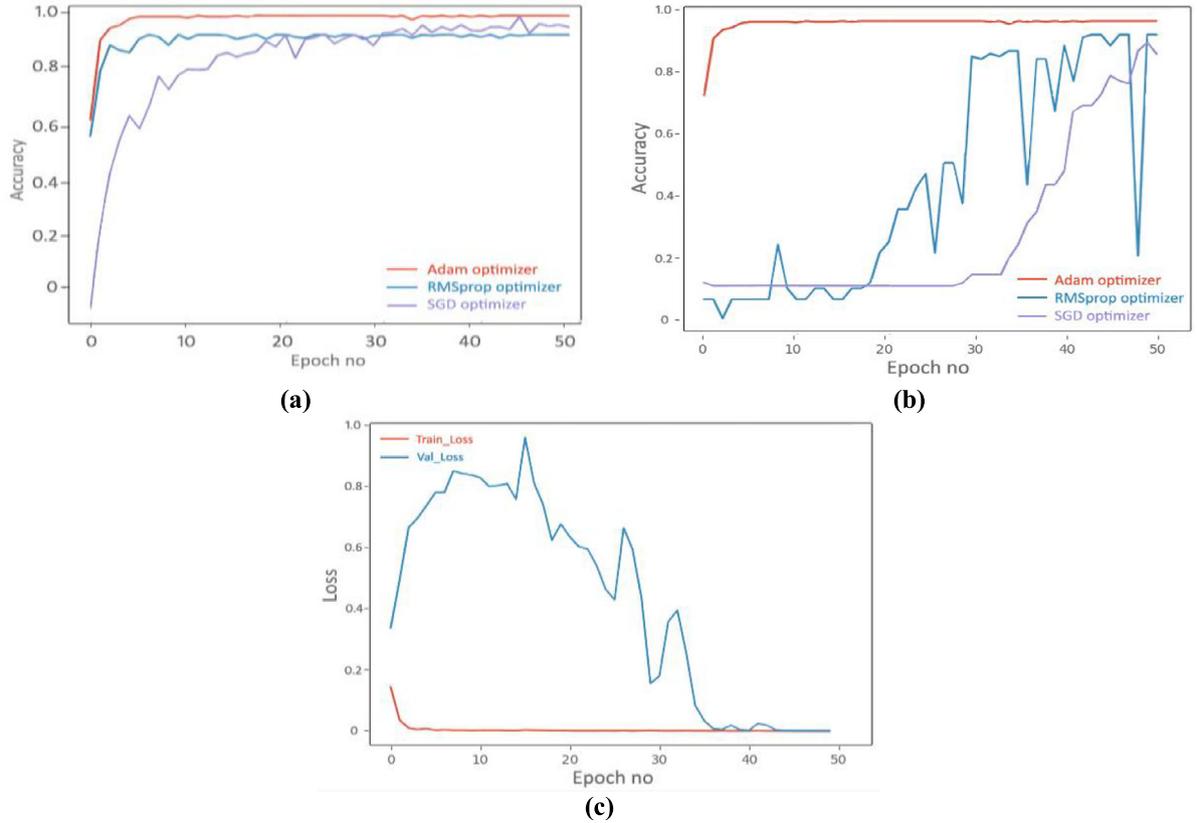


Fig. 4. (a) Training accuracy of our proposed AWDMC-Net with different optimizers; (b) Validation accuracy of our proposed AWDMC-Net with different optimizers; (c) Training and validation loss of our proposed AWDMC-Net using categorical cross entropy loss.

of representative atmospheric/weather degraded conditions of the real world (i.e., fog, haze, dust, rain, poor illumination, and clear days), along with the corresponding atmospheric/weather class labels. Training images are arbitrarily shuffled before feeding the proposed network for training. The training image sequences are split into a training set and validation set at a 4:1 ratio along with their corresponding class labels. The learning rate and weight decay were set to 0.001 and 0.0002 (experimentally), respectively. Moreover, random initialization of weights was performed. We empirically compared the training and validation accuracy of our proposed model with three widely used optimizers, i.e., Adam (Kingma and Ba, 2014), RMSprop (root mean square propagation) (DataDrivenInvestor, 2021) and SGD (stochastic gradient descent) (Sutskever et al., 2013) (as shown in Fig. 4). It has been observed from Fig. 4 that the maximum learning accuracy is obtained using the Adam optimizer (Kingma and Ba, 2014); therefore, we have used Adam (Kingma and Ba, 2014) as an optimizer in our proposed network. A total of 50 epochs with a batch size of 32 images are used in our proposed work, such that the softmax loss is minimized (as described in Section 4.3).

The implementation of our proposed architecture is conducted using Python platform (Tensor-flow and Keras library) on an Nvidia GeForce GTX TITAN XP GPU based system with 64 GB installed memory (RAM). Fig. 4 represents the training/ validation loss obtained per epoch of our proposed model using the categorical cross entropy loss function (Rusiecki, 2019).

#### 4.4. Pruning of convolutional kernels

After training our proposed network (i.e., AWDMC-Net) until it converges on the atmospheric/weather classification task, less important convolutional kernels (sometimes pronounced as filters) are pruned and further retrained to recover the prediction performance of the network. One of the main advantages of filter level pruning is that

no alteration of the network structure is performed. In the literature, there are several empirical criteria to evaluate the significance of each convolutional kernel in the network and prune them accordingly (Luo et al., 2018; Li et al., 2016; Hu et al., 2016). Each of these pruning criteria essentially only considers the filter weights for calculating the score of each filter, but does not take into account the influence of the corresponding output feature maps in the network. Therefore, the selection of relevant convolutional filters/kernels adaptively and competently is still a difficult task. In our work, we proposed a new pruning method to compute the score of each filter, called the “entropy guided mean- $l_1$  norm”. Generally, entropy is a frequently used metric in information theory to measure disorder or uncertainty. The higher the entropy value, the more information the system has. Depending upon this phenomenon, the feature maps with more discriminative information will have higher entropy values as compared to the feature maps with less discriminative information. Thus, on the basis of this observation, our proposed pruning criteria evaluates the filter importance globally by comprehensively considering filters and the entropy values of their corresponding output feature maps. Let  $w_{ij}$  be the weight of the  $i$ th filter at the  $n$ th layer and  $\mathbf{I}_{\mathbf{F}_M}^{n+1}$  be the corresponding input feature map to the  $(n+1)$ th layer obtained from  $i$ th filter. Then the entropy guided mean- $l_1$  norm based pruning criterion (our proposed pruning method) are expressed in Eq. (9).

$$S_{ni} = \frac{1}{N} \sum_{j=1}^N \|w_{ij}\|_1 \times \text{Entropy}(\mathbf{I}_{\mathbf{F}_M}^{n+1}) \quad (9)$$

Here,  $N$  represents the number of input channels at the  $n$ th layer. Depending on this scoring criterion, the proposed network adaptively eliminates all the irrelevant or less important convolution kernels according to their individual scores, as computed in Eq. (9) (i.e., filters with the smallest values of  $S_{ni}$  are discarded from the network). In

**Table 7**

Overall performance of our pruning criteria to reduce parameters and increase prediction performance on our proposed AWDMC-Net atmospheric/weather classification network.

Pruning method	Parameters			Prediction performance		
	Original (L)	Pruned (L)	Percentage	Accuracy (%)	Specificity (%)	Sensitivity (%)
Mean- <sub>1</sub> (Li et al., 2016)		3.66	2%	93.01 ± 1.21	93.03 ± 1.59	93.92 ± 1.93
Mean- <sub>2</sub> (Li et al., 2016)	3.74	3.68	2%	92.36 ± 1.34	91.21 ± 2.11	93.42 ± 1.45
APoZ (Hu et al., 2016)		3.61	4%	92.96 ± 1.67	92.21 ± 1.45	93.18 ± 2.13
<b>Proposed pruning method</b>		<b>3.55</b>	<b>5%</b>	<b>93.85 ± 1.65</b>	<b>93.79 ± 2.33</b>	<b>94.18 ± 1.89</b>

Bold Face—Most Outer Performed Classified Conditions; L—Lakhs.

our filter pruning scenario, if the current filter is less important for output feature map generation,  $S_{ni}$  will approach towards 0. The prediction results are almost unaffected by removing this filter. A higher  $S_{ni}$  value, on the other hand, indicates that the present dimension is more essential. During pruning processes, the number of convolution kernels discarded/eliminated is different for each convolution layer. However, due to the presence of the shortcut connections in the proposed network, the pruning of convolutional filters/kernels from each layer in the network is quite complicated. This is because the channel numbers of skip connection blocks ( $SCB_n$ ) and the corresponding output of convolution blocks ( $CB_n$ ) after pruning in a stage (i.e., addition layer) need to be consistent to make the sum operations valid. Due to the structural constraints, after eliminating  $\mathcal{P}$  percentage of filters from the proposed network of those having lowest scores, the number of channels/filters in the two subsequent layers to be added in the addition layer are matched for the dimension. If the dimension does not match (i.e., the two subsequent layers to be added do not have the same number of feature maps), then the dimensions of the two layers are matched by eliminating the number of filters that have the lowest scores, as per Eq. (9), from the layer that has more filters than the other layer to be added in the corresponding addition layer. The advantages of our pruning process are self-evident. First, it is a global criteria that can assess the influence of all filters at the same time. Second, there is no direct relationship between importance score and filter placement, i.e., the number of filters that should be deleted is totally dependent on the scores.

Finally, all filter scores will be sorted in ascending order across all levels. The top ‘ $T$ ’ percentage of filters will be eliminated, resulting in a compact model that has been trimmed. In practice, the value of ‘ $T$ ’ is determined by the computational or storage budget available. For the proposed network, the influence of the number of pruned convolution kernels (i.e., in percentage) based on our proposed pruning criteria has been studied on the image sequences of our created E-TUVD dataset, as described in Table 3. The impact of the percentage of pruned convolution kernels on the proposed network is displayed in Fig. 5. Here, the  $X$ -axis indicates the percentage of filters pruned from the proposed network, and the  $Y$ -axis indicates the accuracy. The training accuracy and validation accuracy were considered for a total of 50 epochs, and the various associated hyper parameters of the network training remained the same as those described in Section 4.3. Fig. 5 shows that the percentage of filters pruned from the proposed network has a strong influence on the accuracy. Careful inspection of Fig. 5 can lead to the following observation: as the pruning percentage increases, the relevant filters are eliminated from the network, thereby decreasing the prediction performance of the network. Therefore, only the suitable selection of a pruning percentage can apprehend the balance between the minimum computational resource and maximum prediction performance. Additionally, from Fig. 5 it can be observed that the proposed network achieves a maximum performance with a 5% pruning percentage of convolution kernels/filters.

Moreover, in order to establish the effectiveness of our proposed pruning criterion (i.e., entropy guided mean- $l_1$  norm based pruning criterion as shown in Eq. (9)), a comparison with respect to the well-known state-of-the-art pruning criteria i.e., mean- $l_1$  (Li et al., 2016), mean- $l_2$  (Li et al., 2016), and APoZ (Average Percentage of Zeros) (Hu et al., 2016), has been studied. Except for APoZ, the remaining two

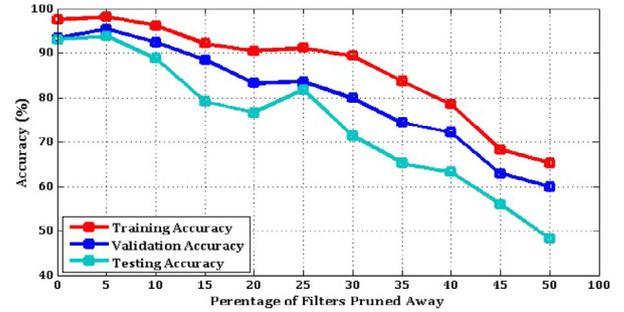


Fig. 5. Effect of pruning of convolutional kernels in learning accuracy using our proposed pruning criterion (Entropy Guided Mean- $l_1$  Norm).

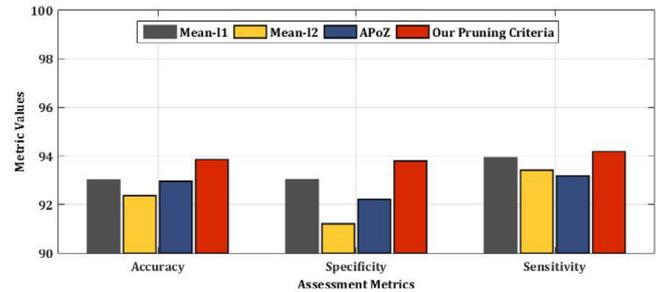


Fig. 6. Effect of using the Proposed Pruning Criterion (i.e., Entropy Guided Mean- $l_1$  Norm) as compared to Mean- $l_1$  (Li et al., 2016); Mean- $l_2$  (Li et al., 2016); and APoZ based pruning criteria (Hu et al., 2016).

state-of-the-art pruning criteria consider filters with higher scores to be more relevant. In Table 7, comparison of our pruning criteria with respect to the network parameters and prediction performance has been reported. The prediction performance was measured in terms of the accuracy, specificity, and sensitivity. These measurement units were obtained from the confusion matrix, which includes: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). From Fig. 6 representing the bar plot of the prediction performance (i.e., testing performance on the testing set of the E-TUVD dataset as provided in Table 3) of the proposed pruning criterion “entropy guided mean- $l_1$  norm” compared to the three state-of-the-art pruning criteria, it can be observed that the proposed pruning criteria has a superior prediction performance compared to the state-of-the-art pruning criteria with an average accuracy, specificity, and sensitivity of 93.85%, 93.79%, and 94.18% respectively. Even though the prediction performance of our proposed pruning method have achieved 0.84% improvement in average accuracy as compared to mean- $l_1$  pruning method (Li et al., 2016) but the network learnable parameters of our proposed AWDMC-Net model thereafter pruning using our proposed pruning method (so as compared to the state-of-the-art pruning methods) is also less (i.e., 3.55 lakhs learnable parameters) thereby reducing the expensive memory utilization and computation costs.

**Table 8**  
Parameter values for data augmentation.

Augmentation parameters	Parameter values
Rotation range	45°, 90°, 135°, 180°, 225°, 270°, and 315°
Horizontal flipping	True
Vertical flipping	True
Rescale	1/255
Zoom range	0.3
Shift range of width	0.3
Shift range of height	0.3

## 5. Experimental results and discussions

In this section, the performance of our proposed network (AWDMC-Net) for the classification of atmospheric/weather degraded image sequences is discussed. Specific illustrations of the experimental results are as follows.

### 5.1. Data augmentation

The main motive of applying data augmentation before feeding to the CNNs is to increase the size of the dataset (i.e., in volume and variety) and to generalize the CNN models with respect to the dataset while preventing the classification accuracy in terms of overfitting. The increase in the training and validation set of the dataset has been performed through different techniques of image processing in order to create new image sequences. These techniques are: flipping, rotation, rescaling, zooming, and shifting. The various parameter values associated with the augmentation techniques adopted in our study are tabulated in Table 8. With these augmentation techniques, we increased the volume of the training dataset (i.e., containing training and validation sets) from 48 000 to 624 000 atmospheric/weather degraded image sequences.

### 5.2. Influence of pruned kernels on the classification accuracy

In this subsection, we report on and discuss the performance of the proposed network before pruning (i.e., AWDMC-Net<sub>WOKP</sub>) and after pruning (i.e., AWDMC-Net<sub>WKP</sub>) of convolutional kernels/filters using our proposed pruning criterion as described in Section 4.4. Table 9 shows the prediction performance of the proposed network using the average accuracy, specificity, and sensitivity. The standard descriptor (i.e., mean  $\pm$  standard deviation) for each considered assessment metric was used to understand the prediction performance of our proposed model. To assess the performance of our proposed model, we tested a testing set of image sequences from E-TUVD (i.e., 12 000 image sequences as described in Table 3) that are not used for training our network. As in Table 9, it can be observed that the proposed network after pruning the convolutional kernels (i.e., 5% of filters are pruned based on our proposed pruning method as described in Section 4.4) achieves a classification performance with an average accuracy, specificity, and sensitivity of 93.85%, 93.79%, and 94.18%, respectively. Moreover, pruning the convolutional kernels from the proposed network significantly increases the atmospheric/weather classification performance with an average percentage-point improvement accuracy of 0.80% compared to the proposed network without pruning convolution kernels/filters. Therefore, for a comparison of the proposed network with the state-of-the-art methods, we have reported the experimental results of the proposed network with adaptive pruning, which has been pronounced as AWDMC-Net in later subsections.

### 5.3. Qualitative evaluation of the proposed AWDMC-Net model in atmospheric/weather conditions of E-TUVD

In this subsection, the performance of our proposed AWDMC-Net model for the classification of each individual atmospheric/weather

condition on our E-TUVD dataset is reported. The prediction performance is tabulated in Table 10 in terms of three assessment metrics: accuracy, specificity, and sensitivity. Similar to the previous subsection, the perception capability of our proposed AWDMC-Net model has been evaluated on the testing set of E-TUVD, containing 2000 image sequences in each class of atmospheric/weather conditions (as shown in Table 3). As displayed in Table 10, the accuracies of our proposed AWDMC-Net model for the prediction of outdoor scenes degraded by fog, haze, dust, rain, poor illumination, and clear day conditions were 96.01%, 93.70%, 91.62%, 92.35%, 94.36%, and 95.08%, respectively. Additionally, it can be observed that the AWDMC-Net model has a superior prediction performance for fog conditions compared to the other five atmospheric/weather conditions, with an average specificity and sensitivity of 96.25% and 95.81%, respectively.

### 5.4. Comparison of the proposed AWDMC-Net model with state-of-the-art pre-trained CNN models

In this subsection, we compared our proposed model (AWDMC-Net) with the pre-trained CNN models applied to the ImageNet Challenge Dataset (Deng et al., 2009) for the classification of outdoor scenes of the E-TUVD dataset degraded by atmospheric/weather conditions. Currently, many pre-trained models of CNN applied on the ImageNet Challenge Dataset (Deng et al., 2009) are available for the research community, along with their learned kernels and weights. There are three procedures to use these pre-trained networks: fixed feature extraction, transfer learning and fine-tuning. From the literature, fifteen well-known and widely used pre-trained CNN models are used in our study: VGG-16 (Russakovsky et al., 2015), VGG-19 (Russakovsky et al., 2015), AlexNet (Krizhevsky et al., 2012), Inception-V3 (Szegedy et al., 2016), GoogleNet (Szegedy et al., 2015), Resnet-101 (He et al., 2016), Resnet-50 (Huang et al., 2017), Resnet-18 (Huang et al., 2017), Densenet-201 (Huang et al., 2017), Xception (Chollet, 2017), MobileNet-V2 (Sandler et al., 2018), Inception-Resnet-V2 (Szegedy et al., 2017), EfficientNet-B0 (Tan and Le, 2019), Darknet-19 (Redmon, 2020), and Darknet-53 (Redmon, 2020). To evaluate and compare the prediction performance of our proposed model (AWDMC-Net) with these state-of-the-art pre-trained models, we used these aforementioned pre-trained CNNs as a fixed feature extraction module and fine tuning module.

**Pre-Trained CNNs as a Fixed Feature Extraction Module.** In our present work, a fixed feature extraction method is adopted by eliminating the fully connected layers from the CNNs that are pre-trained on the ImageNet dataset (Deng et al., 2009) while maintaining the remaining network. After feature extraction using these pre-trained networks, a support vector machine (SVM) classifier (Cristianini and Shawe-Taylor, 2000) with four different kernels (i.e., Linear, Radial Basis Function (RBF), Polynomial, and Gaussian) and k-fold cross validation were appended on the fixed feature extractor, resulting in the classification of atmospheric/weather degraded scenes on E-TUVD dataset. To evaluate the performance of CNNs as a feature extractor, 12 000 similar frames of E-TUVD used for testing our proposed CNN models were used. The performance of the pre-trained CNNs and a comparison of our proposed model (i.e., AWDMC-Net) against the four kernels of the SVM are illustrated in Table 11. The prediction performance has been reported in terms of its accuracy, sensitivity, and specificity. Similar to Section 5.2, the standard descriptor (i.e., mean  $\pm$  standard deviation) for each assessment metric was used to measure the prediction performance. The two best performing CNN models for the classification of atmospheric/weather degraded frames against each of the classifiers are represented by bold underlined faces and bold faces in Table 11, respectively. From Table 11, the following observations can be made:

1. Different CNN based feature sets have shown a superior prediction performance for different kernels of SVM. Among all four kernels of the SVM classifiers, SVM\_Polynomial showed

**Table 9**

Recognition accuracy of our proposed AWDMC-Net model with and without adaptive pruning.

AWDMC-Net with different combinations	Recognition accuracy (Mean $\pm$ SD)		
	Accuracy (%) ( $\uparrow$ )	Specificity (%) ( $\uparrow$ )	Sensitivity (%) ( $\uparrow$ )
AWDMC-Net <sub>WKP</sub>	<b>93.85 <math>\pm</math> 1.65</b>	<b>93.79 <math>\pm</math> 2.33</b>	<b>94.18 <math>\pm</math> 1.89</b>
AWDMC-Net <sub>WOKP</sub>	93.05 $\pm$ 2.21	91.87 $\pm$ 2.32	94.12 $\pm$ 1.05

AWDMC-Net<sub>WKP</sub>—AWDMC-Net with Kernel Pruning; AWDMC-Net<sub>WOKP</sub>—AWDMC-Net without Kernel Pruning; SD—Standard Deviation;  $\uparrow$ —Higher Value Indicates Better Performance.

**Table 10**

Recognition accuracy of our proposed AWDMC-Net model in different atmospheric/weather degraded conditions of E-TUVD.

Atmospheric/Weather conditions	Recognition accuracy		
	Accuracy ( $\uparrow$ )	Specificity ( $\uparrow$ )	Sensitivity ( $\uparrow$ )
Fog Condition	<b>96.01%</b>	<b>96.25%</b>	<b>95.81%</b>
Haze Condition	93.70%	95.03%	94.87%
Dust Condition	91.62%	91.14%	90.94%
Rain Condition	92.35%	93.05%	92.92%
Poor Illumination Condition	94.36%	91.20%	94.83%
Clear Day Condition	<b>95.08%</b>	<b>96.12%</b>	<b>95.66%</b>

Bold Face and Underlined—Most Outer Performed Classified Conditions; Bold Face—Second Most Outer Performed Classified Conditions.

a better prediction performance for all the extracted feature sets using pre-trained CNNs and our proposed model with an average accuracy, specificity, and sensitivity of 79.62%, 80.79, and 79.61%, respectively. Conversely, SVM\_Gaussian showed a lower prediction performance for all the extracted feature sets using CNNs with an average accuracy, specificity, and sensitivity of 74.06%, 73.54, and 75.29%, respectively.

2. Comparing the prediction performance of all the CNN models as a fixed feature extraction module (as shown in Table 11), our proposed CNN model showed a superior prediction performance for all the kernels of the SVM classifier. For SVM\_Linear, the average accuracy, specificity and sensitivity were 89.56%, 89.64%, and 92.70%, respectively; for SVM\_Polynomial, the average accuracy, specificity and sensitivity were 91.40%, 92.33%, and 94.76%, respectively; for SVM\_RBF, the average accuracy, specificity and sensitivity were 90.44%, 87.60%, and 94.02%, respectively; and for SVM\_Gaussian, the average accuracy, specificity and sensitivity were 88.45%, 86.79%, and 91.02%, respectively.

**Pre-Trained CNNs as a Fine-Tuning Module.** On the other hand, CNN with a fine-tuning module is an effective strategy to train the network using given data of interest. In our comparative analysis, we have also fine-tuned the aforementioned pre-trained CNN networks where the last fully connected layer in each of the considered pre-trained networks is replaced by six neurons. Each of the neurons for particular pre-trained CNNs in the last fully connected layer epitomizes six classes of atmospheric/weather conditions of the E-TUVD dataset. Here, we first trained the base networks on a base dataset (i.e., ImageNet challenge dataset Deng et al., 2009) and then the generic features learned from this large dataset (i.e., pre-defined weights) are transferred to the target network (i.e., maintaining the entire network) to be trained on our E-TUVD dataset. For a fair comparison, fine-tuned CNNs are trained on similar augmented training and validation sets of E-TUVD (as described in Section 5.1) that are used for training our proposed network. Meanwhile, to evaluate and compare the prediction performance of pre-trained CNNs as a fine-tuning module and our proposed network, we used a similar test set of 12000 image sequences from the E-TUVD dataset used for testing our proposed model. The training and validation accuracy for each fine-tuned pre-trained CNN model and our proposed model are shown in Table 12. In addition, the testing performance of the respective models in terms of their accuracy, specificity, and sensitivity are shown. Similar to the previous subsection, the two best performing CNN models are represented by bold underlined and bold faces. Table 12 leads to the following observations:

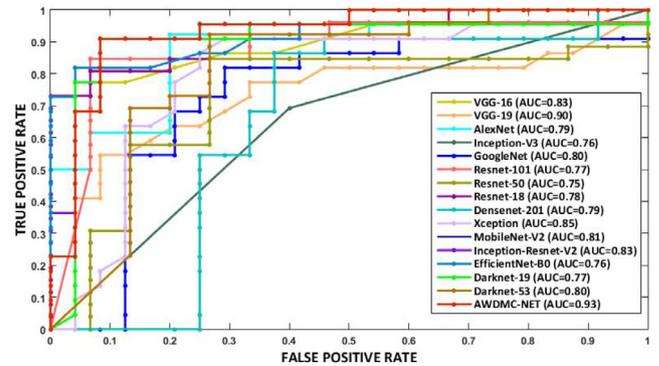


Fig. 7. ROC curves for comparison of fine-tuned pre-trained CNN Models and our proposed CNN model (AWDMC-Net) for atmospheric/weather classification task on E-TUVD dataset.

1. In comparing the prediction performance of each fine-tuned pre-trained CNN model, VGG-19 (Russakovsky et al., 2015) showed the best prediction performance with an average accuracy, specificity, and sensitivity of 90.16%, 93.18%, and 90.13%, respectively. Moreover, in comparison with the testing performance of all the fine-tuned CNN models (i.e., for the best performing epoch), our proposed AWDMC-Net model gives the highest classification performance with an average accuracy, specificity, and sensitivity of 93.85%, 93.79%, and 94.18%, respectively.
2. In addition to the classification performance, the analysis of ROC (Receiver Operating Characteristic) curves for all the pre-trained CNN models and our proposed CNN model for classifying atmospheric/weather conditions of the E-TUVD dataset is also plotted in Fig. 7. Here, AUC (Area Under the Curve) is also considered as the measure of the classification performance for the considered CNN models. As illustrated in Fig. 7, among all the CNN models, AWDMC-Net (i.e., our proposed CNN model) has the highest AUC value of 0.93, which indicates the superiority of its prediction performance over the pre-trained CNN models in classifying the six categories of atmospheric/weather conditions in the E-TUVD dataset (i.e., fog, haze, dust, rain, poor illumination, and clear day conditions).

##### 5.5. Comparison of the proposed AWDMC-Net model with the state-of-the-art methods on available benchmark datasets and E-TUVD dataset

To verify the robustness of the proposed model, in this subsection we compared its prediction performance with the state-of-the-art methods reported in the literature for atmospheric/weather classification tasks. For fair comparison, the prediction performance of the proposed AWDMC-Net has been evaluated on the available benchmark datasets used by most researchers for similar tasks. The benchmark datasets used in our present work are the RFS dataset (Guerra et al., 2018), MWI dataset (Zhang and Ma, 2015), REalistic Single Image DEhazing (RESIDE) dataset (Li et al., 2018), MWD dataset (Guerra et al., 2018), Image2Weather dataset (Chu et al., 2016), and WeatherDataset-4 (Zhao et al., 2019). The detailed key characteristics associated with each of these benchmark datasets are provided in Section 2.2. The

**Table 11**  
Accuracy, Specificity and Sensitivity of pre-trained CNNs and our AWDMC-Net model as a feature extraction on E-TUVD.

CNN architecture(s)		SVM-Linear (Mean $\pm$ SD)			SVM-Polynomial (Mean $\pm$ SD)		
		Accuracy (%)	Specificity (%)	Sensitivity (%)	Accuracy (%)	Specificity (%)	Sensitivity (%)
VGG-16	Russakovsky et al. (2015)	84.78 $\pm$ 3.12	<b>88.41 <math>\pm</math> 2.03</b>	81.84 $\pm$ 3.04	84.93 $\pm$ 2.86	85.33 $\pm$ 2.10	82.02 $\pm$ 1.11
VGG-19	Russakovsky et al. (2015)	<b>85.70 <math>\pm</math> 2.50</b>	81.34 $\pm$ 1.89	<b>86.46 <math>\pm</math> 1.65</b>	<b>88.42 <math>\pm</math> 2.01</b>	<b>89.45 <math>\pm</math> 1.82</b>	<b>88.05 <math>\pm</math> 1.09</b>
AlexNet	Krizhevsky et al. (2012)	76.86 $\pm$ 1.54	80.34 $\pm$ 0.96	75.00 $\pm$ 3.47	78.01 $\pm$ 1.19	83.34 $\pm$ 1.06	77.20 $\pm$ 0.98
Inception-V3	Szegedy et al. (2016)	65.54 $\pm$ 1.05	65.92 $\pm$ 3.01	64.81 $\pm$ 2.13	71.30 $\pm$ 3.20	74.97 $\pm$ 2.53	70.83 $\pm$ 2.40
GoogleNet	Szegedy et al. (2015)	76.85 $\pm$ 2.33	73.57 $\pm$ 1.84	74.33 $\pm$ 1.42	76.88 $\pm$ 2.02	79.68 $\pm$ 2.14	74.60 $\pm$ 1.56
Resnet-101	He et al. (2016)	74.16 $\pm$ 2.84	71.82 $\pm$ 3.11	78.32 $\pm$ 2.53	76.74 $\pm$ 1.18	74.07 $\pm$ 1.25	78.72 $\pm$ 1.05
Resnet-50	Nair and Hinton (2010)	70.82 $\pm$ 1.33	71.15 $\pm$ 1.87	68.97 $\pm$ 1.61	73.67 $\pm$ 1.32	73.40 $\pm$ 1.10	72.98 $\pm$ 1.21
Resnet-18	Nair and Hinton (2010)	71.54 $\pm$ 0.96	73.18 $\pm$ 1.02	70.83 $\pm$ 1.31	73.69 $\pm$ 1.05	75.82 $\pm$ 0.87	72.99 $\pm$ 1.19
Densenet-201	Nair and Hinton (2010)	76.17 $\pm$ 2.13	74.21 $\pm$ 2.01	76.52 $\pm$ 1.55	76.88 $\pm$ 1.34	76.03 $\pm$ 1.14	78.33 $\pm$ 1.32
Xception	Chollet (2017)	70.01 $\pm$ 1.20	67.34 $\pm$ 1.54	71.95 $\pm$ 1.12	73.31 $\pm$ 1.21	71.54 $\pm$ 1.06	75.72 $\pm$ 1.14
MobileNet-V2	Sandler et al. (2018)	81.71 $\pm$ 1.42	80.67 $\pm$ 1.32	84.86 $\pm$ 1.10	84.24 $\pm$ 2.23	85.91 $\pm$ 1.67	84.96 $\pm$ 1.89
Inception-Resnet-V2	Szegedy et al. (2017)	69.56 $\pm$ 2.06	69.56 $\pm$ 2.21	71.95 $\pm$ 1.45	72.05 $\pm$ 2.08	71.00 $\pm$ 2.10	72.95 $\pm$ 2.18
EfficientNet-B0	Tan and Le (2019)	83.30 $\pm$ 0.98	86.72 $\pm$ 1.05	80.07 $\pm$ 1.11	84.54 $\pm$ 1.11	88.93 $\pm$ 1.41	84.00 $\pm$ 1.23
Darknet-19	Redmon (2020)	82.05 $\pm$ 1.16	85.69 $\pm$ 1.35	79.53 $\pm$ 0.83	83.31 $\pm$ 3.01	86.87 $\pm$ 2.56	80.56 $\pm$ 2.01
Darknet-53	Redmon (2020)	81.33 $\pm$ 0.66	83.80 $\pm$ 0.74	79.50 $\pm$ 1.37	84.50 $\pm$ 1.44	84.00 $\pm$ 2.01	85.05 $\pm$ 1.32
<b>AWDMC-Net [Our]</b>		<b>89.56 <math>\pm</math> 1.36</b>	<b>89.64 <math>\pm</math> 2.05</b>	<b>92.70 <math>\pm</math> 2.21</b>	<b>91.40 <math>\pm</math> 0.86</b>	<b>92.33 <math>\pm</math> 1.08</b>	<b>94.76 <math>\pm</math> 1.34</b>

CNN architecture(s)		SVM-RBF (Mean $\pm$ SD)			SVM-Gaussian (Mean $\pm$ SD)		
		Accuracy (%)	Specificity (%)	Sensitivity (%)	Accuracy (%)	Specificity (%)	Sensitivity (%)
VGG-16	Russakovsky et al. (2015)	82.90 $\pm$ 1.10	84.33 $\pm$ 0.87	81.01 $\pm$ 1.06	78.84 $\pm$ 1.30	79.27 $\pm$ 1.22	78.85 $\pm$ 1.54
VGG-19	Russakovsky et al. (2015)	<b>86.44 <math>\pm</math> 0.94</b>	87.21 $\pm$ 1.06	<b>86.20 <math>\pm</math> 1.12</b>	<b>82.95 <math>\pm</math> 1.64</b>	<b>81.28 <math>\pm</math> 2.18</b>	83.36 $\pm$ 1.19
AlexNet	Krizhevsky et al. (2012)	76.90 $\pm$ 1.44	78.46 $\pm$ 1.54	73.52 $\pm$ 1.19	74.21 $\pm$ 2.12	71.27 $\pm$ 1.32	75.77 $\pm$ 1.13
Inception-V3	Szegedy et al. (2016)	68.03 $\pm$ 1.23	66.63 $\pm$ 1.63	69.98 $\pm$ 1.30	68.01 $\pm$ 0.89	70.03 $\pm$ 1.20	67.38 $\pm$ 1.08
GoogleNet	Szegedy et al. (2015)	75.43 $\pm$ 2.16	75.92 $\pm$ 1.29	73.81 $\pm$ 1.18	71.67 $\pm$ 1.11	70.42 $\pm$ 1.12	73.48 $\pm$ 1.33
Resnet-101	He et al. (2016)	73.12 $\pm$ 0.39	70.58 $\pm$ 1.17	75.15 $\pm$ 1.02	70.44 $\pm$ 1.34	69.38 $\pm$ 2.04	73.35 $\pm$ 1.19
Resnet-50	Nair and Hinton (2010)	72.40 $\pm$ 0.96	71.66 $\pm$ 1.88	72.30 $\pm$ 2.22	70.07 $\pm$ 2.56	69.40 $\pm$ 1.45	72.85 $\pm$ 1.16
Resnet-18	Nair and Hinton (2010)	71.06 $\pm$ 0.82	74.00 $\pm$ 1.34	69.85 $\pm$ 2.28	69.97 $\pm$ 1.38	71.71 $\pm$ 1.86	70.00 $\pm$ 2.01
Densenet-201	Nair and Hinton (2010)	74.83 $\pm$ 1.33	72.71 $\pm$ 2.31	76.49 $\pm$ 1.05	70.02 $\pm$ 0.93	73.65 $\pm$ 1.35	68.36 $\pm$ 1.22
Xception	Chollet (2017)	71.22 $\pm$ 1.10	67.57 $\pm$ 1.54	73.98 $\pm$ 1.26	69.78 $\pm$ 2.27	69.53 $\pm$ 2.45	71.74 $\pm$ 2.50
MobileNet-V2	Sandler et al. (2018)	82.33 $\pm$ 1.24	83.59 $\pm$ 0.82	80.45 $\pm$ 1.11	73.33 $\pm$ 2.05	74.84 $\pm$ 1.11	70.96 $\pm$ 2.37
Inception-Resnet-V2	Szegedy et al. (2017)	70.08 $\pm$ 2.10	68.09 $\pm$ 1.32	72.18 $\pm$ 2.21	68.59 $\pm$ 1.51	67.86 $\pm$ 1.06	71.20 $\pm$ 1.43
EfficientNet-B0	Tan and Le (2019)	83.52 $\pm$ 2.15	84.73 $\pm$ 1.65	81.26 $\pm$ 1.19	75.32 $\pm$ 1.39	77.92 $\pm$ 1.28	71.70 $\pm$ 1.07
Darknet-19	Redmon (2020)	83.05 $\pm$ 1.32	<b>87.94 <math>\pm</math> 1.03</b>	80.00 $\pm$ 0.91	78.02 $\pm$ 1.22	75.01 $\pm$ 1.04	79.95 $\pm$ 1.14
Darknet-53	Redmon (2020)	81.79 $\pm$ 1.21	84.06 $\pm$ 1.10	79.95 $\pm$ 1.23	75.31 $\pm$ 0.89	68.26 $\pm$ 2.34	<b>84.65 <math>\pm</math> 1.05</b>
<b>AWDMC-Net [Our]</b>		<b>90.44 <math>\pm</math> 1.10</b>	<b>87.60 <math>\pm</math> 0.97</b>	<b>94.02 <math>\pm</math> 1.35</b>	<b>88.45 <math>\pm</math> 1.03</b>	<b>86.79 <math>\pm</math> 0.96</b>	<b>91.02 <math>\pm</math> 1.13</b>

Bold Face and Underlined—Most Outer Performed among Pre-Trained CNNs (state-of-the-art CNNs); Bold Face—Second Most Outer Performed Method among Pre-Trained CNNs (state-of-the-art models); SD—Standard Deviation.

**Table 12**  
Performance comparison of our proposed AWDMC-Net and state-of-the-art pre-trained CNN models as a fine tuning module on E-TUVD.

CNN architecture(s)		Performance of pre-trained CNN as fine tuning module and our AWDMC-Net				
		Training accuracy	Validation accuracy	Testing performance (Mean $\pm$ SD)		
				Accuracy (%)	Specificity (%)	Sensitivity (%)
VGG-16	Russakovsky et al. (2015)	95.45%	89.54%	85.85 $\pm$ 1.64	89.16 $\pm$ 1.10	87.22 $\pm$ 1.32
VGG-19	Russakovsky et al. (2015)	<b>95.83%</b>	<b>90.23%</b>	<b>90.16 <math>\pm</math> 1.35</b>	<b>93.18 <math>\pm</math> 0.98</b>	<b>90.13 <math>\pm</math> 1.16</b>
AlexNet	Krizhevsky et al. (2012)	93.64%	86.94%	81.47 $\pm$ 1.10	78.45 $\pm$ 1.35	79.93 $\pm$ 1.29
Inception-V3	Szegedy et al. (2016)	93.27%	78.96%	78.30 $\pm$ 1.28	76.73 $\pm$ 1.22	81.65 $\pm$ 1.31
GoogleNet	Szegedy et al. (2015)	94.67%	81.55%	80.76 $\pm$ 0.97	79.45 $\pm$ 2.34	83.26 $\pm$ 1.34
Resnet-101	He et al. (2016)	95.96%	78.54%	77.57 $\pm$ 0.75	80.21 $\pm$ 1.16	77.38 $\pm$ 0.98
Resnet-50	Huang et al. (2017)	94.80%	78.56%	76.82 $\pm$ 1.12	76.44 $\pm$ 1.05	75.88 $\pm$ 1.32
Resnet-18	Huang et al. (2017)	94.96%	78.04%	77.51 $\pm$ 1.23	79.12 $\pm$ 1.31	76.74 $\pm$ 1.06
Densenet-201	Huang et al. (2017)	95.64%	81.34%	80.97 $\pm$ 1.18	82.58 $\pm$ 1.70	79.16 $\pm$ 1.20
Xception	Chollet (2017)	94.77%	88.91%	87.85 $\pm$ 1.26	89.16 $\pm$ 0.97	87.22 $\pm$ 0.92
MobileNet-V2	Sandler et al. (2018)	94.46%	86.57%	84.66 $\pm$ 1.20	85.74 $\pm$ 1.36	83.53 $\pm$ 1.34
Inception-Resnet-V2	Szegedy et al. (2017)	93.21%	91.80%	85.42 $\pm$ 1.34	89.16 $\pm$ 1.05	85.05 $\pm$ 1.42
EfficientNet-B0	Tan and Le (2019)	95.55%	86.67%	79.04 $\pm$ 2.16	82.48 $\pm$ 1.15	79.96 $\pm$ 1.31
Darknet-19	Redmon (2020)	95.34%	88.34%	78.45 $\pm$ 3.21	81.65 $\pm$ 1.08	78.04 $\pm$ 1.03
Darknet-53	Redmon (2020)	93.98%	86.69%	81.56 $\pm$ 1.73	79.87 $\pm$ 1.00	83.32 $\pm$ 1.11
<b>AWDMC-Net [Our]</b>		<b>98.16%</b>	<b>95.40%</b>	<b>93.85 <math>\pm</math> 1.65</b>	<b>93.79 <math>\pm</math> 2.33</b>	<b>94.18 <math>\pm</math> 1.89</b>

Bold Face and Underlined—Most Outer Performed among CNN Method; Bold Face—Second Most Outer Performed CNN Method.

prediction performance of the proposed AWDMC-Net model for the classification of atmospheric/weather conditions on these available benchmark datasets is tabulated in Table 13 in terms of the average accuracy. Additionally, the state-of-the-art methods and the obtained accuracies on these similar datasets used for each research article are reported in Table 13. The proposed AWDMC-Net model can predict the classes of atmospheric/weather conditions with an average accuracy of 91.73% on the RFS dataset (Guerra et al., 2018), 92.72% on the MWI dataset (Zhang and Ma, 2015), 90.20% on the RESIDE dataset (Li

et al., 2018), 88.31% on the MWD dataset (Guerra et al., 2018), 89.36% on the Image2Weather dataset (Chu et al., 2016), and 91.04% on the WeatherDataset-4 dataset (Xia et al., 2020). Moreover, it can also be observed from Table 13 that the prediction performance of the proposed model (i.e., AWDMC-Net) was higher than that of the methods proposed by Zhang and Ma (2015), Zhang et al. (2016b), Chu et al. (2017), Guerra et al. (2018), and Xia et al. (2020) for all the considered datasets, except for the method proposed by Lin et al. (2017) on the MWD dataset with an average accuracy of 94.10%.

**Table 13**

Comparison of our proposed AWDMC-Net model with state-of-the-art weather image classification algorithms on available datasets.

Method, Year		Average accuracy					
		RFS (Guerra et al., 2018)	MWI (Zhang and Ma, 2015)	RESIDE (Li et al., 2018)	MWD (Lin et al., 2017)	Image2Weather (Chu et al., 2016)	WeatherDataset-4 (Xia et al., 2020)
Z. Zhang et al. 2015	Zhang and Ma (2015)	–	59.44%	–	–	–	–
Z. Zhang et al. 2016	Zhang et al. (2016b)	–	71.39%	–	–	–	–
W.T. Chu et al. 2017	Chu et al. (2017)	–	–	–	–	80.00%	–
D. Lin et al. 2017	Lin et al. (2017)	–	–	–	<b>94.10%</b>	–	–
J.C.V. Guerra et al. 2018	Guerra et al. (2018)	80.70%	–	–	–	–	–
J. Xia et al. 2020	Xia et al. (2020)	–	–	–	–	–	88.30%
<b>AWDMC-Net [Our]</b>		<b>91.73%</b>	<b>92.72%</b>	<b>90.20%</b>	<b>91.31%</b>	<b>89.36%</b>	<b>91.04%</b>

Bold Face—Most Outer Performed Method.

**Table 14**

Comparison of our proposed AWDMC-Net model with state-of-the-art weather image classification algorithms on E-TUVD dataset.

State-of-the-art methods		Performance measures		
		Accuracy (%)	Specificity (%)	Sensitivity (%)
M. Elhoseiny et al. 2015	Elhoseiny et al. (2015)	81.47 ± 1.10	78.45 ± 1.35	79.93 ± 1.29
C. Lu et al. 2016	Lu et al. (2014)	88.77 ± 0.98	89.13 ± 1.06	86.06 ± 1.13
X. Li et al. 2017	Li et al. (2017)	83.82 ± 1.33	80.39 ± 1.72	86.32 ± 1.56
J.C.V. Guerra et al. 2018	Guerra et al. (2018)	89.11 ± 2.14	91.03 ± 1.04	87.01 ± 1.88
M.R. Ibrahim et al. 2019	Ibrahim et al. (2019)	80.15 ± 1.02	76.97 ± 1.87	85.13 ± 1.04
B. Zhao et al. 2019	Zhao et al. (2019)	81.37 ± 2.11	77.83 ± 2.26	85.56 ± 1.64
Q.A. Al-Haija et al. 2020	Al-Haija et al. (2020)	<b>89.52 ± 1.31</b>	<b>91.06 ± 1.76</b>	<b>87.64 ± 2.04</b>
<b>AWDMC-Net [Our]</b>		<b>93.85 ± 1.65</b>	<b>93.79 ± 2.33</b>	<b>94.18 ± 1.89</b>

Bold Face and Underline—Most Outer Performed Method; Bold Face—Second Most Outer Performed Method.

**Table 15**

Performance (Training, Validation, and Testing performance) of our proposed AWDMC-Net model with atmospheric/weather image classification benchmark datasets.

Benchmark datasets		Training performance		Performance measures		
		Training accuracy	Validation accuracy	Accuracy (%)	Specificity (%)	Sensitivity (%)
MWI	Zhang and Ma (2015)	95.86%	93.89%	92.75 ± 1.16	88.98 ± 1.17	93.54 ± 2.02
RESIDE	Li et al. (2018)	94.20%	91.06%	90.89 ± 2.32	89.53 ± 2.56	91.11 ± 1.12
MWD	Lin et al. (2017)	96.43%	93.95%	92.79 ± 2.04	90.34 ± 1.04	94.86 ± 2.21
WeatherDataset-4	Xia et al. (2020)	96.95%	94.12%	91.66 ± 1.05	90.05 ± 2.09	93.45 ± 1.02

Moreover, comparison of the proposed AWDMC-Net model with the state-of-the-art deep learning based atmospheric/weather classification methods has been performed on our E-TUVD dataset. The state-of-the-art deep learning based methods used for comparison are: Zhang and Ma (2015), Lu et al. (2014), Li et al. (2017), Guerra et al. (2018), Ibrahim et al. (2019), Zhao et al. (2019), and Al-Haija et al. (2020). Each of these compared methods are fine-tuned so as to classify the six classes of atmospheric/weather conditions of the E-TUVD dataset. To evaluate and compare the prediction performance of state-of-the-art CNN based methods and our proposed network, similar training, validation, and testing set from the E-TUVD dataset as mentioned in Table 3 are used. The models after training the state-of-the-art methods are selected for testing based on the models for which training and validation accuracy is maximum. Table 14 reports the testing performance of the compared methods and our proposed AWDMC-Net model. The testing performance has been measured using the average accuracy, specificity, and sensitivity with the standard descriptor (i.e., mean ± standard deviation). Here also, in Table 14 the most outer performed method is represented by the boldface and underline and the second most outer performed method is represented by the bold face. It can be noticed from Table 14 that method proposed by Al-Haija et al. (2020) has achieved better prediction performance as compared to the other state-of-the-art methods with an average accuracy, specificity, and sensitivity of 89.52%, 91.06%, and 87.64% respectively. Also, it can be perceived that our proposed AWDMC-Net model is observed to be the best performing method with respect to all the state-of-the-art methods with an average accuracy, specificity, and sensitivity of 93.85%, 93.79%, and 94.18% respectively.

In real-world outdoor scenes, images acquired from various classes of atmospheric/weather conditions could be highly imbalanced. To validate the effectiveness of the proposed AWDMC-Net model with

respect to the imbalanced dataset, we have conducted experiments on some of the imbalanced datasets used by the research community for atmospheric/weather classification tasks. Similar to Table 13 of the revised manuscript, the used datasets are: MWI (Zhang and Ma, 2015), RESIDE (Li et al., 2018), MWD (Lin et al., 2017), and WeatherDataset-4 (Xia et al., 2020). These datasets are basically imbalanced with respect to the atmospheric/weather conditions. To address the class imbalance problem for atmospheric/weather classification tasks, we have used the method proposed by Huang et al. (2020) to balance the unbalanced classes of atmospheric/weather degraded conditions. From the four variants of their considered method, we have used the +GAN+ENN variant for data augmentation (i.e., balancing the classes of the considered atmospheric/weather conditions). This method adopts Edited Nearest Neighbor (ENN) (Wilson, 1972) with deep convolutional generative adversarial networks (DCGANs) (Radford et al., 2015) which basically cleans the unreliable data generated by DCGANs. To quantify the effectiveness of our proposed AWDMC-Net model on the balanced dataset generated by Huang et al. (2020), we have fine-tuned the last layer of our proposed AWDMC-Net depending upon the atmospheric/weather conditions present in the considered benchmark datasets thereby retaining the remaining network same. For training, validation, and testing of the proposed network (i.e., AWDMC-Net) on these balanced benchmark datasets, the network parameters are fixed as mentioned for training and testing our proposed AWDMC-Net on E-TUVD dataset (i.e., mentioned in Section 4.3). The prediction performance of the proposed AWDMC-Net model on a balanced set of atmospheric/weather classification benchmark datasets has been reported in Table 15. The training and validation accuracy for the best performed epoch for each of the considered atmospheric/weather degraded benchmark datasets is reported in Table 15. It can be observed that the proposed AWDMC-Net model can well predict the

**Table 16**  
Performance of our proposed AWDMC-Net model with respect to the Noise.

Noise		Performance measures		
		Accuracy (%)	Specificity (%)	Sensitivity (%)
Gaussian Noise	Rosin and Collomosse (2012)	87.43 ± 2.47	83.47 ± 1.22	89.95 ± 2.06
Salt and Pepper Noise	Rosin and Collomosse (2012)	86.18 ± 2.56	87.91 ± 2.07	87.86 ± 2.87
Poisson Noise	Rosin and Collomosse (2012)	86.02 ± 1.93	87.91 ± 1.94	85.01 ± 2.11
Speckle Noise	Rosin and Collomosse (2012)	88.38 ± 2.12	87.05 ± 2.53	89.98 ± 2.32

**Table 17**

Comparison of our proposed AWDMC-Net model with State-of-the-art weather image classification algorithms on noise imposed frames of E-TUVD dataset.

State-of-the-art methods		Accuracy (%)
M. Elhoseiny et al., 2015	Kroemer et al. (2010)	76.05
C. Lu et al., 2016	Lu et al. (2014)	<b>83.37</b>
X. Li et al., 2017	Li et al. (2017)	77.58
J.C.V. Guerra et al., 2018	Liu et al. (2016)	80.52
M.R. Ibrahim et al., 2019	Ibrahim et al. (2019)	72.33
B. Zhao et al., 2019	Zhao et al. (2019)	73.26
Q.A. Al-Haija et al., 2020	Al-Haija et al. (2020)	83.01
<b>AWDMC-Net [Our]</b>		<b>87.00</b>

Bold Face and Underline—Most Outer Performed Method; Bold Face—Second Most Outer Performed Method.

atmospheric/weather conditions in each of the considered balanced set of benchmark datasets with an average accuracy of 92.75% on the MWI dataset (Zhang and Ma, 2015), 90.89% on the RESIDE dataset (Li et al., 2018), 92.79% on the MWD dataset (Lin et al., 2017), and 91.66% on the WeatherDataset-4 dataset (Xia et al., 2020).

### 5.6. Influence of noise in the performance of the proposed AWDMC-Net model

Presence of noise in the atmospheric/weather degraded images is obvious when there are some issues in the sensors used to acquire the atmospheric/weather degraded images online. The effectiveness of our proposed AWDMC-Net model with respect to the noise externally imposed in the atmospheric/weather degraded frames is measured. For effective quantitative measurements of the proposed AWDMC-Net models, four types of noises i.e., Gaussian Noise (Rosin and Collomosse, 2012), Salt and Pepper Noise (Rosin and Collomosse, 2012), Poisson Noise (Rosin and Collomosse, 2012), and Speckle Noise (Rosin and Collomosse, 2012) are included in the atmospheric/weather degraded frames. In our proposed study, the experiment is conducted for a noise variance of 0.01 (i.e., applicable for Gaussian and Speckle noise). Table 16 reported the prediction performance (i.e., testing performance) of the proposed AWDMC-Net model on noise imposed atmospheric/weather degraded frames of E-TUVD dataset. In addition, comparison of the prediction performance of the proposed model with the state-of-the-art atmospheric/weather classification methods has been reported in Table 17. It can be observed from Tables 16 and 17 that the prediction performance of all the state-of-the-art atmospheric/weather classification methods and our proposed model decreases with the noise imposed atmospheric/weather degraded images of E-TUVD dataset. Also, in Table 17 it can be observed that our proposed model has achieved better prediction performance with the aforementioned four types of noises as compared to the state-of-the-art methods with an average accuracy, specificity, and sensitivity of 87.00%, 86.59%, and 88.20% respectively.

### 5.7. Running time

Due to the suitability of the atmospheric/weather classification tasks in real-time applications, including traffic signals and other supportive driving of automobiles, the computational time required for the classification of atmospheric/weather images is also a very significant parameter. In this subsection, we will compare the running time of the proposed AWDMC-Net model with respect to the state-of-the-art CNN

**Table 18**

Running time (in Seconds) comparison for predicting the atmospheric/weather conditions on a testing set of E-TUVD.

CNN Architecture(s)		Time (in Seconds)
VGG-16	Russakovsky et al. (2015)	119.04
VGG-19	Russakovsky et al. (2015)	146.21
AlexNet	Krizhevsky et al. (2012)	117.11
Inception-V3	Szegedy et al. (2016)	85.47
GoogLeNet	Szegedy et al. (2015)	60.57
Resnet-101	He et al. (2016)	77.69
Resnet-50	Huang et al. (2017)	58.31
Resnet-18	Huang et al. (2017)	58.50
Densenet-201	Huang et al. (2017)	57.55
Xception	Chollet (2017)	58.02
MobileNet-V2	Sandler et al. (2018)	70.25
Inception-Resnet-V2	Szegedy et al. (2017)	81.13
EfficientNet-B0	Tan and Le (2019)	63.45
Darknet-19	Redmon (2020)	65.56
Darknet-53	Redmon (2020)	61.47
<b>AWDMC-Net [Our]</b>		<b>54.31</b>

models. Table 18 lists the running time (in seconds (sec)) of each of the CNN models and our proposed CNN model for the prediction of atmospheric/weather conditions on the testing set of the E-TUVD dataset. Each image sequence from the testing set of E-TUVD had a resolution of 1920 × 1080 pixels. For comparison, each of these CNN models was tested on a CPU platform workstation with an Intel®Xeon®Processor E5-1620 v3 @3.50 GHz and 64 GB of installed memory (RAM). From Table 18, it can be observed that an excessively large number of trainable parameters of VGG-16 (Russakovsky et al., 2015), VGG-19 (Russakovsky et al., 2015), and AlexNet (Krizhevsky et al., 2012) leads to a very slow running time on the CPU, i.e., 119.04 s, 146.21 s, and 117.11 s, respectively. Conversely, the running time of our proposed model is only 54.31 s, which is comparatively less than that of the other CNN models/methods. Therefore, our proposed AWDMC-Net model is superior compared to other models both with respect to accuracy, and computational speed, and thus can be extensively applied to various sophisticated equipment in the field of computer vision tasks.

## 6. Conclusion

The presence of atmospheric/weather effects in outdoor scenes strongly affects the performance of computer vision systems due to the high loss in the color contrast and poor visibility. Therefore, the classification of atmospheric/weather degraded outdoor scenes is essential to effectively remove and thereby increase the visibility of the scenes. From this perspective, we proposed a new CNN network, namely, AWDMC-Net, for classification of different atmospheric/weather degraded outdoor scenes. Extensive experiments showed that our proposed network can well maintain the prediction performance and classifies atmospheric/weather conditions better compared to the other standard pre-trained CNN and state-of-the-art methods on our E-TUVD dataset and also on similar benchmark datasets. Henceforth, our proposed network has the potential to be incorporated into various real-time applications comprised of outdoor video monitoring and other driving systems. In our present study, we have designed the proposed AWDMC-Net model for classification of single atmospheric/weather conditions (i.e., fog, haze, dust, rain, poor illumination, clear day) pertaining to the scenes. In future, the proposed network will be extended by considering the blended atmospheric/weather degraded scenes (such as Fog + Rain, Rain + Haze, Fog + Poor Illumination/Low Light, etc.) for atmospheric/weather classification tasks.

## CRediT authorship contribution statement

**Sourav Dey Roy:** Methodology, Software, Validation, Investigation, Data creation, Formal analysis, Writing – original draft, Writing – review & editing, Visualization, Funding acquisition. **Mrinal Kanti Bhowmik:** Conceptualization, Methodology, Formal analysis, Project administration, Funding acquisition, Resources, Writing – review & editing, Visualization, Supervision.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This work was done in Computer Vision Laboratory of Computer Science and Engineering Department, Tripura University (A Central University), Suryamaninagar-799022, Tripura (W), India. The work was supported by the Defence Research and Development Organization (DRDO), Government of India, under Grant Number: ERIP/ER/DG-MED & CoS/990116507/M/01/1716, Dated: 6th March, 2018. The second author is grateful to DRDO, Government of India. The first author is also grateful to Council of Scientific & Industrial Research (CSIR), Government of India for providing him Senior Research Fellowship (SRF) under CSIR-SRF Fellowship Grant Number: 09/714(0020)/2019-EMR-I, Dated: 01/04/2019.

## References

- Al-Haija, Q.A., Smadi, M.A., Zein-Sabatto, S., 2020. Multi-class weather classification using ResNet-18 CNN for autonomous IoT and CPS applications. In: Proc. 2020 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE, pp. 1586–1591.
- An, J., Chen, Y., Shin, H., 2018. Weather classification using convolutional neural networks. In: Proc. 2018 International SoC Design Conference (ISODC). IEEE, pp. 245–246.
- Ancuti, C., Ancuti, C.O., De Vleeschouwer, C., 2016. D-hazy: A dataset to evaluate quantitatively dehazing algorithms. In: Proc. 2016 IEEE International Conference on Image Processing (ICIP). IEEE, pp. 2226–2230.
- Anghinolfi, D., Cannata, G., Mastrogiovanni, F., Nattero, C., Paolucci, M., 2013. On the problem of the automated design of large-scale robot skin. *IEEE Trans. Autom. Sci. Eng.* 10 (4), 1087–1100.
- Bodenhausen, L., Fugl, A.R., Jordt, A., Willatzen, M., Andersen, K.A., Olsen, M.M., Koch, R., Petersen, H.G., Krüger, N., 2014. An adaptable robot vision system performing manipulation actions with flexible objects. *IEEE Trans. Autom. Sci. Eng.* 11 (3), 749–765.
- Bore, N., Ekekrantz, J., Jensfelt, P., Folkesson, J., 2018. Detection and tracking of general movable objects in large three-dimensional maps. *IEEE Trans. Robot.* 35 (1), 231–247.
- Cai, Y., Luan, T., Gao, H., Wang, H., Chen, L., Li, Y., Sotelo, M.A., Li, Z., 2021. YOLOv4-5D: An effective and efficient object detector for autonomous driving. *IEEE Trans. Instrum. Meas.* 70, 1–13.
- Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2017. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* 40 (4), 834–848.
- Chen, Z., Yang, F., Lindner, A., Barrenetxea, G., Vetterli, M., 2012. How is the weather: Automatic inference from images. In: Proc. 2012 19th IEEE International Conference on Image Processing. IEEE, pp. 1853–1856.
- Chollet, F., 2017. Xception: Deep learning with depthwise separable convolutions. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition. pp. 1251–1258.
- Chopra, S., Hadsell, R., LeCun, Y., 2005. Learning a similarity metric discriminatively, with application to face verification. In: Proc. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 1. IEEE, pp. 539–546.
- Chu, W.T., Zheng, X.Y., Ding, D.S., 2016. Image2weather: A large-scale image dataset for weather property estimation. In: Proc. 2016 IEEE Second International Conference on Multimedia Big Data (BigMM). IEEE, pp. 137–144.
- Chu, W.T., Zheng, X.Y., Ding, D.S., 2017. Camera as weather sensor: Estimating weather information from single images. *J. Vis. Commun. Image Represent.* 46, 233–249, Elsevier.
- Cristianini, N., Shawe-Taylor, J., 2000. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge University Press.
- DataDrivenInvestor, 2021. Overview of different Optimizers for neural networks. [Online]. Available: <https://medium.com/datadriveninvestor/overview-of-different-optimizers-for-neural-networks-e0ed119440c3>.
- Deng, J., Dong, W., Socher, R., Li, L., Li, K., Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 248–255.
- E-TUVD, 2020. Extended Tripura University Video Dataset (E-TUVD) [Online]. Available: <http://mkbhowmik.in/eTuvd.aspx>.
- Elhoseiny, M., Huang, S., Elgammal, A., 2015. Weather classification with deep convolutional neural networks. In: Proc. 2015 IEEE International Conference on Image Processing (ICIP). IEEE, pp. 3349–3353.
- Gao, B., Pavel, L., 2017. On the properties of the softmax function with application in game theory and reinforcement learning. arXiv preprint arXiv:1704.00805.
- Garg, K., Nayar, S.K., 2004. Detection and removal of rain from videos. In: Proc. 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2004), Vol. 1. IEEE, p. 1.
- Guerra, J.C.V., Khanam, Z., Ehsan, S., Stolkin, R., McDonald-Maier, K., 2018. Weather classification: A new multi-class dataset, data augmentation approach and comprehensive evaluations of convolutional neural networks. In: Proc. 2018 NASA/ESA Conference on Adaptive Hardware and Systems (AHS). IEEE, pp. 305–310.
- Guo, D., Zhu, L., Lu, Y., Yu, H., Wang, S., 2018. Small object sensitive segmentation of urban street scene with spatial adjacency between object classes. *IEEE Trans. Image Process.* 28 (6), 2643–2653.
- Hadsell, R., Chopra, S., LeCun, Y., 2006. Dimensionality reduction by learning an invariant mapping. In: Proc. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), Vol. 2. pp. 1735–1742.
- He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask R-CNN. In: Proc. IEEE International Conference on Computer Vision. IEEE, pp. 2961–2969.
- He, K., Sun, J., Tang, X., 2011. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (12), 2341–2353.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition. Springer, pp. 770–778.
- Hu, H., Peng, R., Tai, Y.W., Tang, C.K., 2016. Network trimming: A data-driven neuron pruning approach towards efficient deep architectures. arXiv preprint arXiv:1607.03250.
- Huang, Y., Jin, Y., Li, Y., Lin, Z., 2020. Towards imbalanced image classification: A generative adversarial network ensemble learning method. *IEEE Access* 8, 88399–88409.
- Huang, G., Liu, Z., Maaten, L.V.D., Weinberger, K.Q., 2017. Densely connected convolutional networks. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition. pp. 4700–4708.
- Ibrahim, M.R., Haworth, J., Cheng, T., 2019. WeatherNet: Recognising weather and visual conditions from street-level images using deep residual learning. *ISPRS Int. J. Geo-Inf.* 8 (12), 549, MDPI.
- Joulin, A., Bach, F.R., Ponce, J., 2010. Discriminative clustering for image cosegmentation. In: Proc. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, pp. 1943–1950.
- Joulin, A., Bach, F.R., Ponce, J., 2012. Multi-class cosegmentation. In: Proc. 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 542–549.
- Kellenberger, B., Marcos, D., Lobry, S., Tuia, D., 2019. Half a percent of labels is enough: Efficient animal detection in UAV imagery using deep CNNs and active learning. *IEEE Trans. Geosci. Remote Sens.* 57 (12), 9524–9533.
- Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*. ACM, pp. 1097–1105.
- Kroemer, O.B., Detry, R., Piater, J., Peters, J., 2010. Combining active learning and reactive control for robot grasping. *Robot. Auton. Syst.* 58 (9), 1105–1116, Elsevier.
- Kurihata, H., Takahashi, T., Ide, I., Mekada, Y., Murase, H., Tamatsu, Y., Miyahara, T., 2005. Rainy weather recognition from in-vehicle camera images for driver assistance. In: Proc. IEEE Intelligent Vehicles Symposium. IEEE, pp. 205–210.
- Lalonde, J.F., Efros, A.A., Narasimhan, S.G., 2010. Detecting ground shadows in outdoor consumer photographs. In: Proc. European Conference on Computer Vision. Springer, pp. 322–335.
- Li, H., Kadav, A., Durdanovic, I., Samet, H., Graf, H.P., 2016. Pruning filters for efficient ConvNets. arXiv preprint arXiv:1608.08710.
- Li, Q., Kong, Y., Xia, S.M., 2014. A method of weather recognition based on outdoor images. In: Proc. 2014 International Conference on Computer Vision Theory and Applications (VISAPP), Vol. 2. IEEE, pp. 510–516.
- Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., Wang, Z., 2018. Benchmarking single-image dehazing and beyond. *IEEE Trans. Image Process.* 28 (1), 492–505.
- Li, X., Wang, Z., Lu, X., 2017. A multi-task framework for weather recognition. In: Proc. 25th ACM International Conference on Multimedia. pp. 1318–1326.
- Lin, D., Lu, C., Huang, H., Jia, J., 2017. RSCM: Region selection and concurrency model for multi-class weather recognition. *IEEE Trans. Image Process.* 26 (9), 4154–4167.

- Lins, R.G., Givigi, S.N., Kurka, P.R.G., 2015. Vision-based measurement for localization of objects in 3-D for robotic applications. *IEEE Trans. Instrum. Meas.* 64 (11), 2950–2958.
- Liu, Y.F., Jaw, D.W., Huang, S.C., Hwang, J.N., 2018. DesnowNet: Context-aware deep network for snow removal. *IEEE Trans. Image Process.* 27 (6), 3064–3073.
- Liu, H., Yu, Y., Sun, F., Gu, J., 2016. Visual–tactile fusion for object recognition. *IEEE Trans. Autom. Sci. Eng.* 14 (2), 996–1008.
- Long, J., Shelhamer, E., Darrell, T., 2016. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (4), 640–651.
- Lu, C., Lin, D., Jia, J., Tang, C.K., 2014. Two-class weather classification. In: *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 3718–3725.
- Luo, J.H., Zhang, H., Zhou, H.Y., Xie, C.W., Wu, J., Lin, W., 2018. Thinet: pruning cnn filters for a thinner net. *IEEE Trans. Pattern Anal. Mach. Intell.* 41 (10), 2525–2538.
- Martinez-Martin, E., Pobil, A.P.D., 2017. Object detection and recognition for assistive robots: Experimentation and implementation. *IEEE Robot. Autom. Mag.* 24 (3), 123–138.
- Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted boltzmann machines. In: *Proc. of the 27th International Conference on Machine Learning (ICML-10)*. ACM, pp. 807–814.
- Narasimhan, S.G., Nayar, S.K., 2002. Vision and the atmosphere. *Int. J. Comput. Vis.* 48 (3), 233–254, Springer.
- Narasimhan, S.G., Nayar, S.K., 2003. Contrast restoration of weather degraded images. *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (6), 713–724.
- Pavlic, M., Rigoll, G., Ilic, S., 2013. Classification of images in fog and fog-free scenes for use in vehicles. In: *Proc. 2013 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 481–486.
- Pitts, W., McCulloch, W.S., 1947. How we know universals the perception of auditory and visual forms. *Bull. Math. Biophys.* 9 (3), 127–147, Springer.
- Prateek Joshi, 2021. *Overfitting In Machine Learning*. [Online]. Available: <https://prateekjoshi.com/2013/06/09/overfitting-in-machine-learning/>.
- Radford, A., Metz, L., Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.
- Redmon, J., 2020. Darknet: Open source neural networks in c, [Online]. Available: <https://pjreddie.com/darknet/>.
- Roser, M., Moosmann, F., 2008. Classification of weather situations on single color images. In: *Proc. 2008 IEEE Intelligent Vehicles Symposium*. IEEE, pp. 798–803.
- Rosin, P., Collomosse, J., 2012. *Image and Video-Based Artistic Stylization*, Vol. 42. Springer Science & Business Media.
- Roy, S.D., Bhowmik, M.K., 2020. Annotation and benchmarking of a video dataset under degraded complex atmospheric conditions and its visibility enhancement analysis for moving object detection. *IEEE Trans. Circuits Syst. Video Technol. (Early Access)* 1.
- Rubio, J.C., Serrat, J., López, A.M., Paragios, N., 2012. Unsupervised cosegmentation through region matching. In: *Proc. 2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 749–756.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning internal representation by back-propagation of errors. *Nature* 323 (323), 533–536.
- Rusiecki, A., 2019. Trimmed categorical cross-entropy for deep learning with label noise. *Electron. Lett.* 55 (6), 319–320, IEEE.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., 2015. Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* 15 (3), 211–252, Springer.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C., 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. In: *Proc. IEEE Conference on Computer Vision and Pattern Recognition*. Springer, pp. 4510–4520.
- Shelhamer, E., Long, J., Darrell, T., 2017. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (4), 640–651.
- Shi, Y., Li, Y., Liu, J., Liu, X., Murphey, Y.L., 2018. Weather recognition based on edge deterioration and convolutional neural networks. In: *Proc. 2018 24th International Conference on Pattern Recognition (ICPR)*. IEEE, pp. 2438–2443.
- Shirmohammadi, S., Ferrero, A., 2014. Camera as the instrument: The rising trend of vision based measurement. *IEEE Instrum. Meas. Mag.* 17 (3), 41–47.
- Song, H., Chen, Y., Gao, Y., 2014. Weather condition recognition based on feature extraction and K-NN. In: *Foundations and Practical Applications of Cognitive Systems and Information Processing*. Springer, pp. 199–210.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15 (1), 1929–1958.
- Sutskever, I., Martens, J., Dahl, G., Hinton, G., 2013. On the importance of initialization and momentum in deep learning. In: *Proc. International Conference on Machine Learning*. ACM, pp. 1139–1147.
- Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A., 2017. Inception-v4, inception-resnet and the impact of residual connections on learning. In: *Proc. Thirty-first AAAI conference on artificial intelligence*. ACM, pp. 4278–4284.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. In: *Proc. IEEE Conference on Computer Vision and Pattern Recognition*. Springer, pp. 1–9.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision. In: *Proc. IEEE Conference on Computer Vision and Pattern Recognition*. Springer, pp. 2818–2826.
- Tan, M., Le, Q.V., 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. *arXiv preprint arXiv:1905.11946*.
- Wang, S., Li, Y., Liu, W., 2018. Multi-class weather classification fusing weather dataset and image features. In: *Proc. CCF Conference on Big Data*. Springer, pp. 149–159.
- Wilson, D.L., 1972. Asymptotic properties of nearest neighbor rules using edited data. *IEEE Trans. Syst. Man Cybern.* 3, 408–421.
- Xia, J., Xuan, D., Tan, L., Xing, L., 2020. ResNet15: Weather recognition on traffic road with deep convolutional neural network. In: *Advances in Meteorology*, Vol. 2020. Hindawi, pp. 1–11.
- Yamashita, R., Nishio, M., Do, R.K.G., Togashi, K., 2018. Convolutional neural networks: an overview and application in radiology. *Insights Imaging* 9 (4), 611–629, Springer.
- Yan, X., Luo, Y., Zheng, X., 2009. Weather recognition based on images captured by vision system in vehicle. In: *Proc. International Symposium on Neural Networks*. Springer, pp. 390–398.
- Zhang, Z., He, Z., Cao, G., Cao, W., 2016a. Animal detection from highly cluttered natural scenes using spatiotemporal object region proposals and patch verification. *IEEE Trans. Multimed.* 18 (10), 2079–2092.
- Zhang, Z., Ma, H., 2015. Multi-class weather classification on single images. In: *Proc. 2015 IEEE International Conference on Image Processing*. IEEE, pp. 4396–4400.
- Zhang, Z., Ma, H., Fu, H., Zhang, C., 2016b. Scene-free multi-class weather classification on single images. *Neurocomputing* 207, 365–373, Elsevier.
- Zhao, B., Hua, L., Li, X., Lu, X., Wang, Z., 2019. Weather recognition via classification labels and weather-cue maps. *Pattern Recognit.* 95, 272–284, Elsevier.
- Zhao, B., Li, X., Lu, X., Wang, Z., 2018. A CNN–RNN architecture for multi-label weather recognition. *Neurocomputing* 322, 47–57, Elsevier.
- Zhao, X., Liu, P., Liu, J., Tang, X., 2011. Feature extraction for classification of different weather conditions. *Front. Electr. Electron. Eng. China* 6 (2), 339–346, Higher Education Press.
- Zhu, Z., Zhuo, L., Qu, P., Zhou, K., Zhang, J., 2016. Extreme weather recognition using convolutional neural networks. In: *Proc. 2016 IEEE International Symposium on Multimedia (ISM)*. IEEE, pp. 621–625.



**Sourav Dey Roy** received the B.E. and M.Tech. degrees in computer science and engineering from Tripura University (A Central University), Suryamaninagar, Tripura, India, in 2013 and 2015, respectively, where he is currently pursuing the Ph.D. degree in computer science and engineering, under the sole supervision of Mrinal Kanti Bhowmik. He has worked as a Junior Research Fellow (JRF) under a Defense Research and Development Organization (DRDO) funded project, Government of India, from June 2018 to April 2019. He is currently working as a Senior Research Fellow (SRF), sponsored by the Council of Scientific and Industrial Research (CSIR), Government of India. His research interests include computer vision, object detection, face and facial expression recognition, medical imaging, and deep learning.



**Mrinal Kanti Bhowmik** received the B.E. degree in computer science and engineering from the Tripura Engineering College 1335 in 2004, the M.Tech. degree in computer science and engineering from Tripura University (A Central University), India, in 2007, and the Ph.D. degree in engineering from Jadavpur University, Kolkata, India, in 2014. He has successfully completed two Department of Electronics and Information Technology (DeitY) (Now Ministry of Electronics and Information Technology (MeitY)) funded projects, one the Department of Biotechnology (DBT)-Twinning project, one Society for Applied Microwave Electronics Engineering and Research (SAMEER) funded project, and one Indian Council of Medical Research (ICMR) project as the Principal Investigator. He is currently the Principal Investigator of a Defense Research and Development Organization (DRDO), Government of India funded project. Since July 2010, he has been serving with the Department of Computer Science and Engineering, Tripura University as an Assistant Professor. He was awarded the Short Term Indian Council of Medical Research (ICMR), Department of Health Research (DHR) International Fellowship from 2019 to 2020 as a Senior Indian Biomedical Scientist for bilateral cooperation in cross-disciplinary research area (i.e., biomedical diagnostic and inferencing systems). His current research interests are in the field of computer vision, security and surveillance, medical imaging, and biometrics.