Multimedia Tools and Applications (2023) 82:15415–15438
https://doi.org/10.1007/s11042-022-13958-x

Smart parking space detection under hazy conditions using convolutional neural networks: a novel approach

Gaurav Satyanath[1] · Jajati Keshari Sahoo[2] · Rajendra Kumar Roul[3]

Received: 17 February 2022 / Revised: 3 June 2022 / Accepted: 13 September 2022 /
Published online: 4 October 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
Limited urban parking space combined with urbanization has necessitated the development of smart parking systems that can communicate the availability of parking slots to the end-users. Towards this, various deep learning based solutions using convolutional neural networks have been proposed for parking space occupation detection. Though these approaches are robust to partial obstructions and lighting conditions, their performance is found to degrade in the presence of haze conditions. Looking in this direction, this paper investigates the use of dehazing networks that improves the performance of parking space occupancy classifier under hazy conditions. Additionally, training procedures are proposed for dehazing networks to maximize the performance of the system under both hazy and non-hazy conditions. The proposed system is deployable as part of existing smart parking systems, where a limited number of cameras are used to monitor hundreds of parking spaces. To validate our approach, we have developed a custom hazy parking system dataset from real-world task-driven test set of RESIDE-β dataset. The proposed approach is tested against existing state-of-the-art parking space detectors on CNRPark-EXT and hazy parking system datasets. Experimental results indicate a significant accuracy improvement of the proposed approach on the hazy parking system dataset.

Keywords Classification · Computer vision · Convolutional neural networks · Deep learning · Hazy parking · Image Dehazing · IoT · Smart parking system

✉ Gaurav Satyanath
gsatyana@andrew.cmu.edu

Jajati Keshari Sahoo
jksahoo@goa.bits-pilani.ac.in

Rajendra Kumar Roul
raj.roul@thapar.edu

[1] Department of Electrical & Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA

[2] Department of Mathematics, BITS Pilani K K Birla Goa Campus, Goa, India

[3] Department of Computer Science & Engineering, Thapar Institute of Engineering & Technology, Patiala, Punjab, India

1 Introduction

The increase in population coupled with rapid urbanization has led to an increase in vehicle footprints in cities. With the urban space being limited, this has put a lot of strain on existing parking systems leading to an increase in time spent by a person searching for a parking slot. According to Giuffre et al. [21], and Shoup et al. [56], around 25% to 40% of urban traffic flow are searching for an effective parking slot. Further, a survey conducted by Lin et al. [37] indicates that the driver searching for a parking slot has an increased possibility of accidents due to their lack of attention to the road. This has necessitated the development of smart parking solutions, which could communicate the availability of parking slots in advance to the end-user, thereby reducing the time spent on the road searching for a vacant parking slot. Further, the current global market value of the parking industry exceeding 2 billion US dollars has also fuelled the development of smart parking solutions. The parking industry is projected to have a compounded growth of 17% over the next decade, of which 65% is projected to be from smart parking solutions.[1] Many research works have been done in this domain [6, 9, 13, 15, 54, 61, 63, 67].

Various existing smart parking solutions have used sensors to determine the occupancy of parking slots and communicate the same to the end-user using Bluetooth/WiFi modules [2, 16, 27, 30, 31, 40, 65]. However, these solutions require the installation and constant sensor maintenance at each parking slot, making it difficult to scale to a larger IoT setting. Vision-based parking slot occupation detection offers an alternative cost-effective and scalable solution where a limited number of cameras is used to monitor hundreds of parking spaces [7, 8, 17, 26, 43, 60]. However, obstruction of parking slots (for example, from trees) and changing weather conditions degrade the performance of such vision-based techniques. This problem was surpassed in [7, 8] by exploiting the robustness of Deep Neural Networks (DNN). They developed a Convolutional Neural Network (CNN) based parking space occupancy classifier for that purpose.

1.1 Motivation

Despite these fine efforts, vacant parking slot detection under hazy conditions is still an open problem. In densely populated areas and areas closer to industries, the atmosphere is polluted with smoke, dust, and other particles that drastically reduce the visibility. Further, during snowfall and rainfall, the visibility is compromised. Hence, detecting parking slot occupancy under hazy conditions is challenging. All the above-mentioned vision-based approaches are not generalizable to such hazy conditions, and their performance degrades in such situations to a large extent.

Working in this direction, the proposed work focuses on improving the accuracy of vision-based parking space occupation detection under both hazy and non-hazy conditions. To the best of our knowledge, the proposed approach is the first work that tackles the problem of parking slot occupation detection under hazy conditions.

1.2 Contributions

The significant contributions of this paper include the following:

[1] Smart Parking Market worth 5.25 Billion USD by 2021. https://www.marketsandmarkets.com/PressReleases/smart-parking.asp

i. A vision-based parking slot occupancy detection system is proposed that consists of the following two networks in series: an end-to-end dehazing network and a parking slot classifier (CNN). For the dehazing network, we follow *All-in-One Dehazing network* ($AOD\text{-}Net$) architecture [35] owing to its lower computational cost than other state-of-the-art dehazing networks. For the parking space occupancy classifier, we follow the $mAlexnet$ architecture [8] owing to its low computational cost and robustness of the architecture to changing weather conditions. The proposed system is robust to partial occlusions, changing weather conditions, and the presence of haze in the image. Further, the system is deployable as part of existing smart parking systems and scalable to IoT settings.
ii. Various training procedures are explored to maximize the system's accuracy under both hazy and non-hazy conditions. These include:
   (a) Inclusion of non-hazy images as part of the training of dehazing network.
   (b) Use of modified loss function during training of dehazing network incorporating a new hyperparameter for tuning the relative performance of the system on hazy and non-hazy images.
   (c) Joint optimization of dehazing network and classifier.
iii. The proposed approach introduces a custom hazy parking system dataset consisting of 5010 real-world hazy occupied and unoccupied parking patches extracted from RTTS (Real-world Task-driven Test Set) subset of RESIDE-β dataset [36]. Optional non-overlapping training and test split are also provided for benchmarking purposes. The dataset includes patches captured under various intensities of haze as well as different types of hazy conditions (e.g., snow, fog, etc.).

Experimental results show that the use of dehazing network significantly improves the parking space classification accuracy (around 10-15%) on the proposed hazy parking system dataset.

The rest of the paper includes the following: Basic preliminaries used for the proposed approach are discussed in Section 2. Section 3 focused on the related works in this domain. The proposed approach is discussed in Section 4, and the experimental work is carried out in Section 5. Finally, the paper is concluded in Section 6 with some future enhancements.

2 Preliminaries

2.1 Atmospheric scattering model

The atmospheric scattering model proposed in [25, 42, 44, 45] for hazy image generation is given in (1).

$$I(x) = J(x)t(x) + A(1 - t(x)). \tag{1}$$

Here A denotes the global atmospheric light, $J(x)$ is the clear image to be recovered, $I(x)$ is the observed hazy image, and $t(x)$ is the transmission matrix given in (2).

$$t(x) = e^{-\beta d(x)} \tag{2}$$

where β denotes the scattering coefficient of atmosphere and $d(x)$ denotes the distance between the camera and the object.

Dehazing methods involve estimating the transmission matrix $t(x)$ using physically grounded information or data-driven methods and the global atmospheric light A using empirical methods. The clear image is then computed using (1).

2.2 Convolutional neural network

Convolution Neural Networks (CNNs) belong to a class of artificial neural networks (ANNs) which are known for their significantly good performance in image classification [24], numeral recognition [49], face recognition [23], object detection [18], *etc.* CNN consists of a sequence of layers which transform the input using filters (kernels). Various existing classical CNN architectures include LeNet-5 [34], AlexNet [32], VGG-16 Net [57], ResNet [59], Inception Net [58], *etc.* Figure 1 depicts the basic structure of CNN which includes three major components *i.e.* convolution layers, pooling layers and fully connected layers as discussed below.

2.2.1 Convolutional layers

Convolutional layers form the core of CNN's. Convolutional layer consists of convolving the given input with a set of weight matrices (3-dimensional:- Height, Width, and Depth). The weight matrices are slided across the width and height of the input. At each position, dot product between the weight matrices and the input is aggregated, and a bias is added to calculate the final output. Each weight matrix is responsible for producing a single feature map. The maps generated by different weight matrices are concatenated along the depth dimension to form the final output. Use of the same weight matrices at various positions of the input (i.e., weight sharing) decreases the number of parameters of the network.

2.2.2 Pooling layers

Generally, every convolution layer is followed by a pooling layer. Pooling layers are introduced to downsize the feature maps to reduce the number of parameters and thereby the computation of the network. The pooling operation involves sliding a two-dimensional window over each channel of the input map and summarizing the content within with a mathematical operation like max or average etc. Pooling layers help the network robust

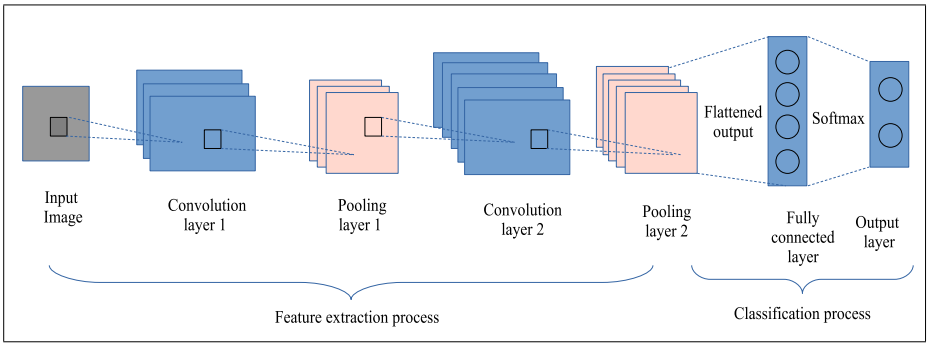

Fig. 1 Structural diagram of CNN

to variations in the position of the features as it involves summarizing the features over a region. This makes the network robust to position the desired object in the image.

2.2.3 Fully connected layer

Fully connected layer mimics the regular neural layer of a Multi-layer Perceptron (MLP), and it performs an affine transformation of the input. Fully connected and pooling layers are generally followed by non-linear activation layers like ReLU, sigmoid etc.

3 Related work

3.1 Image Dehazing

3.1.1 Dehazing networks

Much effort has been used to calculate the transmission matrix $t(x)$ accurately. He et al. [25] calculated the transmission matrix by discovering the dark channel prior (DCP) of the image. Zhu et al. [66] estimated the transmission map using a linear model to map the scene depth of the hazy image whose parameters are learned in a supervised way using colour attenuation prior.

Recently CNN's have been used for haze removal. Ren et al. [51] have proposed a Multi-Scale-Convolutional-Neural-Network (MSCNN) which takes a hazy image as input and outputs a coarse-scale transmission matrix which is then refined by the second fine-tuning network. In [12], a dehazing network called Dehazenet is proposed, which takes a hazy image as input and outputs the corresponding transmission matrix. Both these approaches then calculate the global atmospheric light A using empirical methods and the clear image using the atmospheric scattering model. However, in all the approaches mentioned above, the errors in the estimation of transmission matrix $t(x)$ and global atmospheric light A will accumulate and amplify each other leading to sub-optimal results [35].

To minimize the errors during restoration, recent works [35, 39, 52] are focused on developing an end-to-end dehazing solution. Li et al. [26] developed a novel end-to-end dehazing network based on CNN called $AOD\text{-}Net$. They studied its use for high-level vision tasks such as object detection and recommended joint optimization of the pretrained dehazing network and the object detector for better performance on hazy images. Similarly, Liu et al. [24] also proposed an end-to-end CNN architecture called "$GridDehazenet$" for single image dehazing. Working in the same direction, Ren et al. [25] suggested an end-to-end encoder-decoder based CNN called *"Gated Fusion Network (GFN)"* for single image dehazing. Though $GridDehazeNet$ [39] and GFN [52] achieve better dehazing results than $AOD\text{-}Net$, they are computationally more expensive. This complicates the implementation of $GridDehazenet$ and GFN at the system's edges for real-time applications.

3.2 Smart parking system

In the early days, smart parking systems were built by embedding sensors at each parking spot. Mustaffa et al. [31] used ultrasonic sensors fixed at the ceiling of each parking spot and displayed the availability at the entrance gate. They displayed directional signage to assist the drivers in finding a parking slot. Idris et al. [27] have suggested the use of ultrasonic

sensors at each parking slot and used a wireless sensor network to transmit this information to the entry gate. Khanna et al. [30] used infrared sensors at the parking slots, and stored the occupancy information in the cloud server that is made available to the users through mobile applications. Recent advances have tried to extend the current approaches to larger parking slots [1, 4, 14, 50]. However, these approaches require the installation and maintenance of sensors at each parking slot incurring huge costs, especially when the number of parking slots is large.

Amato et al. [8] employed mAlexNet and mLeNet CNN architectures for parking space detection which proved to be effective for various datasets including CNRPark, PKLot and an unseen test set. The results showed better performance of mAlexNet as compared to mLeNet. But their model needs generalization for night, foggy, or snowy conditions. Nyambal et al. [48] presented a CNN based model for performing real-time classification of parking space. The obtained results showed an accuracy of 93.64% using LeNet and 95.49% using AlexNet architecture. The quality of parking lot images used in their work needs to be improved for better feature extraction. Xiang et al. [64] proposed a novel approach for parking occupancy detection based on Haar-AdaBoosting algorithm and CNN. The proposed model resulted in lower false positive rate and an accuracy of more than 95%. The classifier's performance is unstable, and it may sometime fails to recognize some special vehicles. Gkolias et al. [22] proposed CNN based approach for parking space detection on streets. The performance comparison was done with complex SVM models where CNN model outperformed all others with an accuracy of 90%. But the model requires appropriate training to extend it for other real-time conditions. Nurullayev et al. [47] applied dilated CNN to identify car occupancy in parking lot (CarNet). The experiments were conducted on publicly available PkLot and CNNPark + EXT datasets and a comparative analysis with AlexNet and other well-known architectures showed the efficacy and reliability of the proposed approach. Further investigations are still required to improve the classification accuracy. Vscekic et al. [55] presented a faster region based CNN (R-CNN) to detect parking space occupancy as empty or occupied. The model achieved satisfactory results but still requires some improvements. Their approach needs improvement in terms of hyperparameter tuning and the size of the dataset used. Naufal et al. [46] employed preprocessed R-CNN (Mask R-CNN) for predicting parking space availability and achieved an accuracy of 73.73%. Their system faces lighting issues and is unable to detect when there are obstructions in front of the car. More review works on smart parking system using CNN can be found in [10, 19, 41].

Vision-based techniques offer a scalable and cost-effective solution as it requires only a few cameras for monitoring the entire parking system. De Almeida et al. [17] developed a new parking system dataset called PKLot consisting of around 700,000 parking space images captured from three different cameras. The authors also proposed two systems for vacancy verification using textual descriptors: one using Local Binary Pattern (LBP) and the other using Local Phase Quantization (LPQ) [5]. Both the systems used SVM classifier and achieved an error rate of 11% on the test set. The cameras are placed at the top of the building to solve the occlusion problems.

True [60] used colour histogram and Difference-of-Gaussian (DoG) using SVM and Linear Discriminant Analysis (LDA) for classification. He achieved an average accuracy of 96.45% on images which are not used during the training phase. However, the problem of partial occlusion due to trees, snow, fog, etc., is not addressed in his work.

With the popularization of deep learning, various Deep Neural Networks (DNN's) have also been proposed for parking space occupation classification [53]. Among deep learning approaches, Convolutional Neural Networks (CNN's) are found to be most effective for

vision-based detection and segmentation tasks [20, 33, 57]. In [8], a parking system dataset called $CNRPark$ consisting of 12,584 parking space patches captured from 2 different viewpoints was introduced. Besides introducing the dataset, they also have proposed the use of a mini version of AlexNet [33] architecture for parking space detection and proved it to be robust to light changes and partial occlusions.

Amato et al. [7] further extended the $CNRPark$ dataset to $CNRPark\text{-}EXT$ dataset consisting of 144,965 parking space patches captured from 9 different cameras under various weather conditions. Using the $CNRPark\text{-}EXT$ dataset, they proved the robustness of $mAlexnet$ architecture for changing camera views and weather conditions. Bura et al. [11] developed a smart parking system where the license plate of the cars was captured at the entrance and used wide eye cameras for monitoring the multiple parking slots. For the real-time occupancy classification, they developed a computationally inexpensive CNN architecture. More survey works in this domain can be found in [3, 29, 37].

All the approaches mentioned above are best suitable for non-hazy conditions, and their performances are severely degraded in the presence of haze. The proposed parking system can overcome such limitations and works well both in hazy and non-hazy conditions. Experimental results show the efficiency of the proposed approach over the conventional approaches.

4 Proposed system

4.1 System overview

The proposed parking system enhances the model proposed in [8] and is further extended to tackle hazy conditions by incorporating a dehazing network into the system's flow. The model proposed in [8] consists of a set of cameras to monitor the entire parking lot. Periodically, images are captured from the cameras to determine the occupancy of parking spaces. From each image, all the parking spaces captured by the camera are extracted using manually created predetermined masks. Each parking space is then classified using the trained CNN to determine its occupancy.

The dehazing network can be employed in two different ways. In the first case, the images captured from the cameras are fed to the dehazing network. The dehazed output is then segmented into individual parking spaces and classified using the CNN. In the second case, the images captured from cameras are first segmented into parking spaces, and then each parking space is fed to the dehazing network followed by the classification process using CNN. Computationally, the former case is more efficient as it uses dehazing network once for all the parking spaces in the image. In contrast, the latter case uses the dehazing network for each parking space. However, the latter case enables jointly optimizing the dehazing network and classifier as a single pipeline that can be used to achieve higher accuracy in hazy conditions.

The choice is dependent on the computational resources available, the number of parking spaces in the parking lot, and the period between successive capture of images by the cameras. In this paper, the second case is adopted for experimental works. Accordingly, the classifier and dehazing network are developed considering the computational feasibility of the system.

The proposed system consists of five stages as follows:

Stage 1: Images of the parking lot are captured using cameras.

Stage 2: The portion of parking slots from each image is then extracted using predetermined masks. Each segmented image is resized to 224×224. Images are rescaled by $\frac{1}{256}$.
Stage 3: Each parking slot is passed through the dehazing network.
Stage 4: The parking slot occupancy is determined using CNN.
Stage 5: Finally, all the classification results are aggregated.

Figure 2 demonstrates the overview of the proposed approach.

The following sections describe how the CNN is used for the parking space classification process and how the dehazing network is used in the proposed model.

4.2 Parking space classification

For parking space classification, the proposed system followed the mAlexnet proposed in [8]. mAlexnet is a miniaturized version of Alexnet proposed in [33], specifically built for binary classification tasks. It is computationally cheaper and is found to achieve accuracies similar to that of Alexnet for binary classification tasks [7]. mAlexnet consists of three convolutional layers and two fully connected layers. Each convolutional layer is followed by ReLU activation and max pooling layers. The fully connected layers FC4 and FC5 are followed by ReLU and SoftMax activations, respectively.

The mAlexnet architecture proposed in [8] uses Local response normalization. However quantitative experiments performed by Z. Wang et al. in [62] on CIFAR-10 dataset using various CNN networks showed batch normalization to be a better normalization technique than local response normalization. Taking inspiration from [62], the proposed system uses batch normalization layer instead of local response normalization layer (LRN) after first two convolutional layers. The details of the modified mAlexnet architecture used in the proposed system are illustrated in Table 1. The first row IS refers to the input size of the feature map represented by $Channels \times Height \times Width$ for convolutional layers and the number of neurons for fully connected layers. For conv1-3, the number and size of filter is specified by "num×width×height+stride" along with ReLU activation and batch normalization. Max pooling layer is denoted by "width×height+stride". The dimensionality and the activation function are shown for fully connected layers. Padding is not used for any of the layers.

4.3 Dehazing network

The proposed system follows the *AOD-Net* proposed in [26]. It combines all the parameters of atmospheric scattering equation by rewriting it as follows.

$$J(x) = \frac{I(x)}{t(x)} - \frac{A}{t(x)} + A \tag{3}$$

$$= K(x)I(x) - K(x) + b, where \tag{4}$$

$$K(x) = \frac{\frac{(I(x)-A)}{t(x)} + (A-b)}{I(x)-1} \tag{5}$$

b is constant bias with a default value of 1. With $t(x)$ and A incorporated into a single parameter $K(x)$, a CNN architecture named as the K-estimation module is trained to learn the parameter $K(x)$ given an input image $I(x)$. It consists of five convolutional layers and combines features from different layers to compensate for information loss during convolution. *"concat1"* combines features of *"conv1"* and *"conv2"*.*"concat2"* combines features

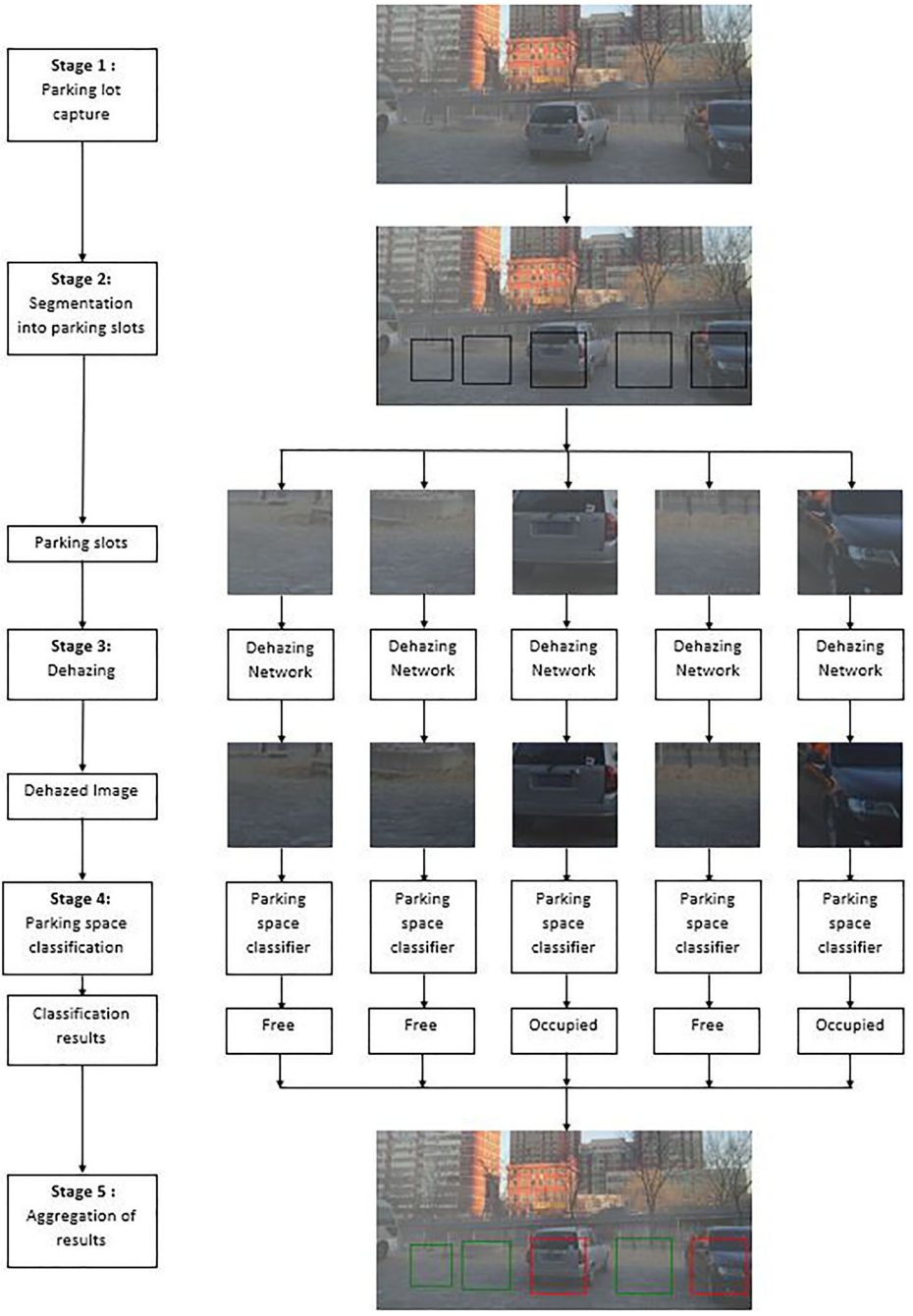

Fig. 2 Overview of the proposed approach

Table 1 Modified mAlexnet architecture used in the proposed approach

Conv1	Conv2	Conv3	Fc4	Fc5
IS: 224x224x3	IS: 16x27x27	IS: 20x11x11	IS: 480	IS : 48
16x11x11+4	20x5x5+1	30x3x3+1	–	–
pool 3x3+2	pool 3x3+2	pool 3x3+2	48	2
Batchnorm	Batchnorm	–	–	–
ReLU	ReLU	ReLU	ReLU	Softmax

of *"conv1","conv2"* and *"conv3"*. *"concat3"* combines features of *"conv1","conv2", "conv3"* and *"conv4"*. Each convolutional layer is followed by a ReLU activation layer.

The K-estimation module architecture is depicted in Fig. 3 [35] and the details are illustrated in Table 2. The first row IS refers to the input size of the feature map represented by $Channels \times Height \times Width$ for convolutional layers. For each convolution layer, the number and the size of filter are specified by "num×width×height+stride, pad" followed by ReLU activation layer. Once $K(x)$ is estimated, the clear image is obtained using (4).

Experimentally, AOD-Net [26] is found to have lower average runtime than other dehazing networks [39]. Given the computational constraints of the system, AOD-Net is the ideal choice when compared to other works such as GridDehazenet [39], Gated Fusion Network [52], MSCNN [51], and Dehazenet [12] owing to its small network architecture.

5 Experimental analysis

5.1 Datasets used for evaluation

5.1.1 Parking lot datasets

For parking space occupancy classification, the proposed system used CNRPark [8], CNRPark-EXT [7] and PKLot [17] datasets. PKLot consists of around 700,000 manually checked and labeled parking space images acquired from parking lots of Federal University of Parana (UFPR) and Pontifical Catholic University of Parana (PUCPR) located in Curitiba, Brazil. The dataset consists of images captured from different parking lots with

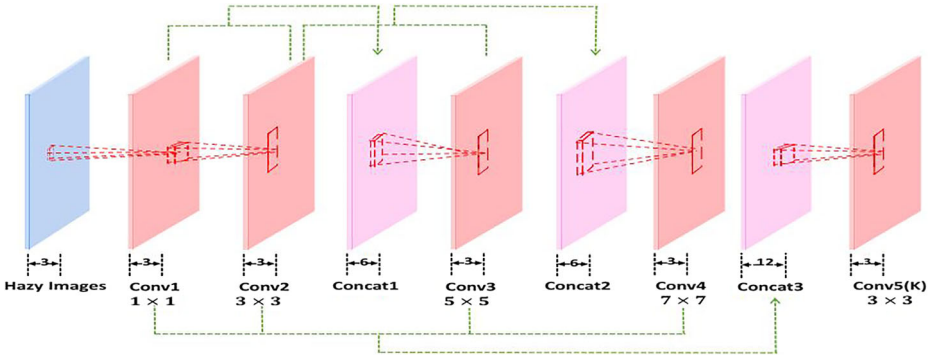

Fig. 3 The network architecture of K-estimation module

Table 2 K-estimation module architecture

Conv1	Conv2	Conv3	Conv4	Conv5
IS: 3x224x224	IS: 3x224x224	IS: 6x224x224	IS: 6x224x224	IS : 12x224x224
3x1x1+1,0	3x3x3+1,1	3x5x5+1,2	3x7x7+1,3	3x3x3+1,1
ReLU	ReLU	ReLU	ReLU	ReLU

distinct features and is divided into three subsets: UFPR04, UFPR05, and PUCPR, corresponding to images captured by cameras at 4^{th} and 5^{th} floor of UFPR, and at 10^{th} floor of PUCPR respectively. With the cameras placed at the top of the buildings, the dataset contains fewer occluded images. The dataset also contains images under different weather conditions (Sunny, Overcast, Rainy) and various illuminations occasioned by weather conditions. The dataset contains both non-segmented full images of the parking lot and pre-segmented parking patches. The slope of the pre-segmented parking space rectangle is rotated such that those with a slope between 0° to 45° were rotated to 0° and those with a slope between 45° to 90° were rotated to 90°. Details of the PKLot dataset are shown in Table 3.

The CNRPark (CNRPark A and CNRPark B) dataset consists of 12,584 images segregated into two non-overlapping subsets: CNRPark A (Images captured using camera A) consisting of 6,171 images and CNRPark B (Images captured using camera B) consisting of 6,413 images. The dataset includes patches captured under various light conditions. It also contains patches partially occluded by trees as well as shadowed by the neighboring cars. This enables us to demonstrate the robustness of the classifier to changing real-life conditions. Further, one can also test the robustness of the classifier to changing camera views by training on images captured from one camera and testing on images captured from different cameras. CNRPark A contains fewer occluded images than CNRPark B.

CNRPark dataset is extended to CNRPark-EXT in [7]. CNRPark-EXT consists of 144,965 image patches captured from 9 different cameras. Those patches are captured under various weather conditions (Sunny, Overcast, and Rainy) as well as at different distances from the camera. The dataset also contains patches partially occluded or shadowed, enabling to train the classifier on various difficult real-life scenarios. The dataset contains both non-segmented full images of the parking lot and pre-segmented parking patches. Further, different training, validation, and test splits are provided to enable common grounds for testing the classification algorithms. Details of the CNRPark and CNRPark-EXT datasets used in the experiments are shown in Table 3.

5.1.2 Single image dehazing datasets

For training the dehazing network, the proposed system used Outdoor Training Set (OTS) of Reside-β dataset [36]. OTS dataset consists 72,135 synthetic outdoor hazy images synthesized from 2061 diverse non-hazy outdoor images.[2] Each non-hazy image was used to synthesize 35 synthetic hazy images by using different pairs of A and β of atmospheric scattering model described in (1). Values of $A \in$ [0.8-1] with increment of 0.05 and $\beta \in [0.04, 0.06, 0.08, 0.1, 0.12, 0.16, 0.2]$. This captures various hazy conditions encoun-

[2]"Beijing realtime weather photos," http://goo.gl/svzxLm.

Table 3 Details of CNRPark, CNRPark-EXT and PKLot datasets [7, 8, 17]

Subsets	Free patches	Busy patches	Total patches
CNRPark A	2549	3622	6171
CNRPark B	1632	4781	6413
CNRPark	4181	8403	12584
CNRPark-EXT	65658	79307	144965
CNRPark-EXT TEST	13549	18276	31825
CNRPark + EXT Train C1-C8	16784	21769	38553
CNRPark + EXT Train	51059	56018	107077
PKLot Train	27314	41744	69058
PKLot Test	275894	248583	524477

tered in real life. The depth map for each image was estimated using the procedure outlined in [38]. To counter the errors during depth estimation, the final hazy images were visually inspected for any irregularities.

5.1.3 Hazy parking system dataset

The proposed system creates a benchmark for the hazy parking system using the Real-time Task-driven Test Set (RTTS) of RESIDE-β dataset [36]. RTTS consists of 4322 images extracted from the Web covering real-world hazy traffic and driving scenarios. Each image is annotated with bounding boxes and object categories. Hazy car images and hazy unoccupied parking space patches are manually extracted to create the hazy parking system dataset. Optional non-overlapping training and test splits are provided to compare classification results under hazy conditions.

The dataset contains 752 unaugmented patches with 502 hazy car patches and 250 hazy free patches. It is splited into non-overlapping 70% training and 30% test set. The test set contains roughly the same number of busy and free patches before augmentation. The images in the training and test set are then augmented using Keras[3] (a deep learning framework) image data pre-processing tool. The hazy car patches are randomly flipped horizontally, and a crop of the image containing 90% of the total height and width is considered. Hazy unoccupied patches are also augmented in a similar setting where additionally, the patches are randomly flipped in a vertical manner. The dataset consists of 5010 augmented labeled patches and captures busy and unoccupied patches under various hazy conditions such as fog, snow, and pollution with varying intensity levels from lightly hazed to heavily hazed patches. Details of the dataset are shown in Table 4. For demonstration purposes, we have shown some of the sample images of the proposed hazy parking system dataset in Fig. 4. Out of 5010 images, 500 images (250 of the training set and 250 of the test set) of the hazy parking system dataset are made available publicly for download.[4]

[3]Chollet, Franccois et al. "Keras." https://keras.io. (2015).
[4]https://github.com/GauravS9776/Hazy-parking-system

Table 4 Details of Hazy parking system dataset

Datasets	Free patches	Busy patches	Total patches
Unaugmented Train	150	380	532
Unaugmented Test	100	122	222
Unaugmented images	250	502	752
Hazy parking system Train	1500	1900	3400
Hazy parking system Test	1000	610	1610
Hazy parking system	2500	2510	5010

5.2 Parking slot classifier

Amato et al. [8] achieved state-of-the-art results using the mAlexnet CNN architecture. The mAlexnet network takes an input image of size 224×224. They trained mAlexnet on Caffe framework [28] using gradient descent with momentum. Momentum and weight decay were set to 0.9 and 5×10^{-4} respectively. The learning rate was individually determined for each experiment and decreased by a factor of 10 after loss stabilization. All the models were

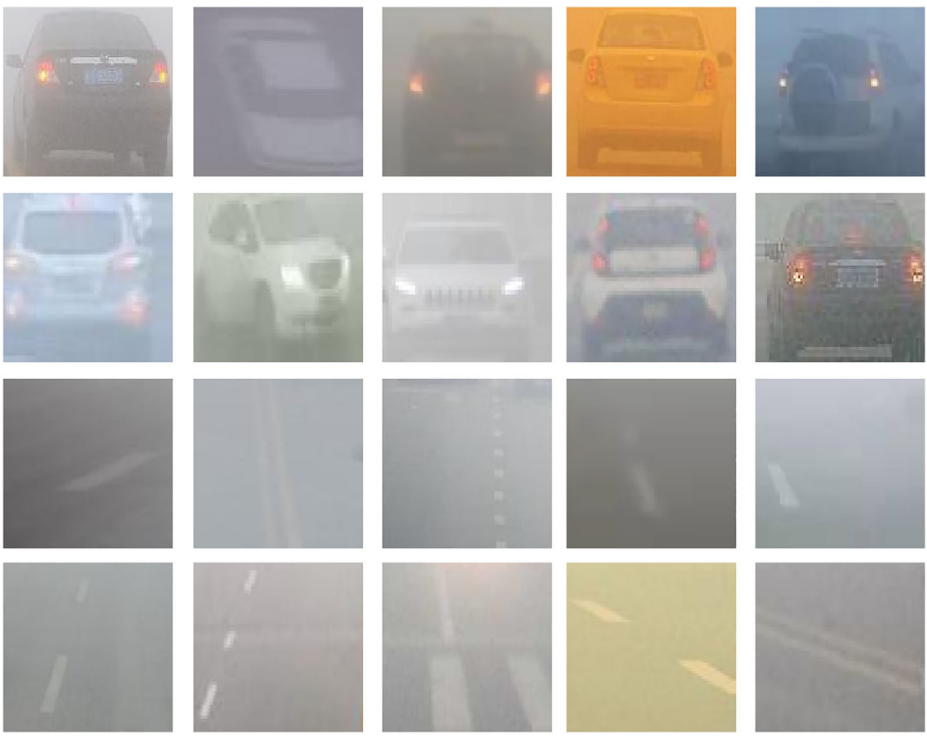

Fig. 4 Sample Images of the proposed Hazy Parking system Dataset: Top two rows - Hazy busy parking patches, Bottom two row - Hazy unoccupied parking patches

trained for 30 epochs. Further in [7], mAlexnet's performance was proven to be comparable to that of Alexnet [33] on CNRPark [8] and CNRPark-EXT [7] datasets while being three times smaller in size. In [7], all the models were trained using gradient descent with momentum for 6 epochs, having an initial learning rate of 0.0008, which was multiplied by 0.75 after every 2 epochs. The batch size was set to 64, momentum to 0.9, and weight decay to 0.0005.

In this work, the proposed system uses the mAlexnet architecture after doing certain modifications to it (shown in Table 1). Training and testing images are shuffled and resized to 224×224 pixels. The modified mAlexnet architecture is trained on Caffe framework using Adam optimization algorithm[5] with the following hyperparameters: $\beta_1 = 0.9$, $\beta_2 = 0.999$, and weight decay=5×10^{-4}. The learning rate is fixed to 0.001 throughout the training process. The training procedure followed in this work is shown in Table 5.

The proposed trained model is compared with the existing mAlexnet model of [8] and [7] using CNRPark, CNRPark-EXT and PKLot datasets (shown in Tables 6 and 7). For comparison, we trained the models on various subsets of CNRPark, CNRPark-EXT, PKLot datasets and are evaluated on non-overlapping test sets. Models shown in Tables 6 and 7 are trained for 10 epochs and 5 epochs respectively. All the models were trained on Ubuntu 18.04 virtual environment with 16 GB RAM running on Intel(R) Core(TM) i5-7300HQ CPU @ 2.50GHz in CPU mode.

It can be observed from Tables 6 and 7 that the proposed model using modified mAlexnet architecture achieves better results compared to the existing Alexnet and mAlexnet pre-trained models of [8] and [7] on subsets of CNRPark, CNRPark-EXT, and PKLot datasets. This can be attributed to two factors: Use of batch normalization instead of local response normalization and use of Adam optimization over gradient descent with momentum for training.

Figure 5 shows the confusion matrix of the proposed mAlexnet and mAlexnet [7] trained on CNRPark-EXT Train set. As can be observed from the figure, the proposed model drastically reduces the number of false positives while slightly improving the performance of false negatives. To gain a further understanding of the improvement in performance brought about by modified mAlexnet, we visualize the false positives of mAlexnet [7] which are correctly classified by the proposed model in Fig. 6. From the figure, it can be inferred that the proposed model is more robust to the presence of water or parking meters at the parking spots. Further, sample false negatives of both the models are shown in Fig. 7. It can be inferred that the misclassification is due to the presence of haze and the partial presence of cars in the image. We conjecture this as one of the main reasons for modified mAlexnet achieving a similar performance as that of mAlexnet [7] in terms of false negatives.

5.3 Dehazing network

5.3.1 Model description

For the dehazing network, the proposed model followed the AOD-Net [35] architecture. Instead of evaluating the dehazing networks on performance metrics such as peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM), they are evaluated by their utility before the parking space classification takes place on hazy and non-hazy datasets. By using dehazing network before the classification process starts will improves the performance of

[5] https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/

Table 5 Training procedure for modified mAlexnet

Optimization Algorithm	Hyperparameters	Initial Learning rate	Batch size	Epochs
Adam	Weight decay: 5×10^{-4} $\beta1$=0.9, $\beta2$=0.999	0.001	64	5-10

Table 6 Comparison of accuracy of proposed model with pretrained models of [8] (bold indicates maximum)

Train	Test	Network	Base learning rate	Accuracy
CNRPark A	CNRPark B	modified mAlexnet	0.001	**87.56%**
		mAlexnet [8]	0.001	86.30%
CNRPark B	CNRPark A	modified mAlexnet	0.001	**90.78%**
		mAlexnet [8]	0.001	90.70%

Table 7 Accuracy comparison of proposed model with the pretrained models of [7] (bold indicates maximum)

Train set	Test set	Network	Accuracy
CNRPark All	CNRPark-EXT Test	modified mAlexnet	**93.92%**
		mAlexnet [7]	93.70%
		Alexnet [7]	93.38%
CNRPark+EXT TRAIN C1-C8	CNRPark-EXT Test	modified mAlexnet	**97.85%**
		mAlexnet [7]	96.44%
		Alexnet [7]	96.74%
CNRPark+EXT TRAIN	CNRPark-EXT Test	modified mAlexnet	**98.51%**
		mAlexnet [7]	97.78%
		Alexnet [7]	97.98%
PKLot Train	PKLot Test	modified mAlexnet	**98.89%**
		mAlexnet [7]	97.96%
		Alexnet [7]	98.76%

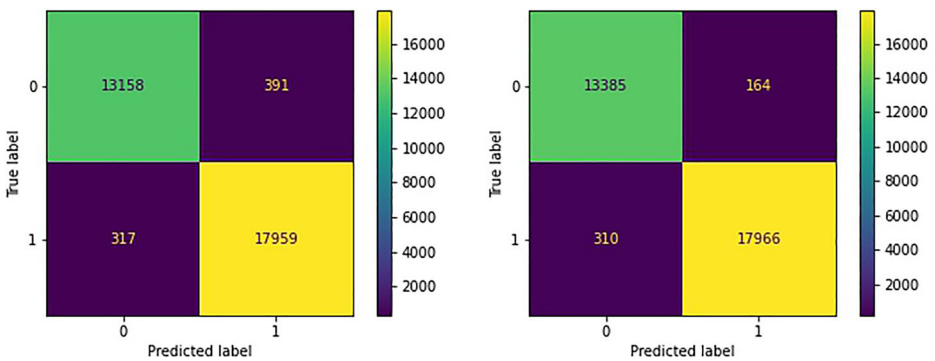

Fig. 5 Confusion matrix for mAlexnet [7] (left) and modified mAlexnet (right) trained on CNRPark-EXT Train and tested on CNRPark-EXT Test set

Fig. 6 Sample false positives mAlexnet [7] which are correctly classified by the posposed model. Both models are trained on CNRPark-EXT Train set and tested on CNRPark-EXT test set

the system on hazy images, but degrades the performance on non-hazy images. Alternative training methodologies for dehazing networks are explored to improve the performance of the system on non-hazy images. For that purpose, we explore four different models, and the details are presented below:

1. *Model 1*: AOD-Net trained on hazy images of OTS dataset followed by parking space classifier.
2. *Model 2*: AOD-Net trained on both hazy and clear (non-hazy) images of OTS dataset followed by parking space classifier.
3. *Model 3* : AOD-Net trained on both hazy and clear (non-hazy) images of OTS dataset using a modified loss function L_1 followed by parking space classifier. The modified loss function L_1 is given by:

$$L_1 = y\lambda L + (1 - y)(1 - \lambda)L \tag{6}$$

Where L denotes the Mean Square Error (MSE) loss function between the output of the network and the ground truth of the image corresponding to the input. $y = 1$, when the input is a clear (non-hazy) image, and $y = 0$, when input is a hazy image. λ is a hyper-parameter whose value ranges from 0 to 1. λ can be used to tune the relative performance of the network on hazy or non-hazy images. For $0 \leq \lambda < 0.5$, the network is trained to perform better on hazy images. For $0.5 < \lambda \leq 1$, the network is trained to perform better on non-hazy images. This gives more flexibility than other three models to train the system according to specific requirements. For experiments, four values of λ are used to train four networks, each randomly chosen from the following four intervals: [0, 0.24], [0.25, 0.49], [0.5, 0.74], [0.75, 1].

We then compare the performance of the above models with modified mAlexnet, mAlexnet [8], and Alexnet [7] trained on CNRPark-EXT train set.

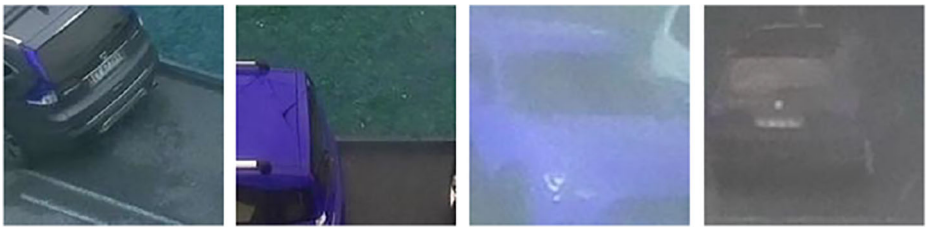

Fig. 7 Sample false negatives of modified mAlexnet and mAlexnet [7] trained on CNRPark-EXT Train set and tested on CNRPark-EXT test set

The combined dehazing network and classification process necessitates the output of AOD-Net, and input of the modified mAlexnet classifier to have the same dimensions. The classifier developed in Section 5.2 takes the inputs of $224 \times 224 \times 3$ images. Further, as the input and output of AOD-Net have same dimensions, the training and testing images are cropped to 224×224.

5.3.2 Model training

AOD-Net is trained using gradient descent with momentum having the following hyperparameters: momentum = 0.9 and weight decay = 0.0001. For non-hazy images, Mean Square Error (MSE) between the output of the network and the ground truth of the image is used as the loss function. For clear (non-hazy) images, the output of the dehazing network is compared with the input for loss calculation. The number of clear (non-hazy) images used during the training phase is kept low compared to that of hazy images to prevent the network from learning an identity function between the input and output. The gradients are constrained within $[-0.1, 0.1]$ to stabilise the training. The model is trained using Caffe framework with a fixed learning rate of 0.001. AOD-Net takes around 4-5 epochs to converge on the OTS dataset. Hence, all the networks are trained for 5 epochs. The same procedure is adopted for all the models except Model 3, where the loss function is altered as given in (6). The dehazing network is trained on Nvidia K80 GPU provided as part of p2.xlarge EC2 instance by Amazon Web Services (AWS) using Jetware's 'Caffe Python 2.7 Nvidia GPU production on Ubuntu' AMI. The training procedure followed is summarized in Table 8.

It is found that joint tuning of pre-trained AOD-Net and modified mAlexnet as a unified pipeline improves the classification performance of the system under hazy conditions. All the trained AOD-Net models are concatenated with modified mAlexnet and are jointly tuned as a single network on hazy parking system training set. The training is performed on Caffe framework using gradient descent with momentum having the following hyperparameters: momentum = 0.9 and weight decay = 0.0001. The model is trained with a fixed learning rate of 0.0001 for 5 epochs, and the gradients are constrained between $[-0.1, 0.1]$ to stabilise the training.

5.3.3 Discussion on empirical results

All the models are evaluated using hazy parking system test set (Table 4) and CNRPark-EXT test set (Table 3) to test the model's performance under both hazy and non-hazy conditions, respectively. The accuracies of the four models on these datasets are displayed in Table 9. The proposed modified mAlexnet model performs extremely well under non-hazy conditions achieving 98.51% accuracy on CNRPark-EXT test set. However, it performs poorly under hazy conditions achieving 74.41% accuracy on hazy parking system test set. Similarly mAlexnet [8] and Alexnet [7] achieves 97.76% and 97.98% on CNRPark-EXT test set while achieving 76.27% and 80.37% on hazy parking test set respectively. In comparison, the proposed AOD-Net+mAlexnet model achieved significant gains of over 10-15% accuracy on hazy parking system test set with a marginal decrease in accuracy of 1-3% on CNRPark-EXT test set.

Model 2 achieves higher accuracy than Model 1 on CNRPark-EXT test set, while achieving lower accuracy on hazy parking test set. However, after joint optimization, Model 1 achieves higher accuracy than Model 2 on hazy parking system test set. Both these observations can be attributed to the fact that AOD-Net in Model 2 was trained on both hazy

Table 8 Training procedure for AOD-Net and joint optimization of AOD-Net and modified mAlexnet pipeline

Network	Optimization Algorithm	Hyperparameters	Initial Learning rate	Batch size	Epochs
AOD-Net	Gradient Descent with momentum	Weight decay: 1×10^{-4} Momentum : 0.9	0.001	64	5
AOD-Net + modified mAlexnet	Gradient Descent with momentum	Weight decay: 1×10^{-4} Momentum : 0.9	0.001	64	10

Table 9 Comparison of model accuracies on Hazy parking system Test set and CNRPark-EXT Test set

Net	Accuracy on Hazy parking system Test	Accuracy on CNRPark-EXT Test
Model 1	86.71%	97.25%
Model 1 after joint optimization	88.39%	96.68%
Model 2	86.40%	97.28%
Model 2 after joint optimization	86.89%	96.72%
Model 3 (λ=0.1916)	86.27%	97.34%;
Model 3 (λ=0.1916) after joint optimization	88.07%	96.88%
Model 3 (λ=0.4291)	86.89%	96.61%
Model 3 (λ=0.4291) after joint optimization	88.63%	96.42%
Model 3 (λ=0.6555)	86.96%	97.38%
Model 3 (λ=0.6555) after joint optimization	86.27%	96.33%
Model 3 (λ=0.8984)	84.78%	97.57%
Model 3 (λ=0.8984) after joint optimization	86.71%	97.08%
Modified mAlexnet	74.41%	98.51%
mAlexnet [8]	76.27%	97.78%
Alexnet [7]	80.37%	97.98%

and clear (non-hazy) images, thereby enhancing system's performance under non-hazy conditions while degrading system's performance under hazy conditions.

In the case of Model 3, λ is a hyperparameter that can be used to tune the relative performance of the network on hazy and non-hazy images. In accordance with (6), it is found that models trained with $\lambda = 0.1916$ and $\lambda = 0.4291$ achieve lower accuracy than models trained with $\lambda = 0.6555$ and $\lambda = 0.8984$ on CNRPark-EXT test set. Models trained with $\lambda = 0.1916$ and $\lambda = 0.4291$ achieve higher accuracy than the model trained with $\lambda = 0.8984$ on hazy parking test set and slightly lower accuracy than model trained with $\lambda = 0.6555$. Further, model 3 trained with $\lambda = 0.6555$ outperforms Model 1 and Model 2 under both hazy and non-hazy conditions.

Joint optimization is found to improve the system's performance on hazy images in almost all of the cases, while it is found to degrade the system's accuracy under non-hazy conditions in all the cases. However, in some cases (Model 1, Model 3 $\lambda = 0.1916, 0.4291, 0.8984$), joint optimization is found to be beneficial as it is found to increase the accuracy on hazy parking test set by 1.5% with around 0.5% decrease in accuracy on CNRPark-EXT test set. Model 3 ($\lambda = 0.4291$) after joint optimization achieves the highest accuracy of 88.63% on hazy parking system test set. Apart from Model 4, Model 3 ($\lambda = 0.8984$) achieves the highest accuracy of 97.57% on CNRPark-EXT test set. This can be attributed to λ value used for training the dehazing network being closer to 1. Model 3 with $\lambda = 0.8984$ achieves around 10% gain on hazy parking system test set with only 1% decrease in accuracy on CNRPark-EXT test set. Model 3 with $\lambda = 0.6555$ after joint optimization achieves over 12% accuracy gain on hazy parking test set and is particularly useful when the target application requires the system to be robust against the presence of haze in the image.

5.4 Runtime analysis

We compare the average runtimes of AOD-Net [35] with other state-of-the-art dehazing networks such as MSCNN [51] Dehazenet [12], GridDehazenet [39] and GFN [52]. Table 10 [39] shows the average runtimes of these models. AOD-Net is atleast 2.5 times faster than the other state-of-the-art dehazing networks.

Table 11 compares the average runtime of the proposed AOD-Net and mAlexnet pipeline with modified mAlexnet, Alexnet [7],mAlexnet [8]. We randomly select 1000 random 224×224 images from each CNRPark-EXT test set and PKLot test set and 200 random 224×224 images from hazy parking system test set to run for each model on the same machine without GPU acceleration. The model runtime is tested using Caffe framework.

Modified mAlexnet and the existing mAlexnet with its compact nature are 20 times faster than Alexnet while achieving similar accuracies on parking space classification, making it an ideal choice for real-time applications. With the use of dehazing network, the average runtime of the proposed architecture is significantly higher than mAlexnet architecture. However, the proposed pipeline of AOD-Net and mAlexnet in series is at least 2

Table 10 Average runtime of state-of-the-art dehazing networks [39]

Network	AOD-Net [35]	Grid-DehazeNet [39]	MSCNN [51]	Dehazenet [12]	GFN [52]
Average runtime (s)	0.08	0.22	0.26	0.30	0.37

Table 11 Average runtime of proposed system(AOD-Net and modified mAlexnet in series), modified mAlexnet , mAlexnte [8] and Alexnet [7] architectures

Network	Average runtime on CNRPark-EXT Test (s)	Average runtime on PKLot Test (s)	Average runtime on Hazy parking system Test (s)	Mean Average runtime (s)
mAlexnet	0.009	0.009	0.009	0.009
modified mAlexnet	0.009	0.009	0.009	0.009
Alexnet	0.181	0.182	0.180	0.181
AOD-Net + modified mAlexnet	0.091	0.090	0.093	0.0907

times faster than Alexnet and other state-of-the-art dehazing networks. The runtime analysis is performed on Ubuntu 18.04 virtual environment with 16 GB RAM running on Intel(R) Core(TM) i5-7300HQ CPU @ 2.50GHz.

6 Conclusion

This paper proposed a parking space occupancy detection model that is robust to hazy conditions. The system pipeline consists of two parts: A dehazing network followed by the parking space classification technique using CNN. The system is evaluated on both hazy and non-hazy images. Modified mAlexnet architecture is trained to classify parking space occupancy and is compared with other state-of-the-art architectures. Using Adam optimization and batch normalization in place of local response normalization, the modified mAlexNet achieved higher classification accuracy on CNRPark and CNRPark-EXT datasets.

For the dehazing network, we trained AOD-Net, an end-to-end dehazing network on OTS dataset, and evaluated its performance by its utility before the classification. We studied the use of clear (non-hazy) images for training the dehazing network and joint optimization of dehazing network. Further, the trade-off between model's performance on hazy and non-hazy images is addressed, and a new loss function L_1 is introduced. We incorporated a hyperparameter λ, which can be used to tune the relative performance of the model on hazy and non-hazy images. This enables us to tune the proposed system according to target requirements.To validate our approaches on hazy images, we have created a new hazy parking system dataset consisting of hazy and unoccupied parking patches extracted from RTTS subset of Reside-β dataset. CNRPark-EXT is used to evaluate the model's performance on non-hazy images.

Experimental results show that the proposed system pipeline achieves significant gains (about 10-15%) over the parking space classification system on hazy images. Experimental results show that the inclusion of non-hazy images during training of AOD-Net does improve the system's performance on non-hazy images but degrades the system's performance on hazy images. Further, joint optimization of the system on hazy images increases the performance under hazy conditions. However, the increase in performance under hazy conditions does not outweigh the decrease in performance under non-hazy conditions. The proposed loss function addresses this trade-off by tuning the hyperparameter λ. Specifically,

training AOD-Net with λ values close to one improves the system's performance under non-hazy conditions, while training with λ value close to zero improves the performance under hazy conditions. However, with the use of dehazing network, the average runtime of the proposed pipeline is significantly greater than that of mAlexnet architecture.

This work can be extended by using quantization techniques, which will reduce the runtime of the proposed system. Further, the parameter λ used in the training of dehazing networks can be extensively tuned for better performance.

Data Availability The dataset generated during and/or analysed during the current study is available from the corresponding author on reasonable request.

Declarations

Conflict of Interest No potential conflict of interest was reported by the authors.

References

1. Ajchariyavanich C, Limpisthira T, Chanjarasvichai N, Jareonwatanan T, Phongphanpanya W, Wareechuensuk S, Srichareonkul S, Tachatanitanont S, Ratanamahatana C, Prompoon N, Pipattanasomporn M (2019) Park king: an iot-based smart parking system. In: 2019 IEEE International Smart Cities Conference (ISC2), pp 729–734. https://doi.org/10.1109/ISC246665.2019.9071721
2. Akhtar ZUA, Wang H (2020) Wifi-based driver's activity recognition using multi-layer classification. Neurocomputing 405:12–25
3. Al-Turjman F, Malekloo A (2019) Smart parking in iot-enabled cities: a survey. Sustain Cities Soc 49:101608
4. Ali G, Ali T, Irfan M, Draz U, Sohail M, Glowacz A, Sulowicz M, Mielnik R, Faheem ZB, Martis C (2020) Iot based smart parking system using deep long short memory network. Electronics, 9
5. Almeida P, Oliveira LS, Silva E, Britto A, Koerich A (2013) Parking space detection using textural descriptors. In: 2013 IEEE International conference on systems, man, and cybernetics, IEEE, pp 3603–3608
6. Alshdadi AA (2021) Cyber-physical system with iot-based smart vehicles. Soft Comput, 1–13
7. Amato G, Carrara F, Falchi F, Gennaro C, Meghini C, Vairo C (2017) Deep learning for decentralized parking lot occupancy detection. Expert Syst Appl 72:327–334
8. Amato G, Carrara F, Falchi F, Gennaro C, Vairo C (2016) Car parking occupancy detection using smart camera networks and deep learning. In: 2016 IEEE Symposium on Computers and Communication (ISCC), IEEE, pp 1212–1217
9. Anagnostopoulos T, Fedchenkov P, Tsotsolas N, Ntalianis K, Zaslavsky A, Salmon I (2020) Distributed modeling of smart parking system using lstm with stochastic periodic predictions. Neural Comput Appl 32:10783–10796
10. Barriga JJ, Sulca J, León JL, Ulloa A, Portero D, Andrade R, Yoo SG (2019) Smart parking: a literature review from the technological perspective. Appl Sci 9:4569
11. Bura H, Lin N, Kumar N, Malekar S, Nagaraj S, Liu K (2018) An edge based smart parking solution using camera networks and deep learning. In: 2018 IEEE International Conference on Cognitive Computing (ICCC), IEEE, pp 17–24
12. Cai B, Xu X, Jia K, Qing C, Tao D (2016) Dehazenet: an end-to-end system for single image haze removal. IEEE Trans Image Process 25:5187–5198
13. Chen C, Liu B, Wan S, Qiao P, Pei Q (2021) An edge traffic flow detection scheme based on deep learning in an intelligent transportation system. IEEE Trans Intell Transp Syst 22:1840–1852
14. Chippalkatti P, Kadam G, Ichake V (2018) I-spark: Iot based smart parking system. In: 2018 International Conference On Advances in Communication and Computing Technology (ICACCT), pp 473–477. https://doi.org/10.1109/ICACCT.2018.8529541
15. Choi J, Min K, Lee Y (2014) An intelligent parking platform of neighborhood ev for autonomous mobility service. Multimedia Tools and Applications, 74
16. Cueva-Fernandez G, Espada JP, García-Díaz V, Gonzalez-Crespo R (2015) Fuzzy decision method to improve the information exchange in a vehicle sensor tracking system. Appl Soft Comput 35:708–716

17. De Almeida PauloRL, Oliveira LS, Britto AS Jr, Silva EJ Jr, Koerich AL (2015) Pklot–a robust dataset for parking lot classification. Expert Syst Appl 42:4937–4949
18. Dhillon A, Verma GK (2020) Convolutional neural network: a review of models, methodologies and applications to object detection. Progress Artif Intell 9:85–112
19. Diaz Ogás MG, Fabregat R, Aciar S (2020) Survey of smart parking systems. Appl Sci 10:3872
20. Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 580–587
21. Giuffrè T, Siniscalchi SM, Tesoriere G (2012) A novel architecture of parking management for smart cities. Procedia Soc Behav Sci 53:16–28
22. Gkolias K, Vlahogianni EI (2018) Convolutional neural networks for on-street parking space detection in urban networks. IEEE Trans Intell Transp Syst 20:4318–4327
23. González-Lozoya SM, de la Calleja J, Pellegrin L, Escalante HJ, Medina MA, Benitez-Ruiz A et al (2020) Recognition of facial expressions based on cnn features. Multimed Tools Applic 79:13987–14007
24. Guo Y, Liu Y, Bakker EM, Guo Y, Lew MS (2018) Cnn-rnn: a large-scale hierarchical image classification framework. Multimed Tools Applic 77:10251–10271
25. He K, Sun J, Tang X (2010) Single image haze removal using dark channel prior. IEEE Trans Pattern Anal Mach Intell 33:2341–2353
26. Huang F, Qi X, Li C, Hu W (2020) Aerial image classification by learning quality-aware spatial pyramid model. Futur Gener Comput Syst 111:271–277
27. Idris MYI, Tamil EM, Noor NM, Razak Z, Fong KW (2009) Parking guidance system utilizing wireless sensor network and ultrasonic sensor. Inf Technol J 8:138–146
28. Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T (2014) Caffe: Convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM international conference on multimedia, pp 675–678
29. Khalid M, Wang K, Aslam N, Cao Y, Ahmad N, Khan MK (2020) From smart parking towards autonomous valet parking: a survey, challenges and future works. J Netw Comput Appl, 102935
30. Khanna A, Anand R (2016) Iot based smart parking system. In: 2016 International Conference on Internet of Things and Applications (IOTA), IEEE, pp 266–270
31. Kianpisheh A, Mustaffa N, Limtrairut P, Keikhosrokiani P (2012) Smart parking system (sps) architecture using ultrasonic detector. Int J Softw Eng Applic 6:55–58
32. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25
33. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inform Process Syst 25:1097–1105
34. LeCun Y et al (2015) Lenet-5, convolutional neural networks. http://yann.lecun.com/exdb/lenet 20: 14
35. Li B, Peng X, Wang Z, Xu J, Feng D (2017) Aod-net: all-in-one dehazing network. In: Proceedings of the IEEE international conference on computer vision, pp 4770–4778
36. 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 28:492–505
37. Lin T, Rivano H, Le Mouël F (2017) A survey of smart parking solutions. IEEE Trans Intell Transp Syst 18:3229–3253
38. Liu F, Shen C, Lin G, Reid I (2015) Learning depth from single monocular images using deep convolutional neural fields. IEEE Trans Pattern Anal Mach Intell 38:2024–2039
39. Liu X, Ma Y, Shi Z, Chen J (2019) Griddehazenet: attention-based multi-scale network for image dehazing. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 7314–7323
40. Liu J, Liu Z, Zhang H, Yuan H, Manokaran KB, Maheshwari M (2021) Multi-sensor information fusion for iot in automated guided vehicle in smart city. Soft Comput, 1–13
41. Mago N, Kumar S, Goyal LM (2021) Real time fuzzy based intelligent parking detection system using deep learning techniques. Int J Fuzzy Syst, 1–9
42. McCartney EJ (1976) Optics of the atmosphere: scattering by molecules and particles. New York
43. Màrmol Soley E, Sevillano X (2016) Quickspot: a video analytics solution for on-street vacant parking spot detection. Multimed Tools Applic 75:17711–17743
44. Narasimhan SG, Nayar SK (2000) Chromatic framework for vision in bad weather. In: Proceedings IEEE conference on computer vision and pattern recognition. CVPR 2000 (Cat. No. PR00662), vol 1, IEEE, pp 598–605
45. Narasimhan SG, Nayar SK (2002) Vision and the atmosphere. Int J Comput Vis 48:233–254
46. Naufal A, Fatichah C, Suciati N (2020) Preprocessed mask rcnn for parking space detection in smart parking systems. Int J Intell Eng Syst 13:255–65

47. Nurullayev S, Lee S-W (2019) Generalized parking occupancy analysis based on dilated convolutional neural network. Sensors 19:277
48. Nyambal J, Klein R (2017) Automated parking space detection using convolutional neural networks. In: 2017 Pattern Recognition Association of South Africa and Robotics and Mechatronics (PRASA-RobMech), IEEE, pp 1–6
49. Rahman M, Islam M, Sassi R, Aktaruzzaman M et al (2019) Convolutional neural networks performance comparison for handwritten bengali numerals recognition. SN Appl Sci 1:1–11
50. Ramasamy M, Solanki SG, Natarajan E, Keat TM (2018) Iot based smart parking system for large parking lot. In: 2018 IEEE 4th International Symposium in Robotics and Manufacturing Automation (ROMA), pp 1–4
51. Ren W, Liu S, Zhang H, Pan J, Cao X, Yang M-H (2016) Single image dehazing via multi-scale convolutional neural networks. In: European conference on computer vision, Springer, pp 154–169
52. Ren W, Ma L, Zhang J, Pan J, Cao X, Liu W, Yang M-H (2018) Gated fusion network for single image dehazing. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 3253–3261
53. Ren Z, Lai J, Wu Z, Xie S (2021) Deep neural networks-based real-time optimal navigation for an automatic guided vehicle with static and dynamic obstacles. Neurocomputing 443:329–344
54. Saharan S, Kumar N, Bawa S (2020) An efficient smart parking pricing system for smart city environment: a machine-learning based approach. Futur Gener Comput Syst 106:622–640
55. Šćekić Z, Čakić S, Popović T, Jakovljević A (2022) Image-based parking occupancy detection using deep learning and faster r-cnn. In: 2022 26th International Conference on Information Technology (IT), IEEE, pp 1–5
56. Shoup DC (2006) Cruising for parking. Transp Polic 13:479–486
57. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556
58. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA (2017) Inception-v4, inception-resnet and the impact of residual connections on learning. In: Thirty-first AAAI conference on artificial intelligence
59. Targ S, Almeida D, Lyman K (2016) Resnet in resnet: generalizing residual architectures. arXiv:1603.08029
60. True N (2007) Vacant parking space detection in static images. https://cseweb.ucsd.edu/classes/wi07/cse190-a/reports/ntrue.pdf
61. Wang Z, Ma Y (2021) Detection and recognition of stationary vehicles and seat belts in intelligent internet of things traffic management system. Neural Comput Appl, 1–10
62. Wang Z, Deng Z, Wang S (2016) Accelerating convolutional neural networks with dominant convolutional kernel and knowledge pre-regression. In: Leibe B, Matas J, Sebe N, Welling M (eds) Computer Vision – ECCV 2016. Springer International Publishing, Cham, pp 533–548
63. Wu H, Pang GK-H, Choy KL, Lam HY (2018) Dynamic resource allocation for parking lot electric vehicle recharging using heuristic fuzzy particle swarm optimization algorithm. Appl Soft Comput 71:538–552
64. Xiang X, Lv N, Zhai M, El Saddik A (2017) Real-time parking occupancy detection for gas stations based on haar-adaboosting and cnn. IEEE Sens J 17:6360–6367
65. Xu Y, Wei M (2021) Multi-view clustering toward aerial images by combining spectral analysis and local refinement. Futur Gener Comput Syst 117:138–144
66. Zhu Q, Mai J, Shao L (2015) A fast single image haze removal algorithm using color attenuation prior. IEEE Trans Image Process 24:3522–3533
67. Zhu L, Yu FR, Wang Y, Ning B, Tang T (2018) Big data analytics in intelligent transportation systems: a survey. IEEE Trans Intell Transp Syst 20:383–398

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.