TOPICAL REVIEW

A Comprehensive Review of Vehicle Detection Techniques Under Varying Moving Cast Shadow Conditions Using Computer Vision and Deep Learning

MUHAMMAD UMAIR ARIF[ID]**[1], MUHAMMAD UMAR FAROOQ**[ID]**[1], RANA HAMMAD RAZA**[ID]**[1],
ZAIN UL ABIDEEN LODHI**[2]**, (Student Member, IEEE),
AND MUHAMMAD ABDUR REHMAN HASHMI**[ID]**[3], (Graduate Student Member, IEEE)**

[1]Department of Electronics and Power Engineering, Pakistan Navy Engineering College (PNEC), National University of Sciences and Technology (NUST), Karachi 75350, Pakistan
[2]Harbin Institute of Technology (HIT), Harbin 150001, China
[3]Department of Electronic and Electrical Engineering, University College London, WC1E 6BT London, U.K.

Corresponding author: Muhammad Umar Farooq (umar.farooq@pnec.nust.edu.pk)

ABSTRACT Design of a vision-based traffic analytic system for urban traffic video scenes has a great potential in context of Intelligent Transportation System (ITS). It offers useful traffic-related insights at much lower costs compared to their conventional sensor based counterparts. However, it remains a challenging problem till today due to the complexity factors such as camera hardware constraints, camera movement, object occlusion, object speed, object resolution, traffic flow density, and lighting conditions etc. ITS has many applications including and not just limited to queue estimation, speed detection and different anomalies detection etc. All of these applications are primarily dependent on sensing vehicle presence to form some basis for analysis. Moving cast shadows of vehicles is one of the major problems that affects the vehicle detection as it can cause detection and tracking inaccuracies. Therefore, it is exceedingly important to distinguish dynamic objects from their moving cast shadows for accurate vehicle detection and recognition. This paper provides an in-depth comparative analysis of different traffic paradigm-focused conventional and state-of-the-art shadow detection and removal algorithms. Till date, there has been only one survey which highlights the shadow removal methodologies particularly for traffic paradigm. In this paper, a total of 70 research papers containing results of urban traffic scenes have been shortlisted from the last three decades to give a comprehensive overview of the work done in this area. The study reveals that the preferable way to make a comparative evaluation is to use the existing Highway I, II, and III datasets which are frequently used for qualitative or quantitative analysis of shadow detection or removal algorithms. Furthermore, the paper not only provides cues to solve moving cast shadow problems, but also suggests that even after the advent of Convolutional Neural Networks (CNN)-based vehicle detection methods, the problems caused by moving cast shadows persists. Therefore, this paper proposes a hybrid approach which uses a combination of conventional and state-of-the-art techniques as a pre-processing step for shadow detection and removal before using CNN for vehicles detection. The results indicate a significant improvement in vehicle detection accuracies after using the proposed approach.

INDEX TERMS Computer vision, convolutional neural networks, deep learning, generative adversarial networks, intelligent transportation system, moving cast shadow removal, vehicle shadows, vehicle detection.

The associate editor coordinating the review of this manuscript and approving it for publication was Shaohua Wan.

I. INTRODUCTION

ITS is becoming increasingly popular because of its advantages in resolution of various traffic monitoring and management issues. Researches conducted under ITS help performing computer-assisted analysis of vehicular traffic for automated monitoring and control. The analyses are generally performed on video feeds data obtained through surveillance cameras already installed on different traffic areas. The data obtained from these cameras facilitate scene understanding in terms of vehicle detection, classification and re-identification [1], [2], [3] etc. This has many real-world applications such as Automatic Number Plate Recognition, queue estimation, speed detection and different anomalies detection including traffic jams and accidents detection etc. To extract accurate information related to vehicles, their successful detection in each frame is required. Some conventional techniques include moving object detection or background subtraction using a steady stream of images. However, as pointed out by Prati *et al.* [5], the presence of vehicles' shadows adds a level of complexity to these techniques. Though, shadows are easily distinguishable by a human eye, but their presence causes error in different detection and classification tasks. It makes them a challenging task because of following two properties:

1) The shadow pixels being significantly different from the background form part of the foreground like the actual vehicle pixels.
2) The shadows move at same speed and maintain same orientation with the associated vehicle.

Resultantly, simple image/video segmentation algorithms perform poorly on traffic video streams because these shadows are often confused as part of vehicles. Many times, the algorithms even merge and classify two vehicles in closer proximity as a single vehicle, as shown in Figure 1.

Similarly, at times, the vehicles are not even detected due to presence of high strength vehicle shadows, as shown in Figure 2. All of these reasons fueled the need to develop algorithms which can detect and remove shadows from video streams and enable accurate detection of vehicles.

Generally, there are two types of shadows i.e., self shadows and cast shadows. Self shadows occur on the object itself when it occludes light from a light source, while cast shadows occur on the ground or any other object nearby. Figure 3 shows difference between both with self shadows and cast shadows marked as red and blue borders, respectively. Due to similar nature of both types of shadows, it is often difficult for the algorithms to distinguish between both. However, this research work focuses only on the detection errors caused by moving cast shadow of vehicles.

There is a vast literature on related approaches that have been reported by the research community. Though, a detailed study yielded only one survey of 2001 by Prati *et al.* [4] which particularly focuses on the traffic paradigm. The other surveys [8], [9] having a broader perspective, reflect only a few shadow detection papers related to this paradigm. Taking
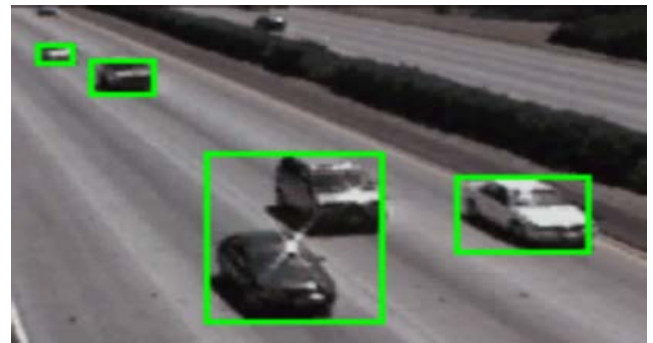
FIGURE 1. Example of two vehicles detected as one due to presence of vehicle shadows in closer proximity [4].

FIGURE 2. Detection of vehicle missed on the right due to presence of a large size, high strength vehicle cast shadow [6].

FIGURE 3. Difference between self and cast shadows [7].

it into consideration, there is a significant gap of more than a decade for a detailed survey which provides a comparative evaluation of cast shadow detection techniques in an urban traffic scenario. This paper provides a detailed analysis and comparison of different traffic paradigm-focused cast shadow detection algorithms reported since 2003. The reported methods have been categorized in line with the methodology adopted by Prati *et al.* to maintain succession [5].

The reported methods have been divided into two main categories of statistical and deterministic approaches. The statistical approaches primarily develop probabilistic models like Gaussian Mixture Model (GMM), Support Vector Machine (SVM) etc. Development of statistical models or classifiers requires supervised training. The classification accuracy of

statistical methods is directly proportional to the quantity and quality of the training dataset. While the quantitative part of the dataset is obvious, the qualitative part improves if the dataset contains traffic video stream recordings under varying conditions of weather effects (rain, clouds, lighting conditions, etc), perspective and vehicle types (varieties of heavy transport vehicles, light transport vehicles, bikes, etc.).

The statistical approaches are further divided into parametric and non-parametric ones. The parametric approaches may use any of the three types of parameters i.e., spatial, spectral and contextual or temporal. The spatial parameter refers to whether a single pixel value will be used to generate the feature vector, or some region/frame of multiple pixels will be used to acquire higher order statistical features like marginals in terms of means and central moments of rows etc. The spectral parameters refer to frequency information. With images, the spectral parameters are derived from the gradients of adjacent pixel values. Spectral parameters may also refer to whether the image is grayscale or Red Green Blue (RGB). If only a still image is used for feature vector generation, then no temporal parameters are exploited. The temporal parameters are derived from a series of still images like a traffic video stream. Functions like background subtraction primarily rely on temporal parameters. On the other hand, the non-parametric methods primarily use the pixel values to generate feature vectors. This apparently fine line between the parametric and non-parametric approaches will be thickened through discussions on the related works in the ensuing paragraphs.

The statistical approaches calculate the probability of a pixel or a set of pixels (an object) belonging to a particular class, e.g., foreground, background or shadow. On the other hand, deterministic methods carry out binary classification of pixels, as to whether or not they belong to a shadow. For binary classification, hard thresholding is done using some criterion. Deterministic approaches can be classified as model based or non-model based. The ones based on models, compare the features extracted from different regions of a scene with a specified model. The positive comparison result indicates presence of shadow and vice versa. However, such methods lack generality since it is difficult to foresee all possible scenarios. Moreover, increasing the number of models increases the algorithm complexity. Nevertheless, the non-model-based approaches use some relatively generalized criterion rather than strict models. The non-model methods may set criterion like ratio and luminance of pixel values in successive frames to decide their membership. Shadow pixels of the same vehicle usually maintain a somewhat constant ratio with the background. Also, they are darker and therefore have low luminance values compared to the background. Another method is to observe whether the object edges and corners change in successive frames. Such changes are inherent to shadows since the vehicle orientation with respect to light source causes changes in the shape of cast shadows.

Many methods normally detect shadow and remove it using the detected masks. To detect and remove shadows, traditional methods depend on hand-crafted features such as color, area, and user interaction. Finlayson *et al.* [10] used the L2-norm to generate color-invariant images and compared the changed image to the original image to set the shadow edges to zero. Guo *et al.* [11] used SVM-based area classifiers and graph cut to segment areas with similar characteristics, such as brightness and texture, and performed shadow area labeling. After that, using the planned illumination model, the shadow removal image was generated by reconstructing the brightness value of each pixel. Based on user inputs for shadow area, Gong and Cosker [12] obtained a fusion image with magnified shadow boundary. After detection, the shadow model was extracted based on the lighting value shift of the shadow boundary.

Other than the traditional methods, machine learning methods have shown distinctive performances for object detection and image classification tasks. Both conventional (such as Decision Trees, SVM, K-means and Naive Bayes etc.) and advanced deep learning architectures (such as VGG16 [13], ResNet101 [14] and R-CNN [15] etc.) are used for various related tasks. There are two main approaches to machine learning, which are supervised and unsupervised learning. The major difference between them is labeled data. The learning-based methods can be classified into supervised and unsupervised learning. In supervised learning, algorithm iteratively learns from the labeled data by making predictions and then automatically makes adjustments according to the ground truth. The model takes time to train, however, the results are usually accurate. On other hand, the unsupervised learning algorithm learns the underlying structure from the unlabeled data without any human intervention. The model usually takes comparatively less time to generate output, however, the results are not that accurate unless human supervision is involved in validation process.

For relative performance analysis of different algorithms, only the work evaluated using same or similar datasets is included in comparison. Some popular datasets include Highway I, II, and III datasets. Details of these datasets are provided in ensuing paragraphs.

The overall contributions of the paper are as followed:

1) An in-depth comparative analysis of conventional and state-of-the-art techniques reported from 2003 till date for moving cast shadow detection and removal with a focus on traffic paradigm for ITS applications.
2) Performance comparison for YOLOv5-based vehicle detection with and without removal of moving cast shadows using a Generative Adversarial Networks (GAN) based approach on a custom dataset.
3) Examination of a hybrid approach using a combination of conventional Computer Vision-based Gamma Correction and state-of-the-art GAN-based shadow removal technique to improve overall vehicle detection performance.

The rest of the paper is organized as follows. Section II discusses the major highlights of similar surveys previously conducted. Section III provides taxonomy that categorizes the shadow detection algorithms. Section IV provides details of the common datasets and evaluation metrics and further discusses strengths and weaknesses of all approaches provided in the taxonomy. Section V provides results and related discussion on proposed vehicle detection with shadow exclusion approach. Finally, Section VI concludes the study and highlights the future direction for this research.

II. RELATED WORKS

Shadow detection is equally important for indoor as well as outdoor scene analysis for segmentation, objection detection and recognition etc. However, the scope of this paper includes shadow detection methods for accurate vehicle detection only. Accordingly, the surveys discussed in this section include only those that have reviewed at-least some outdoor shadow detection methods applicable to traffic video streams.

One of the pioneer works is that of Prati *et al.* reported in 2003 [5], which presented the first comprehensive evaluation of different shadow detection approaches. The survey [5] selected 20 research papers from the four categories of algorithms, i.e., Statistical Parametric (SP), Statistical Non-Parametric (SNP), Deterministic Model (DM) based and Deterministic Non-Model (DNM) based. Two quantitative and seven qualitative evaluation metrics were used. The quantitative metrics included the True Positive Rates of shadow and object detection named as Shadow Detection Rate and Shadow Discrimination Rate defined by equation 1 and 2, respectively.

$$\eta = \frac{TP_S}{TP_S + FN_S} \qquad (1)$$

$$\zeta = \frac{TP_F}{TP_F + FN_F} \qquad (2)$$

where, True Positives and False Negatives are represented by TP and FN, respectively. Shadow is represented by *S* and Foreground is represented by *F*. The TP_F is obtained by subtracting the number of detected shadow points that correspond to a foreground object from the number of ground-truth points of the foreground object.

The qualitative metrics include robustness to noise, detection of indirect cast shadows and penumbra etc. Four different representative algorithms selected from each of the category were implemented and evaluated using a benchmark dataset of indoor and outdoor video sequences. Based upon the results, authors at [5] suggested that statistical methods performed better in indoor environment, whereas the deterministic ones performed optimally in outdoor environments. Coping with noise in traffic video streams requires some image/videos pre-processing like filtering etc. and post processing steps. The authors at [5] advised using non-parametric and non-model-based methods for more generalized applications involving large number of object classes and backgrounds. Whereas the parametric and model-based

methods are preferred for use in relatively controlled environments where more assumptions can be made.

Najdavi *et al.* [8] reviewed 37 works reported between 1998 to 2010 using a novel 4-layer taxonomy. The top two layers indicated whether the algorithm was object and/or environment dependent. The third layer described if the feature extraction during implementation used pixel values remaining in the spatial domain of the image or some transformation to other domains like Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT) etc. The bottom layer indicated whether the dataset contained mono-chrome or color images. The taxonomy of [5] gave information about the classifier algorithm and the feature extraction method in terms of the broad pseudo-code. However, the taxonomy of [8] provided details of possible application domains i.e., generalized, or controlled environment in terms of object and environment. Secondly, the only algorithm detail extractable from their taxonomy was whether a transform was employed during the feature extraction phase or not. Thirdly, it referered to the dataset color features. While [5] taxonomy was more algorithm oriented, [8] was rather application oriented. While differing in taxonomy, [8] evaluated some representative works using the quantitative evaluation metrics and test datasets used by [5] except Highway I, II. However, [8] discussed some additional qualitative metrics like performance in varying conditions of color space, illumination, texture of the foreground, static or dynamic scenes, geometric models and computational complexities in both hardware and software platforms. The authors at [8] suggested that the most desired shadow detection methods were the ones which could perform in more generalized conditions being independent of the object and environment. The overall recommendations based on [5], [8] are as follows:

- The methods involving transforms are computationally inexpensive and better suited for real time applications.
- Region based feature extraction from multiple pixels or transforms provides better noise robustness.
- Algorithms using shadow color models, texture models and geometric models perform well only under certain controlled conditions
- Transform methods generally provide better performance in applications requiring object and environment independency.

The last comprehensive survey on shadow detection techniques was published in 2012 by Andres *et al.* [9]. The work [9] reviewed shadow detection methods reported in almost the same time period as reviewed by [8]. However, unlike [8] that used an application-based taxonomy, [9] categorized the works based upon the type of features extracted for shadow detection. The rationale behind the taxonomy of Andres *et al.* [9] was that certain features are more effective for detecting shadows as compared to others. The possible features that can be used include intensity, chromacity, physical properties, geometry, textures and temporal features. Accordingly, 33 works reported between 2003 and 2010 were categorized into four classes. The four classes

corresponded to the relatively better performing features of chromacity, physical features, geometry and textures. The methods using textures were sub-divided into those using small and large regions. Some representative works from each class were then evaluated using a bigger dataset as compared to [5], [8]. The dataset included Campus, Hallway, Highway I, Highway III, Lab, Room and Caviar. The quantitative evaluation was done using the same metrics of Shadow Detection Rate and Discrimination Rate as used by [5] and [8]. The metrics used for qualitative evaluation were also similar to those of [5], [8] including, shadow independence, object independence, penumbra detection, robustness to noise, detection/discrimination trade-off and computational load. In terms of time efficiency, the chromacity based methods performed best, whereas the small region texture-based methods remained the least time efficient. The geometry and physical feature-based methods took around 10-15% additional computation time as compared to the chromacity methods, whereas the large region texture methods took almost twice the time. Generally, the large region texture-based methods showed highest shadow detection rate, whereas the geometry-based methods had least accuracies. Even based on qualitative metrics, the texture-based methods performed best while the geometry-based methods remained least performing. The chromacity and physical feature-based methods' performance remained higher than geometry methods but less than texture methods. Some discussion was also provided on the tracking performance of different methods. Based upon the application requirements in terms of time efficiency and accuracy, the taxonomy of [9] provided a useful guideline to select the most appropriate method.

Both surveys [8], [9] tried to cover the gap between the work of A. Prati et al. [5] till 2010. However, none used the same taxonomy but categorized the works of cast shadow detections using novel application and algorithm related categories. A useful contribution of this survey is that it maintains succession to the work of A. Prati et al. [5] by adopting the same taxonomy and reviewing the new significant works reported from 2003 till date. To the best knowledge of authors, no review covers the research reported after 2012. Moreover, while the previous surveys [5], [8], [9] evaluated shadow detection methods in indoor and outdoor environments, this work presents a more focused review for shadow detection in traffic video streams for ITS applications.

III. TAXONOMY OF CAST SHADOW DETECTION AND REMOVAL ALGORITHMS
This section discusses different cast shadow detection and removal algorithms organized in a taxonomy. The shadow detection techniques can be divided into classical and state-of-the-art algorithms as depicted in Figure 4. Sec III-A discusses the classical shadow detection and removal algorithms. While, Sec III-B discusses the state-of-the-art algorithms for detection and removal of shadows.

A total of 70 papers evaluated using datasets of urban traffic scenes have been shortlisted so as to give a
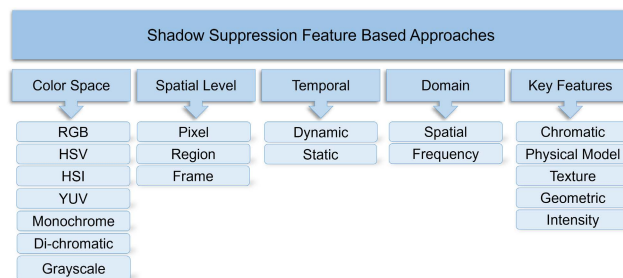

FIGURE 4. Taxonomy of traffic paradigm-focused cast shadow detection and removal algorithms.

comprehensive review of the work done in this area as of now. It was observed that despite increasing use of state-of-the-art deep neural networks for almost all vision tasks, there are limited publications that are focused on shadow detection techniques in context of urban traffic scenarios. There are few publications that are focused on shadow detection techniques based on Pulse Coupled Neural Networks (PCNN) [16], [17]. However, their results in context of traffic paradigm are not available, Therefore, a limited number of papers for these state-of-the-art techniques are discussed in this review paper. Nonetheless, these approaches are important and should be evaluated for urban traffic datasets in future.

Similarly, Chung et al. [18] provided the shadow detection scheme based on the successive thresholding on the hue over intensity ratio value to detect the shadow from Aerial images. This is basically a statistical non-parametric shadow removal approach for aerial images. This work proved to be more efficient compared to the earlier work of Tsai [19]. Authors modified the ratio map provided by [19] to make the ratio values of the shadow and the non-shadow pixels far apart by making use of both global and local thresholding. Possible shadow regions were grouped together by applying connected component process, thus, applying the thresholding technique in the iterative manner for extraction of true shadow pixel from combined shadow regions. Since the scope of this paper is limited to road side surveillance cameras, therefore, shadow detection on aerial images is not included in this review.

Furthermore, the taxonomy detailed in Tables 1, 2, 3, 4 has a horizontal level classification based on significant algorithmic differences while the vertical level classification entails features like color space, spatial level, temporal, domain and key features. This second level of classification, as shown in Figure 5 is based on factors which significantly affect the shadow detection results. The proposed second layer taxonomy derives inspiration from the work of Andres et al. [9], which correctly points out that the selection of appropriate features far greatly affects the results as compared to the employed algorithm. The traffic datasets on which the results of the papers have been reported are also included in appendices. Lack of standardization of datasets is evident from these tables. Similarly, a column for key technique used in

TABLE 1. Statistical (parametric) approach based shadow detection techniques.

Paper	Dataset	Color Space	Spatial	Temporal	Key feature	Domain	Technique Used
Friedman and Russell (1997) [21]	Road Watch traffic project	RGB	Pixel	Static	Intensity, Physical	Spatial Domain	Shadow pixel estimation using Incremental EM Algorithm
Mikic et al. (2000) [22]	Highway I	RGB	Pixel	Static	Chromatic	Spatial Domain	Shadow pixel estimation using Luminance based Turbo Segmentation
Bevilacqua (2003) [23]	Custom	Mono-chrome	Region	Static	Intensity	Spatial Domain	Multi-gradient operation on a division image to find shadow regions
Porikli and Thornton (2005) [24]	Custom	RGB	Pixel	Static	Intensity, Physical	Spatial Domain	Adaptive shadow flow method using Bayesian update technique
Joshi et al. (2007) [25]	Intelligent Room, Highway I, II, III,	RGB	Pixel	Static	Geometric	Spatial Domain	Blob-level reasoning using geometric constraints
Liu et al. (2007) [26]	Highway I, Highway II, Custom video	HSV	Pixel, Region	Dynamic	Physical	Spatial Domain	Pixel level GMM, region level Markov random fields and nearest neighbour tracking
Martel-Brisson and Zaccarin (2007) [27]	Highway III	YUV	Pixel	Dynamic	Physical	Spatial Domain	GMM based multi distribution statistical learning process
Pei and Wang (2009) [28]	Highway I, Intelligent Room, Laboratory	Mono-chrome	Region	Dynamic	Physical	Spatial Domain	GMM or generation of background image and PCA to extract features
Huang and Chen (2009) [29]	Highway I, Highway II, Hallway	RGB	Pixel	Dynamic	Chromatic	Spatial Domain	Confidence-rated GMM approach for shadow detection
Lin et al. (2010) [30]	Highway I and II, Custom	Mono-chrome	Pixel	Static	Chromatic	Spatial Domain	Shadow removal and extraction of foreground objects using GMM
Qi et al. (2014) [31]	Highway I, Intelligent Room, Laboratory etc.	RGB	Pixel	Dynamic	Intensity	Spatial Domain	Initial moving patches are extracted using GMM. Then, LBP is used to separate moving object pixels from initial moving pixels
Jiang et al. (2013) [32]	Campus, Hallway, Lab, Room, CAVIER	YUV	Pixel	Static	Chromatic, Texture	Spatial Domain	Texture based adaptive threshold estimator
Ouivirach and Dailey (2013) [33]	Hallway, Laboratory, Highway	HSV	Pixel	Dynamic	Chromatic	Spatial Domain	Maximum Likelihood based on color information
Russell and Zou (2013) [34]	Highway I, Intelligent Room, Lab and Campus	RGB	Region	Dynamic	Chromatic, Physical	Spatial Domain	Clustering approach to find similar patterns of spatial and temporal color constancy among pixels
Khare et al. (2014) [35]	Campus, Intelligent Room Video, Lab Video, Highway I	HSV	Pixels	Dynamic	Chromatic	Frequency Domain	HSV based Discrete Wavelet Transform (DWT)
Xiang et al. (2014) [36]	Highway I, Hallway, Intelligent Room, Campus	RGB	Region	Dynamic	Intensity	Spatial Domain	Local Intensity Ratio Model (LIRM) via GMM to deal with influences of illumination variations and shadows
Russell et al. (2015) [37]	Highway I, II and III	All	Frame	Static	Intensity	Spatial Domain	Illumination Direction Estimation and Shadow-Object Line Analysis
Valiere et al. (2015) [38]	Highway I, Custom	RGB	Pixel, Region	Static	Chromatic, Geometry	Spatial Domain	Vehicle Edges identification and removal with morphological operations
Dai et al. (2016) [39]	Highway I	RGB	Pixel	Static	Chromatic	Spatial Domain	GMM based separation with feature extractions using, chromacity, saliency, texture etc
Farou et al. (2017) [40]	Highway I	RGB, HSV, YUV	Pixel	Static	Chromatic	Spatial Domain	GMM based shadow suppression
Garg et al. (2019) [41]	Highway I, Lakeside I, PayalLebar I	RGB	Block	Static	Intensity	Spatial Domain	Predefined vehicle-sized blocks and interior edge features
Shi and Liu (2019) [42]	Highway I, Highway III, Custom	HSV	Region	Static	Chromatic	Spatial Domain	Combination of GMM, GFM and Bayes Classifier for foreground and background.
Sun et al. (2020) [43]	Custom	HSV	Region	Static	Color	Spatial Domain	Vehicle detection using optical flow and a shadow removal technique based on HSV
Zhang et al. (2021) [44]	Custom	Grayscale	Region	Static	Texture, Edges	Spatial Domain	Utilization of edge information, grayscale information and prior knowledge for shadow detection and removal.
Ghahremannezhad et al. (2021) [45]	Highway I, Highway III, Interstate-280	RGB, HSV	Pixel, Region	Static	Chromaticity, Illumination	Spatial Domain	Pixel, Region based and Statistical modeling approaches for shadows detection.
Zhou et al. (2021) [46]	Custom	HSV	Pixel, Region	Static	Chromaticity, Brightness similarity	Spatial Domain	Combination of color features and HLGP features used for shadow detection.
Shi (2021) [47]	Custom	RGB, YIQ, YCbCr	Pixel, Region	Static, Temporal	Intensity	Spatial Domain	Shadow region detected using Gaussian density.

each of these research papers has been added separately for better understanding of the classification methods.

A. CLASSICAL ALGORITHMS

A detailed highlight of classical shadow detection approaches based on statistical and deterministic techniques is given below.

1) STATISTICAL APPROACHES: PARAMETRIC

Statistical parametric approach assumes probabilistic distribution of sampled data based on parameters that are fixed. These algorithms provide best results when the assumptions are accurate. Therefore, selection of parameters is one of the major tasks in implementation of these algorithms. A total of 27 papers have been reviewed in this section related to

TABLE 2. Statistical (non-parametric) approach based shadow detection techniques.

Paper	Dataset	Color Space	Level	Temporal	Key feature	Domain	Technique Used
Rittscher et al. (2000) [48]	Custom	Monochrome	Pixel	Static	Intensity	Spatial Domain	Probabilistic approach based on HMM
Siala et al. (2004) [49]	CR, Highway I	RGB	Pixel	Static	Chromatic	Spatial Domain	Support Vector Domain Description (SVDD) algorithm to classify shadow pixels in the 3D color ratios space.
Wang et al. (2008) [50]	Highway I	RGB	Pixel	Dynamic	Intensity, Physical	Spatial Domain	HMM along with background subtraction. Maximum Likelihood Estimation is used to avoid local maximum.
Martel-Brisson and Zaccarin (2008) [51]	Highway I, Highway II	RGB	Pixel	Static	Physical	Spatial Domain	Unsupervised Kernel based approach and correlation between the model and spatial gradients
Joshi and Papanikolopoulos (2008) [52]	Highway I, III, IV, Intelligent Room	RGB	Pixel	Dynamic	Chromatic, Intensity, Physical	Spatial Domain	Detection and removal of shadows using Support Vector Machine (SVM) and co-training algorithm
Vargas et al. (2010) [53]	Highway I, Custom	Grayscale	Frame	Static	Chromatic	Spatial Domain	Quotient image with edge density estimation.
Amato et al. (2011) [54]	Highway I, Highway II	RGB, DHSV	Region	Static	Physical	Spatial Domain	Intensity variation background features in the shadows
Meher and Murty (2013) [55]	Custom	Monochrome	Region	Static	Intensity	Spatial Domain	Minimization of shadow regions search space using PCA
Wang and Zhang (2014) [56]	Custom	RGB	Region	Static	Haar-like features	Spatial Domain	Haar-like features with Adaboost used to train Haar Detector for detection of shadows
Gomes et al (2016) [57]	Highway 1, Highway 3	HSV	Pixel, Region	Static	Physical	Spatial Domain	A hypergraph segmentation is done containing chromatic and gradient information. Stochastic voting performed to classify.
Wang et al. (2017) [58]	Highway I, Highway II, Highway III	RGB	Region	Static	Intensity, Chromatic	Spatial Domain	Extraction of shadow regions using background subtraction, followed by PCA to reduce the algorithm complexity
Yang and Siu (2017) [59]	Custom	HSV	Image	Static	Chromatic	Spatial Domain	Learning and non-learning-based approaches for detection of shadow patches using HOG+SVM, 3-stage Cascade and Modified decision tree configurations
Yi et al. (2019) [60]	Highway I, Hallway, Laboratory, Campus etc	RGB, LRGB, HSV	Pixel, Region	Static	Geometry, Intensity, Chromatic	Spatial Domain	Separation of background (i.e., shadow) and foreground objects using ELM
Zhu and Yin (2020) [61]	Custom	LAB	Image	Static	Intensity	Spatial Domain	Shadow image trained with SVM which is then divided into shadow and light regions
Kan and Wang (2021) [62]	Highway I, Intelligent Room, Lab etc	RGB	Pixel, Region	Static	Intensity, Chromatic	Spatial Domain	Pixel and regional level features used to train S-ELM for classification of foreground and background pixels
Lu et al. (2022) [63]	Custom	N/A	Pixel	Static	Point cloud	Spatial Domain	Shadows detected and eliminated using a combination of Relief-F algorithm and Kernel Density Estimation.
Anandhalli et al. (2022) [64]	KITTI, Highway	RGB	Region	Static	Corner points	Spatial Domain	Corner points were used to accurately detect vehicles even with long shadows and illumination changes.
Sahoo and Nanda (2022) [65]	Highway I and Campus etc	RGB	Pixel, Region	Static, Temporal	LBP features, Gabor features	Spatial Domain	Background modeling using ST-KDE model was performed. LBP and Gabor features were combined to train the model. Residual shadows were classified by entropy map and threshold.

statistical parametric approach (See Table 1). Summary of these papers is given below:

Friedman and Russell [21] classified a pixel based on contributions from three distributions i.e., road, shadows, and vehicles. The pixel values were applied on a model with two settings i.e., intensity levels and RGB values. These mixture models were learned by incremental expectation maximization EM algorithm and then labelled. Each pixel

TABLE 3. Deterministic (model based) approach based shadow detection techniques.

Paper	Dataset	Color Space	Level	Temporal	Key feature	Domain	Technique Used
Koller et al. (1993) [66]	Test Videos	Mono-chrome	Pixel	Static	Geometry	Spatial Domain	Intraframe matching process with a state MAP-update step
Onoguchi (1998) [67]	N/A	Mono-chrome	Pixel	Static	Physical (Height)	Spatial Domain	Height information utilized for excluding shadow areas on road plane.
Yoneyama et al. (2003) [68]	N/A	Mono-chrome	Region	Dynamic	Geometric	Spatial Domain	Shadow cast region distinguished through a joint model of shadow and object
Salvador et al. (2004) [69]	MPEG-7 Highway Test Video	RGB	Pixel	Static	Chromatic	Spatial Domain	Invariant Feature points by Photometric and geometric properties of Shadows
Nadimi and Bhanu (2004) [70]	Custom	RGB, CiCb	Pixel	Dynamic	Physical models	Spatial Domain	Dichromatic reflection model, and reflectance-based analysis model and Spatiotemporal albedo ratio test
Sun and Li (2010) [71]	Sequence 1 of a road crossing, Sequence 2 of Jingzhu Highway	RGB, HSI	Pixel	Static	Chromatic	Spatial Domain	HIS and c1c2c3 Threshold Gaussian Shadow model
Chacon-Murguia and Gonzalez-Duarte (2012) [72]	PETS, CAVIAR	HSV	Pixel	Dynamic	Intensity	Spatial Domain	SOM based architecture to deal with dynamic backgrounds and elimination of shadows
Yang et al. (2018) [73]	Highway I, Room, Laboratory, Campus	RGB	Image	Static	Chromatic	Spatial Domain	NMF and BNMF based moving shadow detection
Hu and Liu (2021) [74]	Custom	RGB	Region	Static	Luminance, Chrominance, Texture	Spatial Domain	Multi-feature differences data was used to form a training sample which was then classified using GLVQ

TABLE 4. Deterministic (non-model based) approach based shadow detection techniques.

Paper	Dataset	Color Space	Level	Temporal	Key feature	Domain	Technique Used
Amamoto and Fujii (1999) [75]	N/A	Mono-chrome	N/A	Dynamic	N/A	Frequency Domain	Discrete Cosine Transform (DCT) used to extract the moving object while eliminating shadows
Stander et al. (1999) [76]	Table tennis, image	Mono-chrome	Region	Static	N/A	Spatial Domain	The shadow detection method is integrated into 2D shape estimation algorithm and the spatial adaptive relaxation scheme
Cucchiara et al. (2001) [77]	Highway I	HSV	Pixel	Static	Chromatic	Spatial Domain	Shadow suppression performed using HSV colorspace information
Cucchiara et al. (2003) [78]	Highway I	HSV	Pixel	Static	Chromatic	Spatial Domain	Object segmentation and shadow detection improved by using HSV colorspace information
Toth et al. (2004) [79]	N/A	LUV	Pixel	N/A	N/A	Spatial Domain	Mean Shift Algorithm for discriminating moving objects from their shadows
Leone and Distante (2007) [80]	Highway I	Color Space	N/A	Static	Texture	Frequency Domain	Automatic segmentation procedure based on adaptive background subtraction
Srividhya et al. (2022) [81]	Custom	RGB	Boundary, Region	Static	Intensity	Spatial Domain	IOOPL algorithm used to eliminate shadow regions.

was classified according to the current mixture models. The weakness in this approach was the poor initialization and labelling which affected the proper identification of shadows.

Mikic *et al.* [22] also used the local pixel-level information but added spatial information. The segmentation was done by comparing the luminance values of each pixel with the mean luminance at that location. A prior probability of the pixel belonging to each of the three classes i.e., background, shadow, and the foreground was estimated using an iterative estimation process called turbo segmentation to propagate neighborhood information. The authors suggested that addition of temporal information would improve the results.

Bevilacqua [23] used the pixel intensity values with due consideration to their membership in the possible foreground
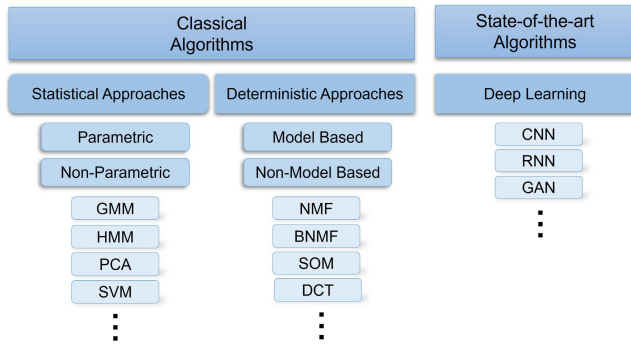
FIGURE 5. Feature based classification of different techniques.

shadow models. The algorithm targeted to search the most probable shadow areas by applying multi-gradient operations on a high-level image derived by dividing the query frame by the background frame. It then performed binary edge matching to remove fewer probable regions using blob analysis.

Porikli and Thornton [24] used the basis of statistical features of shadows cast by objects. After selecting shadow pixels, a Bayesian model was used to form multivariate shadow models. A comparison of the online EM algorithm with Bayesian update was discussed in which the latter showed better results in terms of maintaining the multimodality of the distribution and better estimation of the variance. This technique was adaptive as it used a recursive learning-based method and educated itself on the features of cast shadows automatically through analysis of the developed Gaussians.

Joshi *et al.* [25] used 4 parameters including 3 error values and a ratio to mark the shadow pixels. The error values were derived from the color scheme and magnitude and direction of the edges. On the other hand, the ratio value was based on intensity. The work employed blob analysis exploiting the geometrical features of foreground and shadow objects. The purpose of blobs was to join together segments of an image that probably belong to the same object, which in this case were the shadows. However, the geometric relationship between the blobs needed to be tuned for each video sequence.

Liu *et al.* [26] used three levels of information to remove shadows. At the very basic level called pixel level, GMM was used to model cast shadow in Hue Saturation Value (HSV) color space; at the region level, four neighborhood pixels were modelled using Markov random fields (MRF) to ascertain if they were shadows or not; while the global level used the nearest neighbor tracking method to classify shadows from objects.

Martel-Brisson and Zaccarin [27] used GMM to describe moving cast shadows on surfaces. Major difference between the two methods was that later one employed multi-distribution statistical learning process. The learning was done by identifying the pixel values, then generating of stable shadow distributions and storing them in the GMM-based Gaussian Mixture Shadow Model (GMSM).

These both techniques provided very promising results for indoor scenarios but had poor efficiency for outdoors as it mislabelled any object (vehicle) that was a shade of gray darker than the road. Clearly, these methods suffered from false detections caused by the chromaticity and luminance features of the outdoor subject, thus, making them not suitable for the case of traffic flow analysis.

Pei and Wang [28] proposed a novel method based on GMM and Principal Component Analysis (PCA) to detect moving cast shadows in a scene. The GMM was used for the generation of the background image, while features were extracted using PCA transformation. Then, feature space was used for the classification of foreground objects and their moving shadows. The experimental results showed satisfactory performance in both indoor and outdoor scenarios.

Huang and Chen [29] proposed a confidence-rated Gaussian mixture learning approach for the detection of moving cast shadows. The spatial information was utilized to improve the detection rate by avoiding misclassifications when the foreground is similar to the background. The model was evaluated on popular datasets including Highway I and II datasets and showed a satisfactory performance against state-of-the-art approaches.

Lin *et al.* [30] proposed an algorithm for the removal of moving vehicle cast shadows and efficient extraction of foreground objects using GMM. The non-shadow pixels from objects in the foreground were extracted using the information of gray levels. Similarly, all useful features for locating objects without shadows were then integrated. A practical example of vehicle counting was demonstrated, which showed good real-time performance.

Qi *et al.* [31] proposed a novel cascade method for cast shadow detection. Firstly, the initial moving patches were extracted using GMM. Then, Local Binary Pattern (LBP) was utilized for the separation of moving object pixels from initial moving pixels. Finally, post-processing was performed for accurate identification of moving shadow pixels by correcting misclassified pixels. The performance of this approach was evaluated against various popular methodologies and satisfactory results were obtained.

Jiang *et al.* [32] used YUV color scheme to separate shadow pixels from object pixels in the foreground. The Y component altered when background contained shadow pixels while the UV fairly remained the same. An adaptive threshold estimator was implemented to achieve an unsupervised shadow detection algorithm. The threshold estimator worked by developing a global texture of the image using prewitt edge detector and convolving its horizontal and vertical masks. The pixels in the global texture were expressed as foreground and background. Standard Gaussian distribution was employed using statistical concepts to derive the estimated thresholds. The detection process was adaptable to several dynamic and complex scenes and did not require any manual interventions.

Ouivirach and Dailey [33] constructed a joint probability model using HSV color space in an offline phase. Maximum

Likelihood approach was used to classify foreground pixels as shadows or objects in the online phase. The possibility of detecting incorrect shadow pixels existed because of the similar color between the object and the background. It was opined that incorporating assumptions of geometric and shadow region shapes might improve the performance of this approach.

Russell and Zou [34] proposed a classification-based method to detect moving cast shadows. The paper used a clustering approach to find similar patterns of spatial and temporal color constancy among pixels. The regions with similar patterns were classified as shadow regions. This approach was proven to be particularly useful for the cases when foregrounds had a similar texture as the backgrounds (i.e., foreground-background camouflage). Good results were shown for both indoor and outdoor environments.

Khare *et al.* [35] also used HSV color space but in Discrete Wavelet Transform (DWT) domain. The proposed method depended only on wavelet coefficients that broke a signal into similar and discontinuous sub-signals which helped to classify shadow pixels. A stable threshold was proposed i.e., Relative standard deviation, which proved to be more useful as compared to standard deviation earlier proposed by Guan [20] for detection of shadows in DWT domain. The approach performed very well for indoor and outdoor sequences but performs relatively poorly for non-stationary backgrounds.

Xiang *et al.* [36] proposed Local Intensity Ratio Model (LIRM) via GMM to deal with the influences of illumination variations and shadows. The proposed approach showed good performance for detection of objects without moving cast shadows.

Russell *et al.* [37] worked on a completely new idea of image-line analysis in contract to the previously used methods based on a single pixel or a pixel composite bunch analysis. It was based on the fact that the light intensity of a casted shadow decreased as we moved away from the boundary between the object and its shadow. It required measurements of illumination direction and intensity to discriminate shadow from object. Practical measurements revealed that differences existed in the intensity of any two parts of a shadow. A window operation searched for this condition, termed as the object-shadow line, and hence the shadow was easily detected and removed. The downside of this approach was the requirement of prior knowledge of illumination direction and intensity, which made it difficult to adapt to a different scene.

Valiere *et al.* [38] proposed a robust real-time method that analyzed traffic video streams for vehicle detection, classification, and tracking. The proposed method first segmented the foreground from the background and then detected and classified vehicle objects in the foreground region. The approach was threefold consisting of background subtraction, moving cast shadow removal, and adjusting the occlusions between vehicles management. A GMM based background subtraction technique was presented to find membership of each pixel. The probability of occurrence of color for a given pixel was calculated which was then used to classify the pixel based on its association with the background or the foreground. After that an edge-based moving shadow removal algorithm was employed which had two main purposes; the first was to discard shadow boundary by preserving the edges of the moving objects. The second was the reconstruction of the moving objects based on the information extracted through edges.

Dai *et al.* [39] proposed a technique that was based on a fusion of multiple features including intensity, color, and texture. These features were used to detect moving cast shadows in segmented foreground images through GMM from selected videos. A score was given to each feature set, which was larger if the feature was better in classifying moving objects and shadows. Lastly, a component labeling algorithm removed the minute errors in classification from the shadow and the object.

Farou *et al.* [40] proposed a method for moving cast shadow detection which utilized chromatic properties of different color spaces. Firstly, the canny filter was used to put boundaries on the shadow and background. Then, an improved GMM separated the moving objects from the background. To get rid of shadow parts, the pixels which met a particular threshold (relating to each of the color spaces) were labelled as part of shadows, which were accordingly removed. This model helped separate lighter shadows.

Garg *et al.* [41] used the vehicle-sized blocks approach to define shadows. The vehicle-sized blocks as candidate regions were classified into vehicle or shadow. This helped to escape the heavy computations required in pixel-based region segmenting. It was more of a top-down approach separating vehicles and their shadows. An interior edge feature was used only to differentiate between complex scenarios where the shadows were close to the vehicle and very thin in nature. It achieved a better accuracy than other techniques but with a 20 times faster architecture on a low-cost platform.

Shi and Liu [42] proposed a novel framework for the detection and removal of cast shadows from foregrounds. The framework used Global Foreground Modelling (GFM), GMM, and Bayes classifier for the classification of foreground and background. Initially, the foreground was obtained with shadows, which was then differentiated based on specified criteria. Then, shadow region detection method was used to detect shadows which were then classified with Gaussian distribution. Aggregated shadow detection was then used to combine all results obtained from previous steps. The model was tested on popular benchmark datasets including Highway I and Highway II etc. The results indicated that the proposed method performed more efficiently for shadow detection task compared to other famous methods.

Sun *et al.* [43] proposed a robust vehicle detection approach that combined optical flow with shadow removal technique to eliminate the interference of shadows. Shadow regions were detected based on color features of shadows in the HSV color space, which were then removed through a

region labeling algorithm. The test results indicated a good performance for daytime detection of vehicles with long shadows.

Zhang *et al.* [44] proposed a vehicle detection method while utilizing shadow detection and elimination to improve detection accuracy. The foreground regions were extracted by a background differential method using edge information. Then, the shadows were eliminated from the foreground regions using grayscale and edge information, as well as prior knowledge. The authors proved the superiority of the proposed approach over various state-of-the-art approaches.

Ghahremannezhad *et al.* [45] proposed an approach for the detection and removal of moving cast shadows that utilized pixel and region-based techniques, as well as statistical modeling for the detection of shadows. GFM method was firstly used to segment moving objects and their shadows. After that, a new region-based approach was proposed which used k-means to perform partition between object and shadow regions. Then, the foreground and background values in different color spaces were used to construct six-dimensional feature vectors, which were modelled using statistics to classify the foreground pixels into shadows and objects. Finally, the results of all these steps were integrated to perform a robust shadow detection.

Zhou *et al.* [46] proposed a shadow suppression approach based on a combination of color features and Histogram of Local Gradient Binary Patterns (HLGP) features. Firstly, the shadow was detected using chromaticity and brightness similarity features. Then, HLGP features were used to provide robustness against illumination. The results indicated an improved vehicle detection accuracy.

Shi [47] proposed four statistical-based models to detect the objects in videos. Firstly, a GFM method was used for foreground object detection. It further used a Local Background Modeling (LBM) method to model the background. Then, Haar wavelet features and temporal information were used to form a 12-dimensional feature vector, which was then classified through Bayes Classifier. Secondly, a shadow region detection method was proposed which used a single Gaussian density to model the shadow class for each pixel. The probability density function of the pixels was estimated using GMM. Thirdly, a model to solve automated road recognition problems using temporal features was proposed. Finally, a driving detection approach was proposed which detected abnormal driving patterns in videos.

2) STATISTICAL APPROACHES: NON-PARAMETRIC

Unlike statistical parametric based algorithms, algorithms in non-parametric statistical approach make no assumptions about the probability distribution of the feature set i.e., number of parameters dependent upon the size of training data. There are total 18 papers that have been identified by the authors for review in SNP approach (Table 2).

Rittscher *et al.* [48] proposed a probabilistic approach based on Hidden Markov Model (HMM), which

discriminated between foreground, background, and shadow regions. Further, probabilistic trackers based on particle filters were also used. The overall approach was used for vehicle tracking applications and showed a good capability to be used as a robust tracker.

Siala *et al.* [49] used a diagonal model for the RGB color space to identify shadow distortion. The authors considered shadow as a case of illumination change and then applied a Support Vector Domain Description (SVDD) algorithm in the space of color ratios which helped to properly discriminate shadow pixels from the foreground. Two different datasets were used to gauge the developed algorithm but the accuracy of the results for shadow detection and discrimination was low especially for the case where foreground detection was clearly affected by strong self-shadows.

Wang *et al.* [50] used HMM along with background subtraction to identify moving cast shadows. In order to initialize the Gaussian observation model of HMM, Maximum Likelihood Estimation (MLE) was employed which avoided the local maximum. Baum-Welsh (BW) estimated the unknown parameters of HMM after the training data was gathered. The Viterbi algorithm was used to decode the state sequences associated with background, shadow, and foreground. The process was computationally expensive, and the training took a considerable amount of time and needed to be tuned for each video.

Martel-Brisson and Zaccarin [51] introduced the non-parametric framework for learning the cast shadows. Analysis of the properties of light sources and object surfaces was carried out to find such regions in the image that correlated to the background with some variations. These regions were probably the shadow regions since shadows generally could not completely hide the background and some features of the background were still identifiable. The major advantage of this method was that it was completely unsupervised and learned the model parameters through scene activity.

Joshi and Papanikolopoulos [52] also proposed a color-based shadow detection method for distinguishing moving cast shadows from objects but incorporated a semi-supervised learning technique. The proposed method devised a set of features useful for classification by leveraging characteristic differences in color and edges in the video frames. This was followed by a learning technique that used support vector machines and a co-training algorithm, which relied on human-labelled data, for shadow detection. The authors demonstrated that the use of semi-supervised learning made the technique robust to varying scene conditions, and once deployed, the proposed technique could automatically be adapted to the varying scene and illumination conditions.

Vargas *et al.* [53] proposed an approach which involved conducting an analysis of the image obtained from the division of the query frame by the background. This process, often lead to the partition of moving vehicle into multiple blobs which could be overcome by applying a subsequent clustering procedure to reunify the separated regions of the

same vehicle. On a quotient image which was obtained from the current frame and the background model.

Amato *et al.* [54] proposed a technique for the detection of shadows that exploited the embedded background features in the shadows. The intensity variations in the shadow region contained significant similarities to that of the background. The shadow regions were detected by dividing the background image values by the values of the queried frame. It was shown that this intensity ratio could be used to identify the low variation segments that exclusively distinguished the shadow from the foreground. This technique detected both the achromatic as well as camouflaged chromatic shadows that occurred due to the similarity in the foreground and the shadow regions.

Meher and Murty [55] proposed a model to improve the detection and classification of moving objects by removal of moving shadows which generated object localization errors. PCA was used to minimize search space for shadow regions. The detected shadow regions were then removed while keeping only moving vehicle regions. The model was further improved by using Scale-invariant Feature Transform (SIFT) features for the classification of vehicles (both with and without shadows).

Wang and Zhang [56] proposed a two-step approach for the detection of vehicles. Firstly, the potential location of vehicles was assumed through extensive searching of shadows under vehicles. Haar-like features with Adaboost were used to train a Haar detector in offline mode for the detection of shadows. The hard sample training approach was used to eliminate false detection. Based on these detected areas, vehicle detection was performed in further steps using a combination of different algorithms including Histogram of Oriented Gradients (HoG), SVM, and K-means. The results showed satisfactory performance with a real-time processing capability.

Gomes *et al.* [57] proposed a shadow detection method having adaptive and non-adaptive versions. Using the Lab color space, a weighted hypergraph was constructed which split partitions that consist of more than one group of foreground pixels. To separate the shadow gradient and color, correlation data was used. Furthermore, HSV color was used to partition regions from shadow and non-shadow. Hypergraph partitioning allowed to classify whole region as a shadow or vice versa depending on the pixels present. Lastly, a shadow mask was obtained that contained all shadow pixels.

Wang *et al.* [58] discussed the influence of shadows on vehicle detection and proposed a shadow elimination method based on PCA. The approach firstly weakened the shadow areas to make them look similar to the background. Then, shadow pixels were separated from the moving vehicle. The results indicated a 10.3% to 13.3% improvement in shadow elimination compared to other conventional algorithms.

Yang and Siu [59] proposed both learning and non-learning approaches for the detection of shadow patches. A cascade detector was used to examine features in the non-learning-based approach, while a modified decision tree was used as a learning-based approach. The modified decision tree was also

compared with SVM and showed better performance than the later. Experimentation showed satisfactory performance with both approaches.

Yi *et al.* [60] proposed a novel shadow detection approach based on Extreme Learning Machine (ELM), which distinguished between the background (i.e., shadow) and foreground objects. Firstly, the pixel and region-level features were extracted from the foreground. Then, the ELM approach classified the shadow and non-shadow points which helped with the detection of the shadow region. Post-processing was performed to further improve the performance of the moving cast shadow detection algorithm. The proposed approach showed good performance in comparison to different state-of-the-art methods.

Zhu and Yin [61] proposed a shadow detection method that used SVM to train an RGB image with the shadow which was then divided into shadow and light regions. The light intensity in the shadow region was adjusted by the elimination of pixel differences between both regions. Image gradient was used on boundary shadows, which was replaced by smooth interpolation for a gradual transition from light to shadow region. The approach successfully detected the shadow regions and reproduced the images without shadows.

Kan and Wang [62] proposed a moving shadow detection approach based on a Semi-supervised Extreme Learning Machine (S-ELM). Firstly, pixel and regional level features were extracted, which were then trained using S-ELM to classify foreground and background pixels. Necessary post-processing was performed to remove the effect of noise and overall results improvement. The proposed method showed good performance against various state-of-the-art methods in both indoor and outdoor environments.

Lu *et al.* [63] proposed a shadow removal method based on point cloud features similarity for shadow-affected vehicles and pedestrians using an event camera. Different point distribution characteristics are presented by each traffic entry, which were then further classified into geometrical, quantitative, and Gaussian projection features. Shadows were detected and removed using the feature weights calculated using the Relief-F algorithm and Kernel Density Estimation. The experimental results indicated a shadow elimination rate of 96.5% in shadow samples.

Anandhalli *et al.* [64] proposed a Shi-Tomasi-based corner detection approach for vehicle detection and tracking under rough climate conditions and shadows etc. Various corner points from vehicle regions were segmented from non-vehicle regions using a background corner point model. Similarly, the foreground corner points belonging to vehicle regions were grouped together by utilizing the Euclidean distances. The results indicated a good vehicle detection performance even with long shadows and illumination changes.

Sahoo and Nanda proposed a method to detect moving objects in a video after the removal of shadows [65]. The background was modelled by the proposed Spatio-Temporal Kernel Density Estimation (ST-KDE) model. The LBP

features of ST-KDE model were then fused in online mode with the Gabor features of the corresponding original frame. The learning weights were calculated based on the scene dynamics. After learning and classification, the residual shadows were eliminated by entropy map and thresholding.

3) DETERMINISTIC APPROACHES: MODEL-BASED

The model-based deterministic approach compares the features extracted from different regions of a scene with a specified model. The positive comparison result indicates the presence of shadow and vice versa. However, such methods lack generality since it is difficult to foresee all possible scenarios. Moreover, increasing the number of models increases the algorithm complexity. Total 9 papers have been identified for review in the DM approach (Table 3).

Koller *et al.* [66] presented a technique to detect and track moving vehicles in video sequence recorded by a stationary camera. The technique exploited a-priori knowledge about the shape and motion of vehicles, using a parameterized model for the intra-frame matching process and a motion model-based recursive estimator for motion estimation. An image analysis method looked for pixels maintaining the same relation in the successive frames of traffic video streams. The vehicles were detected using the assumption that the related pixel belonging to the same vehicle will maintain their relation in terms of separation from each other. An illumination model was included so that shadow edges of vehicles could be accounted for during the matching process. The combination of multiple techniques enabled the tracking of vehicles under complex illumination conditions. However, based on a parameterized vehicle model, the technique was not easily adaptable to various kinds of vehicles.

Onoguchi [67] proposed a method for the elimination of shadows casted by moving objects. Assuming that the shadows casted by moving objects lie on the road plane, the proposed technique leveraged height information to eliminate shadows. Two cameras were placed such that their common visual fields include the area to be monitored. The image obtained from one of the cameras was inversely projected to the road plane and the image projected on the road plane was transformed to the view from the other camera using pre-estimated image transformation parameters that were obtained by indicating several corresponding points between the images acquired from the two cameras. The true shadows were identified through analysis of the areas that were occupied in the transformed and other camera images. This allowed for shadow areas to be removed by simply subtracting between these images. While the proposed technique was independent of the object type and road color, it required shadows to be on the flat road plane and in the visibility of both cameras. Moreover, it was not easily adaptable to a new surveillance area as, in addition to the correct placement of cameras at different locations, the image transformation parameters also needed to be estimated for every scene.

Yoneyama *et al.* [68] proposed a model-based technique for the detection of vehicles in a highway monitoring system. The proposed shadow elimination technique was based on a six-vertex joint 2D shadow/vehicle model of six types projected to a 2D image plane. The parameters of vehicle and shadow models were estimated from the input video by luminance analysis and without the need of light source and camera calibration information. The authors claimed that the algorithm was of low computational complexity as it did not perform any 3D image analysis. Moreover, the shadow region was distinguished from the vehicle via the determination of parameters of the joint model instead of two separate models, thus further reducing computational complexity. Based on experimental results, the authors demonstrated that the proposed technique performed well irrespective of the camera orientation, calibration, and variations in the light sources.

Salvador *et al.* [69] proposed a shadow segmentation technique applicable to both, still and moving shadows. The proposed technique used spectral and geometric properties of shadows in a scene. First, the probable shadows were identified using the assumption that shadows increased intensity of the regions upon which they were casted. This was followed by further verification based on physical and geometric features of shadows. Based on the extracted information, the final stage performed a binary decision corresponding to the acceptance or rejection of a region as a shadow. The authors demonstrated that the proposed technique could be applied to a large class of scenes without requiring any change in parameters by evaluating the proposed technique on different kinds of scenes.

Nadimi and Bhanu [70] proposed an approach for separating moving cast shadows from the moving objects that relied solely on physical models for shadow and object detection and did not make any assumptions about surface geometries and textures, types and shapes of shadows, objects, and backgrounds. The proposed technique was based on a Spatio-temporal albedo test and dichromatic reflection model. Multiple illumination sources with different Spectral Power Distribution (SPD) were incorporated. The proposed technique utilized a temporally extended spatio albedo ratio test for surface segmentation. While authors demonstrated that the proposed technique was robust to varying background surfaces, foreground materials, and illumination conditions, they conceded that it required the Spectral Power Distribution (SPD) of each source of illumination to be constant.

Sun and Li [71] proposed a method for detection of moving cast shadows of vehicles using combined color models. Using the observation inference of shadow pixel intensities remaining lower than the object pixel for majority of observations, the ratio of the hue and intensity in the HSI color space was used to identify the object pixels in foreground. Three typical photometric color invariants with color models were defined as $c1$, $c2$, $c3$. Then, they employed the theory of photometric color invariants in the color model to distinguish the dark (similar to shadows) and colorful object pixels from the shadow pixels. The two images obtained from these methods were then synthesized to obtain a rough shadow image.

Finally, post-processing was applied to correct shadow detection failure and correct object detection failure. The authors selected two sequences (one containing road crossing, other from Jingzhu Highway), with shadow, and reported better performance by comparing the proposed method with two well-known models i.e., SNP and DNM.

Chacon-Murguia and Gonzalez-Duarte [72] proposed an adaptive object detection approach focused on dynamic backgrounds. Self-Organizing Map (SOM) based architecture was utilized to deal with the dynamic backgrounds and elimination of shadows. The framework automatically adjusted the main parameters, making it capable to work without human intervention. The performance evaluation was conducted on nine different videos with varying backgrounds and overall satisfactory performance was reported.

Yang *et al.* [73] proposed two shadow detection approaches based on Non-negative Matrix Factorization (NMF) and Block Non-negative Matrix Factorization (BNMF). The performance of both algorithms was evaluated and the detection results of BNMF proved to be better than the NMF method. The BNMF method allowed the inclusion of new samples and classes without the need for re-execution which significantly lowered the computational complexity. The algorithm not only successfully detected moving cast shadow areas, but also classified different object types.

Hu and Liu [74] proposed a shadow elimination method based on multi-feature differences between the shadow region and the corresponding background region. The approach incorporated luminance, chrominance, and texture differences of the foreground and background regions as features. The feature differences were utilized to form a training sample, which was then fed into the Generalized Learning Vector Quantization (GLVQ) model to find whether the pixel belonged to shadow or not. Experiments were conducted on a custom road monitoring video. The results indicated better performance of the approach compared to single-feature-based methods.

4) DETERMINISTIC APPROACHES: NON-MODEL-BASED

The non-model based deterministic approaches use some relatively generalized criteria rather than strict models. The non-model methods may set criteria like ratio and luminance of pixel values in successive frames to decide their membership. Shadow pixels of the same vehicle usually maintain a somewhat constant ratio with the background. Also, they are darker and therefore have low luminance values compared to the background. A total of 7 papers have been identified for review in the DNM approach. (Table 4)

Amamoto and Fujii [75] proposed a method for the detection of vehicles on a road. The proposed approach combined background and time differences to detect the varying regions in the image. The detected varying regions were further classified into moving objects, stationary objects, and variations in the image due to illumination variation. The variations that occured due to varying illumination were used to update the

background. In order to separate the detected objects from their shadows, the image was first converted to the spatial frequency domain by using the DCT. Suggesting that the shadow of an object simply varied the pixel values uniformly in comparison to the background, the authors inferred that a significant dc component indicates an object shadow, whereas a significant ac component indicated a moving object. The proposed method extracted the moving object by using the ac component. The authors reported a shadow elimination rate of 95.5% on an evaluation dataset. However, there were also some "dropouts" in the object regions, where parts of the object were eliminated with the shadow, specifically in the case of a black car.

Stander *et al.* [76] proposed an intelligent method for the detection of moving shadows. The technique was based on four assumptions. Firstly, the hypothesis said that the shadows were formed by the strongest of the light sources illuminating a scene. The next assumption said that the image stream was captured from a static camera and therefore saw the same background with the possibility of moving objects in the foreground. The third assumption said that the background scene was static and not dynamic. The final assumption correctly assumed the distance between the light source and the moving objects to be significantly larger compared to the distances between the objects and the background. Resultantly, based upon the aforementioned 4 assumptions, a binary decision on the membership of pixel whether or not belonging to a shadow was taken. The proposed method was based on assumptions that did not hold in real-world applications. The authors conceded that for their specific technique, the shadows that were weak, that had a highly structured background, or that had contours as sharp as object edges, the assumptions did not hold. Such shadows could not be detected. Moreover, the method could only detect shadows that were moving.

Cucchiara *et al.* [77] proposed a method for the segmentation of moving objects based on object-level classification of moving objects, ghosts, and shadows using motion and shadow information to extract objects and their shadows from the background model while retaining their ghosts. The method defined an approach for shadow detection and suppression based on color analysis in the HSV color space. The proposed method was independent of any prior knowledge of the scene.

Cucchiara *et al.* proposed an approach, in which the shadow regions were detected based on the comparison of texture descriptors and photometric properties [78]. The proposed approach was largely unaffected by differing light conditions and object classes because it relied on the texture descriptors which remain unchanged. Their updated HSV color space-based technique for detection of objects, ghosts (artifacts on shadow boundaries), and shadows detection integrated object-level knowledge into a statistical background model. The pixels belonging to moving objects, shadows, and ghosts were processed differently in order to supply an object-based selective update. The authors suggested that

when a shadow was casted on a background, the hue and saturation components changed only within certain limits, and that the difference in saturation was an absolute difference and the difference in hue was an angular difference. The presence of a shadow was established based on these considerations.

Toth *et al.* [79] proposed an algorithm that discriminated against moving objects from their shadows. The proposed method first divided the changed region into sub-regions that consisted of pixels having similar color properties using a non-parametric mean shift algorithm. This was followed by a significance test that classified each of the pixels as belonging to either a shadow or an object. Finally, the global and local information from the first two steps was combined to obtain a refined change mask that represented the object. The proposed system was largely independent of lighting conditions and could be adapted for outdoor applications.

Leone and Distante [80] proposed a technique for shadow detection of moving objects that uses an automatic segmentation procedure based on adaptive background subtraction. Utilizing the fact that shadows are half transparent regions that retain features of the underlying background surface, the approach labels as shadows the regions having a substantially unchanged structure with respect to the reference background frame.

Srividhya *et al.* [81] proposed a vehicle detection and segmentation approach using a delta learning algorithm. It also addressed the problem of a shadow being recognized as part of the object itself. The approach utilized Inner-Outer Outline Profile (IOOPL) algorithm to eliminate shadows. IOOPL extracted the image with a shadow and then allowed the objects (vehicles) in that image to be multi-layered after the application of Gaussian smoothing. These multiple layers were then used to mark the objects which helped to detect exact shadow-less object boundaries by separating background, shadow, and objects.

B. STATE-OF-THE-ART ALGORITHMS

The learning-based approaches utilize computational methods to learn information from the collected data (known as training data). This information is used to make predictions without relying on a predetermined equation. As highlighted before, there are limited publications using state-of-the-art algorithms which are focused on shadow detection techniques in traffic paradigm. Therefore, only a total of 9 papers have been reviewed in this section (Table 5). Summary of each paper is provided as followed:

Vicente *et al.* [82] presented a large-scale dataset training approach for detection of shadows. The datasets contained a variety of scenes and image types including vehicles. A semantic-aware patch-level CNN model was then used to train on shadow patches while also incorporating the image-level semantic information.

Li *et al.* [83] proposed a novel Faster R-CNN based automatic and accurate vehicle and shadow regions detection model from Mobile Mapping System (MMS) images. The

results indicated a good recall of around 96.3%. The model was successfully able to identify vehicle and shadow regions even with different shadow directions and partial occlusions etc.

Bakr *et al.* [84] proposed a Mask R-CNN-based approach for shadow detection which automatically extracted the shadow features and also performed object detection. The distinctive features were extracted using a deep residual network (ResNet-101). Then, Region Proposal Network (RPN) was used to predict ROIs and the classes which contained foreground objects. A segmentation mask for each detected class was then generated through the fully convolutional network. The proposed algorithm was tested on various popular datasets including vehicle-related Highway I dataset and achieved an average detection rate of 96.81% without any additional post-processing.

Fang *et al.* [85] presented a dataset that targeted to find both shadow and object instances and then paired them. A Light-guided Instance Shadow-object Association-based framework was proposed for the automatic prediction of boxes and masks of shadow and object instances. These predicted instances were then paired up and matched with the predicted shadow-object associations for the generation of final evaluation results. A new evaluation metric was also proposed to perform evaluations against various baseline frameworks.

Chen *et al.* [86] explored shadow detection in dynamic scenes by collecting a video shadow detection dataset, ViSha, which contained different classes including vehicles. Authors also proposed a baseline model Triple Cooperative Video Shadow Detection Network (TVSD-Net) which made use of parallel networks in a cooperative way to learn inter-video and intra-video shadow discriminative properties. The proposed method showed a good performance against different state-of-the-art relevant methods.

Cao *et al.* [87] proposed an Object-aware Shadow Detection Network (OSD-Net) model for retention of key objects in a complex scenario. Firstly, large shadow areas were detected by the shadow detection module, which used ResNeXt-101 as the backbone network, followed by Direction-aware Spatial Context (DSC) module. The key target objects were then segmented using Mask R-CNN network. Finally, both networks were combined to predict the shadow mask.

Bao *et al.* [88] proposed a deep learning model for shadow detection of moving ground targets on video Synthetic Aperture Radar (SAR) data. Five different tools were utilized to guarantee the excellent performance of the proposed model, which focused on feature extraction, elimination of clutters, computing the speed of moving targets, matching shadow locations and shapes, and hard mining techniques to boost the background discrimination capacity of the model. The experimental results indicated better performance compared to other state-of-the-art methods while sacrificing slight detection speeds within the acceptable ranges.

Peng *et al.* [89] proposed an approach for automatic smoky vehicle detection in videos while distinguishing this smoke

TABLE 5. State-of-the-art algorithms for shadow detection.

Paper	Dataset(s)	Level	Temporal	Key Feature	Domain	Technique Used
Vicente et al. (2016) [82]	Custom	Patch, Image	Static	CNN-based features	Spatial	Semantic-aware patch-level CNN model for training of shadow patches
Li et al. (2018) [83]	Custom	Region	Static	CNN-based features	Spatial	Vehicle and shadow detection based on Faster R-CNN from MMS images
Bakr et al. (2020) [84]	Highway I, Bungalows, Bus station etc.	Pixel	Static	CNN-based features	Spatial	A Mask R-CNN based approach for shadows and objects detection
Fang et al. (2020) [85]	Custom	Patch	Static	CNN-based features	Spatial	Light-guided Instance Shadow-object Association based framework to find shadow and object instances
Chen et al. (2021) [86]	Custom	Pixel	Static	CNN-based features	Spatial	A baseline model TVSD-Net to learn shadow discriminative properties
Cao et al. (2021) [87]	Virtual KITTI 2, KITTI, UQ St Lucia Stereo Vehicular Dataset	Region	Static	CNN-based features	Spatial	Shadow detection by ResNeXt-101 network as backbone, followed by DSC module. Segmentation of targets using Mask RCNN and generation of fusion masks for prediction.
Bao et al. (2022) [88]	Sandia National Laboratories (SNL) video SAR data	Region	Static	CNN-based features	Spatial	Deep learning model applied on video SAR images to detect shadows of moving ground targets.
Peng et al. (2022) [89]	Custom, SEU-SmokeVeh etc	Region	Static	CNN-based features	Spatio Temporal Domain	Smoke-vehicle matching using IoU method for identification of false positives, which were then eliminated using a deep learning model.
Arora et al. (2022) [90]	Custom	Pixel	Static	CNN-based features	Spatial Domain	Pixel probabilities calculated for identification of shadows.

from shadow regions due to cluttered roads. The smoke regions identified through a deep learning model were passed through a smoke-vehicle matching module which made the smoke region and the certain vehicle a pair, based on their Intersection-over-Union (IoU) ratios. The same matching module was also proved to be helpful in identification of other non-vehicle regions. Finally, a light-weight 3D model was used to eliminate these false positives and further refine the results in spatial temporal space.

Arora *et al.* [90] proposed a moving vehicle detection during both day and night times using the Fast Region-based CNN (Fast R-CNN) deep learning model. The proposed work showed good performance even in the presence of long shadows and other rough conditions. The method used three Gaussian mixtures related to vehicle, road, and shadow for all background pixels to find whether the pixels belonged to foreground or background. The current pixel probability was then calculated to compute a foreground mask for the identification of the desired area. Then, Kalman Filtering was used to identify the position features of the moving vehicles. Finally, the Fast R-CNN model utilized these features for the successful detection and classification of moving vehicles.

IV. COMPARATIVE ANALYSIS OF SHADOW DETECTION AND REMOVAL APPROACHES

A. DATASETS
The reported works have used a variety of datasets for evaluation in indoor and outdoor environments. Keeping in view the scope of this review, the datasets containing traffic scenarios are identified. The most popular of these are Highway I,

Highway II and Highway III [80]. The mentioned datasets are courtesy of *Computer Vision and Robotics Research Laboratory* of *University of California San Diego (UCSD)*. Following are the characteristic parameters of this dataset:

- Moving and cast shadows
- Presence of only one light source i.e. Sun
- High light source intensity
- Fixed position of light source
- Static camera and background
- Frontal camera viewing angle
- Medium speed of vehicles
- Cast shadows of three types i.e., small, medium, large
- Varying vehicle types
- Three viewing angles i.e., near, medium and far field
- No vehicle-to-vehicle occlusions

Ground truth information in terms of vehicle classification, time duration and vehicle count of the selected benchmark datasets is mentioned in Table 6. Additional detailed characteristics of the datasets on which most of the works are done in literature are presented in Table 7.

B. EVALUATION METRICS
Evaluating any shadow detection algorithm in a systematic way requires algorithm assessment at two key frontiers; good detection and good discrimination. An algorithm for shadow detection must be able to have a high probability to detect a shadow with minimum chances of missing any shadow point. The algorithm must also be able to clearly discriminate between shadow and non-shadow points, to have little to no chance of erroneously classifying a non-shadow point as a shadow point.

TABLE 6. Ground truth of standard datasets.

Dataset Name	Highway I	Highway II	Highway III
Duration	00:00:29	00:00:35	00:03:49
Cars	77	40	158
Bikes	-	-	3
Bus	-	-	1
Van	11	5	15
Truck	2	1	3
SUV	14	6	2
Total Vehicles	104	52	182

Classifying shadow points as a part of foreground or background is termed as False Negatives (FN); and keeping FN to a minimum level is a key characteristic of good detection. Likewise, classifying foreground or background points as shadow points is termed as False Positives (FP); and keeping FP to a minimum level is a key characteristic of good discrimination [5]. For evaluation of algorithms involving moving object detection, Onoguchi [67] came up with the proposition of two assessment parameters; False Alarm Rate (FAR) and Detection Rate (DR) defined in equations 3 and 4 respectively. These below mentioned parameters are defined based on TP which are the total number of shadow points identified correctly.

$$FAR = \frac{FP}{TP + FP} \tag{3}$$

$$DR = \frac{TP}{TP + FN} \tag{4}$$

The Onoguchi parameters were deemed insufficient by Prati *et al.* [5] for evaluation of a shadow detection algorithm; because these parameters do not help in determining whether the identified shadow point belongs to the background object or the foreground object. Utilization of shadow detection for rendering improvements to the moving object detection algorithms makes the first case problematic. The reason behind it is that the FPs which belong to the background do not have any effect in detecting the object or determining its shape. Taking this consideration into account, Prati *et al.* [5] upgraded these parameters; thereby coining the metrics of shadow detection rate η and shadow discrimination rate ζ already defined by equations 1 and 2, respectively.

C. DISCUSSION BASED ON TAXONOMY

In this section, the performance of the techniques is analyzed to understand if they can be used as a preprocessing step. Results are directly obtained from the respective research works and are compiled and presented in Table 8. The average efficiency values of considered papers are shown in Table 9. Figure 6 shows the detection and discrimination performance for all considered datasets. Bars in the figure represent average value of the detection rate (η) and discrimination (ζ) rate for each category.

TABLE 7. Highlights of the highway I, II, and III datasets.

Dataset	Highway I	Highway II	Highway III
Video Frame			
Shadow Sample			
Sequence Type	Outdoor	Outdoor	Outdoor
Video Length	00:00:29	00:00:35	00:03:49
Frame Rate	14 frames/second	14 frames/second	10 frames/second
Image Size	320 x 240	320 x 240	320 x 240
Shadow Strength	Medium	Medium	Medium
Shadow Size	Large	Small	Medium
Shadow Type	Moving Caste Shadow	Moving Caste Shadow	Moving Caste Shadow
Object Class	Vehicle	Vehicle	Vehicle
Object Types	Sedan/hashback, Van, Truck, SUV	Sedan, Van, Truck, SUV	Sedan, Van, Truck, SUV, Bike, Bus
Depth of field	near-field-view	medium-field-view	Far-field-view
Object Speed	Fast	Medium	Medium
Direction	Incoming vehicles	Outgoing vehicles	Incoming vehicles
Noise Level	Low	Low	Low
Camera position	Top view	Top inclined view	Top view
Vehicle to Vehicle Occlusion	None	None	None

TABLE 8. Experimental results of selected papers.

| Paper	Category	Results (Average)					
		Highway I		Highway II		Highway III	
		η	ζ	η	ζ	η	ζ
Mikic et al. (2000) [22]	SP	59.59%	84.70%	46.93%	91.49%	-	-
Joshi et al. (2007) [25]		84.11%	96.69%	-	-	88.57%	93.78%
Liu et al. (2007) [26]		72.13%	79.71%	-	-	-	-
Martel-Brisson and Zaccarin (2007) [27]		63.30%	71.30%	-	-	58.51%	44%
Pei and Wang (2009) [28]		86.94%	89.97%	-	-	-	-
Huang and Chen (2009) [29]		70.83%	82.37%	76.50%	74.51%	82.05%	90.47%
Qi et al. (2014) [31]		82.67%	97.41%	-	-	11l-	-
Ouivirach and Dailey (2013) [33]		93.38%	45.27%	-	-	-	-
Russell and Zou (2013) [34]		85.4%	88.7%	-	-	-	-
Xiang et al. (2014) [36]		74.6%	66.8%	-	-	-	-
Dai et al. (2016) [39]		85.97%	94.68%	-	-	-	-
Farou et al. (2017) [40]		79.29%	98.78%	-	-	79.08%	94.68%
Garg et al. (2019) [41]		84.56%	97.23%	-	-	-	-
Shi et al. (2019) [42]		89%	94%	-	-	90%	76%
Ghahremannezhad et al. (2021) [45]		90%	93%	-	-	90%	86%
Siala et al. (2004) [49]	SNP	77.51%	67.53%	-	-	-	-
Wang et al. (2008) [50]		92.10%	89.20%	-	-	-	-
Martel-Brisson and Zaccarin (2008) [51]		70.50%	84.40%	-	-	68.40%	71%
Joshi and Papanikolopoulos (2008) [52]		88.21%	97.0%	-	-	92.33%	95.47%
Amato et al. (2011) [54]		81.5%	85%	-	-	72%	75%
Gomes et al. (2016) [57]		87.92%	94.67%	-	-	69.10%	85.78%
Yi et al. (2019) [60]		88.24%	94.66%	-	-	-	-
Kan and Wang (2021) [62]		80.11%	96.42%	-	-	-	-
Sahoo and Nanda (2022) [65]		87.53%	94.23%	-	-	-	-
Yang et al. (2018) [73]	DM	94.55%	85.19%	-	-	-	-
Stander et al. (1999) [76]	DNM	75.49%	62.38%	60.24%	72.50%	-	-
Cucchiara et al. (2001) [77]		48.7%	82.7%	-	-	-	-
Cucchiara et al. (2003) [78]		69.72%	76.93%	54.07%	78.93%	-	-
Bakr et al. (2020) [84]	SOTA	95.10%	99.23%	-	-	-	-

TABLE 9. Average efficiency values of both conventional and state-of-the-art approaches on standard datasets.

Classification	Highway I		Highway II		Highway III	
	η (avg)	ζ (avg)	η (avg)	ζ (avg)	η (avg)	ζ (avg)
Classical – SP	80.12%	85.37%	61.72%	83.00%%	81.36%	80.82%
Classical – SNP	83.73%	89.23%	N/A	N/A	75.45%	81.81%
Classical – DM	94.55%	85.19%	N/A	N/A	N/A	N/A
Classical – DNM	64.64%	74.00%	57.16%	75.72%	N/A	N/A
State-of-the-Art	95.10%	99.23%	N/A	N/A	N/A	N/A

State-of-the-Art based approaches provide best results for both shadow detection and discrimination, with an average of more than 95% for Highway I dataset. Though, this category contains a total of 9 papers, but the results are available for Highway I only from the work of Bakr *et al.* [84]. The DM based approach provides best results for both η and ζ, with an average of around 90% for Highway I dataset. This is because it uses more assumptions in its algorithm compared to other approaches. On the other hand, DNM based approach has a very low average performance for Highway I and Highway II datasets because DNM based systems are not able to better classify the large shadows compared to other approaches. However, for Highway III dataset, which consists of small shadows, the average results for both SP and SNP based approaches are almost the same. SP approach performs good in most cases. However, its detection rate is relatively low for Highway II dataset. Major drawback in this approach is the selection of parameters. SNP approach shows diverse output for given datasets. It achieves good η and ζ for Highway I

and Highway III datasets. However, performance of η does not exceed more than 85%.

Comparing overall performance of the conventional algorithms, it can be deduced that model based approaches perform better than the other approaches. The only downside is that deterministic model based approach increases the complexity and processing time as compared to non-model based approaches as it requires modeling of every class. This becomes difficult particularly for the scenarios like different viewing angles, multiple vehicle types and diverse lighting conditions or specific environments like Highway I, which has large shadow sizes. In comparison, the State-of-the-art approaches provide excellent results for outdoor traffic scenarios particularly in Highway I dataset which has large medium strength shadows. Therefore, in the case of moving cast shadow scenarios, it can be concluded that state-of-the-art approaches are relatively better in terms of performance. A qualitative evaluation, based on the additional metrics of some selected papers from each section of taxonomy is given

Detection Rate (η)

Discrimination Rate (ζ)
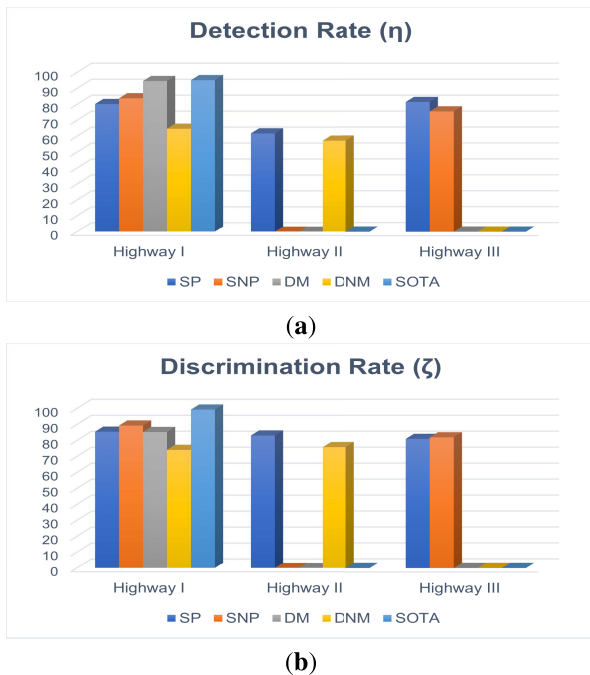
FIGURE 6. Comparison of shadow detection performance by dataset (a) Detection rate and (b) Discrimination rate.

in Table 10. The methods are rated as low (L), medium (M) or high (H) according to five criterias. These five criterias include robustness to noise, computational complexity, shadow sharpness, illumination independence and scene independence. Overall, mostly SP approaches are relatively more robust to noise as well as less computationally expensive. For the case of illumination independence, almost all the approaches have same average capacity to deal with varying illumination conditions i.e., few algorithms in each category perform better than others for different light conditions. State-of-the-Art approaches seem to be more versatile and have shown better results for different scenarios than other approaches.

V. CASE STUDY: SHADOW REMOVAL USING GAN-BASED PRE-TRAINED GHOST-FREE SHADOW REMOVAL APPROACH FOR IMPROVED CNN-BASED VEHICLES DETECTION

Based on comparison of conventional and state-of-the-art based algorithms discussed in literature, it is proved that the state-of-the-art approaches are indeed better than the former. The conventional approaches in comparison are less capable of picking up the diversity and hence cannot cater to the changes. A state-of-the-art shadow detection and removal approach was used to study effects on vehicle detection. Dual Hierarchical Aggregation Network and Shadow Matting GAN based pre-trained ghost free shadow removal model [91] has been implemented for the investigation of shadow detection and elimination using a customized and complex dataset with three different views i.e., front, rear and

side views. YOLOv5 model has been used to predict vehicle classes with and without shadow removal incorporated. Evaluation of the obtained results is given below.

Using the front view dataset (see Figure 7 a), the classification accuracy of most of the vehicles has been improved but the accuracy of the car at the front remained the same, i.e., 82%. However, after removing the shadow, rickshaw which is categorized as truck, has been falsely classified as a car.

Nevertheless, using the rear view dataset (see Figure 7 b), the overall contrast has been increased where shadow was detected, although shadow seems to be visible but increasing the contrast has improved the YOLO detection and classification results. In addition, the vehicle got correctly classified as car which was previously predicted as bus in the image with shadow. However, YOLO failed to recognize a motorcycle which was previously detected in the image without shadow removal.

On the other hand using side view dataset (see Figure 7 c), contrast has been increased with the detection of shadow; however the shadow is still visible but it has enlightened the road area which has helped the YOLO model in better detection with lower false prediction. Moreover, the classification accuracy either remained the same or further improved. Also, a person is detected as a vehicle, and it has been categorized as car which is clearly a wrong detection as well as false classification.

The proposed approach has successfully improved the vehicle detection with varying shadow sharpness and sizes etc. However, there exists some issues and problems faced with the CNN-based models. These models still show misdetection or less confident mapping of boundaries in the images captured with very dull surroundings. Therefore, we further propose a hybrid solution involving gamma correction, which is a conventional approach, followed by vehicle detection through CNN-based models. Since the previous approach showed significant results, the same pre-trained ghost free shadow removal model approach was applied on the images obtained after the gamma correction. The results have been compared with and without shadow removal incorporated. The block level description of the proposed approach is shown in Figure 8.

Using front view images, four different examples (see Figure 9) have been reported below which illustrates the classification and detection outcomes of the YOLOv5 technique before and after the shadow removal using Gamma correction method with pre-trained ghost free shadow removal model. It can be seen that the Gamma correction with 1.5 value has increased the contrast and enhancement of the image. It has detected and removed the shadows near the trees, cars and motorcycle. Then applying YOLO to the shadow removal images has improved the detection and classification results in comparison with the results on original images.

From example 1, it can be seen that the YOLO has detected a truck, which is a false positive result. Similarly, it has not detected the motorcycle as well. In contrast, the shadow

TABLE 10. Qualitative evaluation of the reviewed papers (High: H, Medium: M, Low: L).

| Paper	Category	Noise robustness	Computational complexity	Shadow sharpness	Illumination independence	Scene independence
Mikic et al. (2000) [22]	SP	M	H	L	L	M
Pei and Wang (2009) [28]		M	H	M	H	H
Huang and Chen (2009) [29]		H	H	M	H	M
Lin et al. (2010) [30]		H	H	M	H	M
Qi et al. (2014) [31]		M	H	M	H	M
Xiang et al. (2014) [36]		M	H	M	H	M
Russell et al. (2015) [37]		M	L	M	H	M
Valiere et al. (2015) [38]		M	-	L	L	L
Dai et al. (2016) [39]		M	L	M	H	M
Farou et al (2017) [40]		H	L	H	H	H
Garg et al. (2019) [41]		M	L	M	L	L
Shi and Liu (2019) [42]		H	H	M	H	M
Rittscher et al. (2000) [48]	SNP	M	M	M	M	M
Wang et al. (2008) [50]		L	L	L	L	L
Joshi and Papanikolopoulos (2008) [52]		M	M	M	H	M
Meher and Murty (2013) [55]		M	H	M	M	H
Gomes et al (2016) [57]		M	L	H	H	M
Wang et al. (2017) [58]		M	H	M	M	H
Zhu and Yin (2020) [61]		M	H	M	M	H
Sun and Li (2010) [71]	DM	H	M	L	M	M
Chacon-Murguia and Gonzalez-Duarte (2012) [72]		M	M	M	M	L
Yang et al. (2018) [73]		M	M	M	H	H
Stander et al. (1999) [76]	DNM	M	H	L	M	L
Cucchiara et al. (2003) [78]		M	L	M	H	H
Leone and Distante (2007) [80]		M	L	L	M	H
Vicente et al. (2016) [82]	SOTA	M	M	H	M	M
Li et al. (2018) [83]		M	M	H	M	M
Bakr et al. (2020) [84]		M	H	H	H	M
Wang et al. (2020) [85]		M	M	H	M	M
Chen et al. (2021) [86]		M	H	H	M	M
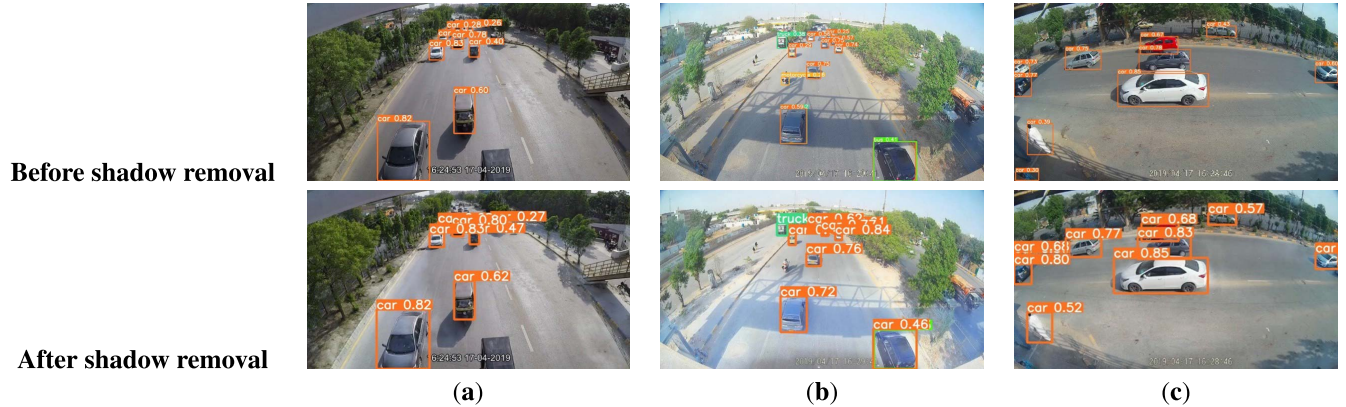
FIGURE 7. Analysis of before and after shadow removal images, (a) Frontal view, (b) Rear view and (c) Side view.

removal image, YOLO has not detected any false positives and successfully detected motorcycle in the image. Whereas in 2nd example, YOLO has detected 6 vehicles and has not detected the right motorcycle in original image; however in the shadow removal image, YOLO has successfully detected 7 vehicles and has detected the right motorcycle in the image. In 3rd and 4th example, YOLO has detected 5 and 9 vehicles, respectively, and has not detected the right motorcycle in original image; whereas in the shadow removal images, YOLO has successfully detected 6 and 10 vehicles, respectively, and has detected the right motorcycle in the image with better classification results.
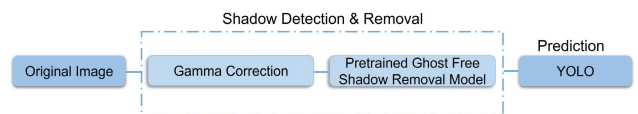

FIGURE 8. Proposed approach for shadow detection and removal.

In front view dataset, shadow removal with the help of Gamma correction has produced some decent results which has helped the YOLO model to increase its detection and classification accuracy.

FIGURE 9. Gamma correction and shadow removal, (a) Original image, (b) Prediction (YOLO) on original image, (c) Gamma correction based shadow removal and (d) Prediction (YOLO) after shadow removal.

VI. CONCLUSION AND FUTURE WORK

In traffic flow analysis of urban traffic video scenes, moving object detection is one of the common yet challenging tasks for vision-based algorithms. Moving cast shadows are a major concern for foreground detection algorithms. Therefore, finding the most suitable approach for the detection of shadows in this scenario is quite challenging.

In this review paper, contributions already made in this field are discussed for comparative evaluation of moving cast shadow detection methods. A total of 70 papers that contain results of urban traffic scenes have been shortlisted from the last three decades to give a broader review of the work done in this area. Following the approach of Prati *et al.* [5], existing techniques of moving cast shadow detection methods are categorized. The characteristics of cast shadows are defined and benchmark datasets used for the specific conditions related to traffic analysis are presented. The lack of standardization of urban traffic datasets is a critical point identified and accordingly needs to be worked on by the research community. The research also presented quantitative and qualitative analysis of the reviewed papers. By overall analysis of all techniques, it was concluded that state-of-the-art techniques performed much better than the other approaches in terms of performance and are recommended approach for removal of shadows. However, there is a trade-off between accuracy and high processing times.

This paper demonstrated the performance of vehicle detection after removal of shadows. A state-of-the-art GAN-based algorithm was used to remove shadows from the traffic camera images having varying shadow strengths. After that, a pre-trained YOLOv5 model was used to evaluate vehicle detection performance on the shadow removed images of different angles i.e., frontal, rear and side views. However, there were still detection issues, due to which this paper proposed a hybrid solution for that. The classical computer vision-based Gamma Correction technique was used in combination with GAN-based model for shadow removal before passing them to the deep learning model for vehicle detection. This proposed hybrid architecture indicated a good performance in removal of shadows and accordingly overall accuracy improvement of vehicles detection. The performance of hybrid solution clearly outperforms the simple shadow detection and removal-based vehicle detection approach.

As part of a future work, the use of transfer learning/fine-tuning techniques for YOLO model can be explored to produce more accurate results than the pre-trained models. The research work can also be extended by use of even more complex datasets with large shadows of stronger strengths. Similarly, advanced and efficient state-of-the-art algorithms can be developed which are equally comparable to conventional algorithms in terms of inference or execution times.

REFERENCES

[1] Z. Wang, J. Zhan, C. Duan, X. Guan, P. Lu, and K. Yang, "A review of vehicle detection techniques for intelligent vehicles," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jan. 5, 2022, doi: 10.1109/TNNLS.2021.3128968.

[2] M. Won, "Intelligent traffic monitoring systems for vehicle classification: A survey," *IEEE Access*, vol. 8, pp. 73340–73358, 2020.

[3] S. D. Khan and H. Ullah, "A survey of advances in vision-based vehicle re-identification," *Comput. Vis. Image Understand.*, vol. 182, pp. 50–63, May 2019.

[4] A. Prati, I. Mikic, C. Grana, and M. M. Trivedi, "Shadow detection algorithms for traffic flow analysis: A comparative study," in *Proc. IEEE Intell. Transp. Syst. (ITSC)*, Aug. 2001, pp. 340–345.

[5] A. Prati, I. Mikic, M. M. Trivedi, and R. Cucchiara, "Detecting moving shadows: Algorithms and evaluation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 7, pp. 918–923, Jul. 2003.

[6] J. Vertens, A. Valada, and W. Burgard, "SMSnet: Semantic motion segmentation using deep convolutional neural networks," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 582–589.

[7] M. Taha, H. H. Zayed, M. E. Khalifa, and T. Nazmy, "Moving shadow removal for multi-objects tracking in outdoor environments," *Int. J. Comput. Appl.*, vol. 97, no. 10, pp. 43–51, Jul. 2014.

[8] N. Al-Najdawi, H. E. Bez, J. Singhai, and E. A. Edirisinghe, "A survey of cast shadow detection algorithms," *Pattern Recognit. Lett.*, vol. 33, no. 6, pp. 752–764, 2012.

[9] A. Sanin, C. Sanderson, and B. C. Lovell, "Shadow detection: A survey and comparative evaluation of recent methods," *Pattern Recognit.*, vol. 45, no. 4, pp. 1684–1695, Apr. 2012.

[10] G. D. Finlayson, S. D. Hordley, and M. S. Drew, "Removing shadows from images," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2002, pp. 823–836.

[11] R. Guo, Q. Dai, and D. Hoiem, "Paired regions for shadow detection and removal," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 12, pp. 2956–2967, Dec. 2013.

[12] H. Gong and D. Cosker, "Interactive shadow removal and ground truth for variable scene categories," in *Proc. BMVC*. Princeton, NJ, USA: Citeseer, 2014, pp. 1–11.

[13] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[15] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, 2015, pp. 91–99.

[16] J. Ji, X. Jiang, and W. Sun, "Shadow detection using double-threshold pulse coupled neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 1971–1975.

[17] B. Min, H. Wei, W. Chen, and L. Yao, "Shadow detection in color aerial sensing images based on adaptive pulse coupled neural network," in *Proc. Int. Conf. Audio, Lang. Image Process.*, Jul. 2014, pp. 521–525.

[18] K. L. Chung, Y. R. Lin, and Y. H. Huang, "Efficient shadow detection of color aerial images based on successive thresholding scheme," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 671–682, Feb. 2009.

[19] V. J. D. Tsai, "Automatic shadow detection and radiometric restoration on digital aerial images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2003, pp. 732–733.

[20] Y. P. Guan, "Spatio-temporal motion-based foreground segmentation and shadow suppression," *IET Comput. Vis.*, vol. 4, no. 1, pp. 50–60, Mar. 2010.

[21] N. Friedman and S. Russell, "Image segmentation in video sequences: A probabilistic approach," 2013, *arXiv:1302.1539*.

[22] I. Mikic, P. C. Cosman, G. T. Kogut, and M. M. Trivedi, "Moving shadow and object detection in traffic scenes," in *Proc. 15th Int. Conf. Pattern Recognit. (ICPR)*, vol. 1, 2000, pp. 321–324.

[23] A. Bevilacqua, "Effective shadow detection in traffic monitoring applications," *J. World Soc. Comput. Graph. (WSCG)*, Czech Republic, Tech. Rep., 2003, pp. 57–64, vol. 11.

[24] F. Porikli and J. Thornton, "Shadow flow: A recursive method to learn moving cast shadows," in *Proc. 10th IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 1, Oct. 2005, pp. 891–898.

[25] A. J. Joshi, S. Atev, O. Masoud, and N. Papanikolopoulos, "Moving shadow detection with low- and mid-level reasoning," in *Proc. IEEE Int. Conf. Robot. Autom.*, Apr. 2007, pp. 4827–4832.

[26] Z. Liu, K. Huang, T. Tan, and L. Wang, "Cast shadow removal combining local and global features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.

[27] N. Martel-Brisson and A. Zaccarin, "Learning and removing cast shadows through a multidistribution approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 7, pp. 1133–1146, Jul. 2007.

[28] L. Pei and R. Wang, "Moving cast shadow detection based on PCA," in *Proc. 5th Int. Conf. Natural Comput.*, vol. 2, 2009, pp. 581–584.

[29] J.-B. Huang and C.-S. Chen, "Moving cast shadow detection using physics-based features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 2310–2317.

[30] C.-T. Lin, C.-T. Yang, Y.-W. Shou, and T.-K. Shen, "An efficient and robust moving shadow removal algorithm and its applications in ITS," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, pp. 1–19, Dec. 2010.

[31] M. Qi, J. Dai, Q. Zhang, and J. Kong, "Cascaded cast shadow detection method in surveillance scenes," *Optik*, vol. 125, no. 3, pp. 1396–1400, 2014.

[32] K. Jiang, A.-H. Li, Z.-G. Cui, T. Wang, and Y.-Z. Su, "Adaptive shadow detection using global texture and sampling deduction," *IET Comput. Vis.*, vol. 7, no. 2, pp. 115–122, 2013.

[33] K. Ouivirach and M. N. Dailey, "Extracting the object from the shadows: Maximum likelihood object/shadow discrimination," in *Proc. 10th Int. Conf. Electr. Eng./Electron., Comput., Telecommun. Inf. Technol.*, May 2013, pp. 1–5.

[34] A. Russell and J. J. Zou, "Moving shadow detection based on spatial–temporal constancy," in *Proc. 7th Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, Dec. 2013, pp. 1–6.

[35] M. Khare, R. K. Srivastava, and A. Khare, "Moving shadow detection and removal—A wavelet transform based approach," *IET Comput. Vis.*, vol. 8, no. 6, pp. 701–717, Dec. 2014.

[36] J. Xiang, H. Fan, H. Liao, J. Xu, W. Sun, and S. Yu, "Moving object detection and shadow removing under changing illumination condition," *Math. Problems Eng.*, vol. 2014, pp. 1–10, Feb. 2014.

[37] M. Russell, J. J. Zou, and G. Fang, "Real-time vehicle shadow detection," *Electron. Lett.*, vol. 51, no. 16, pp. 1253–1255, 2015.

[38] P. Valiere, L. Khoudour, A. Crouzil, and D. N. T. Cong, "Robust vehicle counting with severe shadows and occlusions," in *Proc. Int. Conf. Image Process. Theory, Tools Appl. (IPTA)*, Nov. 2015, pp. 99–104.

[39] J. Dai, Y. Li, J. Gao, and X. Wei, "Adaptive weighted moving shadow detection based on multiple features," Revista de la Facultad de Ingeniería UCV, Tech. Rep., 2016, pp. 68–80, vol. 31. [Online]. Available: https://oa.mg/work/10.21311/002.31.8.07

[40] B. Farou, H. Rouabhia, H. Seridi, and H. Akdag, "Novel approach for detection and removal of moving cast shadows based on RGB, HSV and YUV color spaces," *Comput. Informat.*, vol. 36, no. 4, pp. 837–856, 2017.

[41] K. Garg, N. Ramakrishnan, A. Prakash, S. Thambipillai, and P. Bhatt, "Rapid technique to eliminate moving shadows for accurate vehicle detection," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2019, pp. 1970–1978.

[42] H. Shi and C. Liu, "A new cast shadow detection method for traffic surveillance video analysis using color and statistical modeling," *Image Vis. Comput.*, vol. 94, Feb. 2020, Art. no. 103863.

[43] W. Sun, M. Sun, X. Zhang, and M. Li, "Moving vehicle detection and tracking based on optical flow method and immune particle filter under complex transportation environments," *Complexity*, vol. 2020, pp. 1–15, Apr. 2020.

[44] J. Zhang, X. Guo, C. Zhang, and P. Liu, "A vehicle detection and shadow elimination method based on greyscale information, edge information, and prior knowledge," *Comput. Electr. Eng.*, vol. 94, Sep. 2021, Art. no. 107366.

[45] H. Ghahremannezhad, H. Shi, and C. Liu, "A new online approach for moving cast shadow suppression in traffic videos," in *Proc. IEEE Int. Intell. Transp. Syst. Conf. (ITSC)*, Sep. 2021, pp. 3034–3039.

[46] T. Zhou, Y.-X. Li, M. Deng, and S.-Y. Pang, "Shadow suppression for vehicle target detection in open-air expressway scenes," in *Proc. 33rd Chin. Control Decis. Conf. (CCDC)*, May 2021, pp. 1229–1236.

[47] H. Shi, "Novel statistical modeling methods for traffic video analysis," Ph.D. dissertations, Digit. Commons New Jersey Inst. Technol., 2021.

[48] J. Rittscher, J. Kato, S. Joga, and A. Blake, "A probabilistic background model for tracking," in *Proc. Eur. Conf. Comput. Vis.* Berlin, Germany: Springer, 2000, pp. 336–350.

[49] K. Siala, M. Chakchouk, F. Chaieb, and O. Besbes, "Moving shadow detection with support vector domain description in the color ratios space," in *Proc. 17th Int. Conf. Pattern Recognit. (ICPR)*, vol. 4, 2004, pp. 384–387.

[50] Y. Wang, H.-D. Cheng, and J. Shan, "Detecting shadows of moving vehicles based on HMM," in *Proc. 19th Int. Conf. Pattern Recognit.*, 2008, pp. 1–4.

[51] N. Martel-Brisson and A. Zaccarin, "Kernel-based learning of cast shadows from a physical model of light sources and surfaces for low-level segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[52] A. J. Joshi and N. P. Papanikolopoulos, "Learning to detect moving shadows in dynamic environments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 11, pp. 2055–2063, Nov. 2008.

[53] M. Vargas, S. Toral, J. Milla, and F. Barrero, "A shadow removal algorithm for vehicle detection based on reflectance ratio and edge density," in *Proc. 13th IEEE Int. Conf. Intell. Transp. Syst.*, Sep. 2010, pp. 1123–1128.

[54] A. Amato, M. G. Mozerov, A. D. Bagdanov, and J. Gonzalez, "Accurate moving cast shadow suppression based on local color constancy detection," *IEEE Trans. Image Process.*, vol. 20, no. 10, pp. 2954–2966, Oct. 2011.

[55] S. K. Meher and M. N. Murty, "Efficient method of moving shadow detection and vehicle classification," *AEU, Int. J. Electron. Commun.*, vol. 67, no. 8, pp. 665–670, Aug. 2013.

[56] H. Wang and H. Zhang, "A hybrid method of vehicle detection based on computer vision for intelligent transportation system," *Int. J. Multimedia Ubiquitous Eng.*, vol. 9, no. 6, pp. 105–118, Jun. 2014.

[57] V. Gomes, P. Barcellos, and J. Scharcanski, "Stochastic shadow detection using a hypergraph partitioning approach," *Pattern Recognit.*, vol. 63, pp. 30–44, Mar. 2017.

[58] J. Wang, J. Han, E. Liu, and Z. Fan, "Traffic video vehicle shadow elimination algorithm based on principal component analysis," *DEStech Trans. Comput. Sci. Eng.*, Oct. 2017.

[59] X.-F. Yang and W.-C. Siu, "Vehicle detection under tough conditions using prioritized feature extraction with shadow recognition," in *Proc. 22nd Int. Conf. Digit. Signal Process. (DSP)*, Aug. 2017, pp. 1–5.

[60] Y. Yi, J. Dai, C. Wang, J. Hou, H. Zhang, Y. Liu, and J. Gao, "An effective framework using spatial correlation and extreme learning machine for moving cast shadow detection," *Appl. Sci.*, vol. 9, no. 23, p. 5042, Nov. 2019.

[61] T. Zhu and X. Yin, "Image shadow detection and removal in autonomous vehicle based on support vector machine," *Sensors Mater.*, vol. 32, no. 6, pp. 1969–1979, 2020.

[62] G. Kan and C. Wang, "Semi-supervised extreme learning machinebased method for moving cast shadow detection," in *Proc. IEEE 4th Adv. Inf. Manag., Communicates, Electron. Autom. Control Conf. (IMCEC)*, Jun. 2021, pp. 1083–1089.

[63] X. Lu, X. Mao, H. Liu, X. Meng, and L. Rai, "Event camera point cloud feature analysis and shadow removal for road traffic sensing," *IEEE Sensors J.*, vol. 22, no. 4, pp. 3358–3369, Feb. 2022.

[64] M. Anandhalli, A. Tanuja, V. P. Baligar, and P. Baligar, "Corner based statistical modelling in vehicle detection under various condition for traffic surveillance," *Multimedia Tools Appl.*, vol. 81, no. 20, pp. 28849–28874, Aug. 2022.

[65] S. Sahoo and P. K. Nanda, "Adaptive feature fusion and spatio-temporal background modeling in KDE framework for object detection and shadow removal," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 3, pp. 1103–1118, Mar. 2022.

[66] D. Roller, K. Daniilidis, and H. H. Nagel, "Model-based object tracking in monocular image sequences of road traffic scenes," *Int. J. Comput. Vis.*, vol. 10, no. 3, pp. 257–281, Jun. 1993.

[67] K. Onoguchi, "Shadow elimination method for moving object detection," in *Proc. 14th Int. Conf. Pattern Recognit.*, vol. 1, 1998, pp. 583–587.

[68] A. Yoneyama, C. H. Yeh, and C.-C.-J. Kuo, "Moving cast shadow elimination for robust vehicle extraction based on 2D joint vehicle/shadow models," in *Proc. IEEE Conf. Adv. Video Signal Based Surveill.*, Jul. 2003, pp. 229–236.

[69] E. Salvador, A. Cavallaro, and T. Ebrahimi, "Cast shadow segmentation using invariant color features," *Comput. Vis. Image Understand.*, vol. 95, no. 2, pp. 238–259, 2004.

[70] S. Nadimi and B. Bhanu, "Physical models for moving shadow and object detection in video," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 8, pp. 1079–1087, Aug. 2004.

[71] B. Sun and S. Li, "Moving cast shadow detection of vehicle using combined color models," in *Proc. Chin. Conf. Pattern Recognit. (CCPR)*, Oct. 2010, pp. 1–5.

[72] M. I. Chacon-Murguia and S. Gonzalez-Duarte, "An adaptive neural-fuzzy approach for object detection in dynamic backgrounds for surveillance systems," *IEEE Trans. Ind. Electron.*, vol. 59, no. 8, pp. 3286–3298, Aug. 2012.

[73] X. Yang, D. Zhou, D. Liu, and R. Yang, "Moving cast shadow detection using block nonnegative matrix factorization," *Bull. Polish Acad. Sci., Tech. Sci.*, vol. 66, no. 2, pp. 229-234, 2018.

[74] Y. Hu and W. Liu, "Shadow elimination based on multiple feature differences and GLVQ," in *Proc. 13th Int. Conf. Measuring Technol. Mechatronics Autom. (ICMTMA)*, Jan. 2021, pp. 872–877.

[75] N. Amamoto and A. Fujii, "Detecting obstructions and tracking moving objects by image processing technique," *Electron. Commun. Jpn. III, Fundam. Electron. Sci.*, vol. 82, no. 11, pp. 28–37, Nov. 1999.

[76] J. Stander, R. Mech, and J. Ostermann, "Detection of moving cast shadows for object segmentation," *IEEE Trans. Multimedia*, vol. 1, no. 1, pp. 65–76, Mar. 1999.

[77] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting objects, shadows and ghosts in video streams by exploiting color and motion information," in *Proc. 11th Int. Conf. Image Anal. Process.*, 2001, pp. 360–365.

[78] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting moving objects, ghosts, and shadows in video streams," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1337–1342, Oct. 2003.

[79] D. Toth, I. Stuke, A. Wagner, and T. Aach, "Detection of moving shadows using mean shift clustering and a significance test," in *Proc. 17th Int. Conf. Pattern Recognit. (ICPR)*, vol. 4, 2004, pp. 260–263.

[80] A. Leone and C. Distante, "Shadow detection for moving objects based on texture analysis," *Pattern Recognit.*, vol. 40, no. 4, pp. 1222–1233, Apr. 2007.

[81] S. Srividhya, C. Kavitha, W. Lai, V. Mani, and O. Khalaf, "A machine learning algorithm to automate vehicle classification and license plate detection," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–12, Jun. 2022.

[82] T. F. Y. Vicente, L. Hou, C.-P. Yu, M. Hoai, and D. Samaras, "Large-scale training of shadow detectors with noisily-annotated shadow examples," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, 2016, pp. 816–832.

[83] Y. Li, M. Sakamoto, T. Shinohara, and T. Satoh, "Object detection from MMS imagery using deep learning for generation of road orthophotos," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XLII-2, pp. 573–577, May 2018.

[84] H. Bakr, A. Hamad, and K. Amin, "Mask R-CNN for moving shadow detection and segmentation," *Int. J. Comput. Inf.*, vol. 8, no. 1, pp. 1–18, 2021.

[85] T. Wang, X. Hu, Q. Wang, P.-A. Heng, and C.-W. Fu, "Instance shadow detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1880–1889.

[86] Z. Chen, L. Wan, L. Zhu, J. Shen, H. Fu, W. Liu, and J. Qin, "Triple-cooperative video shadow detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 2715–2724.

[87] J. Cao, C. Peng, Y. Li, and S. Du, "A shadow detection method for retaining key objects in complex scenes," in *Proc. 13th Int. Conf. Knowl. Smart Technol. (KST)*, Jan. 2021, pp. 90–95.

[88] J. Bao, X. Zhang, T. Zhang, and X. Xu, "ShadowDeNet: A moving target shadow detection network for video SAR," *Remote Sens.*, vol. 14, no. 2, p. 320, Jan. 2022.

[89] X. Peng, X. Fan, Q. Wu, J. Zhao, and P. Gao, "Video-based smoky vehicle detection with a coarse-to-fine framework," 2022, *arXiv:2207.03708*.

[90] N. Arora, Y. Kumar, R. Karkra, and M. Kumar, "Automatic vehicle detection system in different environment conditions using fast R-CNN," *Multimedia Tools Appl.*, vol. 81, no. 13, pp. 18715–18735, May 2022.

[91] X. Cun, C.-M. Pun, and C. Shi, "Towards ghost-free shadow removal via dual hierarchical aggregation network and shadow matting GAN," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, no. 7, pp. 10680–10687.
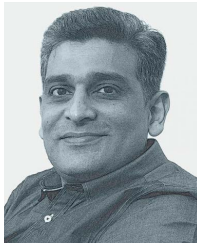
MUHAMMAD UMAIR ARIF received the B.E. degree in electrical engineering from the NED University of Engineering and Technology, Pakistan, in 2009, and the M.S. degree in fuzzy controls and the Ph.D. degree in computer vision from the National University of Sciences and Technology (NUST), Pakistan, in 2013 and 2022, respectively. His work is focused on computer vision solutions for intelligent transportation systems and their applications on embedded systems.

MUHAMMAD UMAR FAROOQ received the M.S. degree in electronics engineering from the Pakistan Air Force (PAF), Karachi Institute of Economics and Technology (KIET). He is currently working as a Lecturer with the Electronics and Power Engineering Department, Pakistan Navy Engineering College (PNEC), National University of Sciences and Technology (NUST). He is also working as a Project Manager with the Video Analytics Laboratory (Partner Laboratory of National Center of Big Data and Cloud Computing), PNEC, NUST. His research interests include artificial intelligence, computer vision, video analytics, medical imaging, and immersive technologies.

ZAIN UL ABIDEEN LODHI (Student Member, IEEE) received the B.E. degree in telecommunication engineering from the Institute of Space Technology, in 2012, and the M.S. degree in control systems from the National University of Sciences and Technology (NUST). He is currently pursuing the Ph.D. degree with the Harbin Institute of Technology (HIT), China. His work is focused on computer vision system and control of swarm robots.

RANA HAMMAD RAZA received the Ph.D. degree in electrical engineering from Michigan State University (MSU), USA, in 2013. He has been working in academia and industry for more than 25 years. He worked at the IBM T. J. Watson Research Center, New York, NY, USA. He is currently an Assistant Professor and the Founding Director of the Video Analytics Laboratory, Virtual Reality Center, Pakistan Navy Engineering College, National University of Sciences and Technology, Karachi. His research interests include artificial intelligence, video analytics, medical imaging, underwater acoustics, and immersive technologies. He was a recipient of the U.S. Fulbright Scholarship.

MUHAMMAD ABDUR REHMAN HASHMI (Graduate Student Member, IEEE) received the bachelor's degree in electronic engineering and the M.S. degree in electrical (communication) engineering from the Pakistan Navy Engineering College, National University of Sciences and Technology, Karachi, in 2007 and 2017, respectively. He is currently pursuing the Ph.D. degree in electronic and electrical engineering with the University College London (UCL), U.K. He has over ten years of experience working in industry. His research interests include RF/microwave antennas and circuit design, environmental sensor networks, underwater acoustics, image processing, and audio signal processing.

- - -