AN ENHANCED THREE-FRAME-DIFFERENCING APPROACH FOR VEHICLE DETECTION UNDER CHALLENGING ENVIRONMENTAL CONDITIONS

ZI YANG

PhD

The Hong Kong Polytechnic University

2019

The Hong Kong Polytechnic University

Department of Land Surveying and Geo-Informatics

An Enhanced Three-Frame-Differencing Approach for Vehicle Detection under Challenging Environmental Conditions

Zi YANG

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

July 2018

Certificate of Originality

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

____________________________________(Signed)

Zi YANG, Lynda (Name of student)

Abstract

Robust and efficient vehicle detection, counting and tracking is an important task in Intelligent Transportation Systems. With the continuous development of computer vision technologies, remarkable progress has been made in vision-based vehicle detection. Comparing to other sensors, vision cameras provide rich information for driving understanding. At the same time, robust feature descriptors and efficient background models have been proposed for the purpose of accurate vehicle detection. In this thesis, a computationally efficient method for vehicle detection, counting and tracking under different environmental conditions is presented, with a special focus on adverse illumination and weather. The general framework is based on enhanced Three-Frame-Differencing (E-TFD). In a given video sequence, three consecutive frames are utilized to generate frame differencing images. With an efficient thresholding and removal of small noise regions, moving vehicles can be extracted in an efficient and accurate manner. Meanwhile, based on extracted regions of interest (ROIs), exact numbers of vehicles can be counted and displayed on the screen. The E-TFD method can detect and count vehicle candidates in both fine and inclement weather conditions, including sunny, rainy, foggy, snowy, blizzard, wet snow and nighttime conditions in this study. To evaluate the E-TFD detection approach, nine videos are collected from different sources. Six videos are selected from two public datasets, CDnet

2014 and KIT dataset, for the purpose of performance analysis. At the same time, three videos are recorded from different roads in Kowloon, Hong Kong using a digital camera. Of all 4532 tested frames, 10059 vehicles can be successfully detected out of 11556 vehicles, showing an average detection rate of 87.1%. The E-FTD method shows a significant improvement of detection rate in adverse conditions and can provide a efficient solution of all-time, all-weather detection, counting and tracking that can in future be embedded into a real-time traffic surveillance system.

Acknowledgements

My deepest gratitude goes to my research supervisor Dr. Lilian PUN for continual guidance and stewardship of my graduate career. She is a devoted scholar, a knowledgeable advisor and a patient mentor. Her supervision helped me in all the time of research and writing of this thesis. I could not have imagined having a better supervisor for my PhD study.

Besides my supervisor, I would like to mention my appreciation to the entire Department of Land Surveying and Geo-Informatics at the Hong Kong Polytechnic University for their insistence on research excellence. I am also grateful to my fellow research group colleagues for their feedback, cooperation and of course friendship. Without their precious support, it would not be possible to conduct this research.

Lastly, I would like to thank my family for supporting me spiritually throughout writing this thesis and my life in general.

Table of Contents

List of Figures

List of Tables

CHAPTER 1 INTRODUCTION

1.1 Background

Intelligent Transport Systems (ITS) is a popular field of research in recent years. By providing innovative services relating to different modes of transport and traffic management and enabling various users to be better informed and make safer, more coordinated and 'smarter' use of transport networks (PARLIAMENT and UNION, 2010), ITS aims to improve transportation safety, mobility, productivity and environmental performance for traffic planners and road users. With continuous urban road development and extensive construction of expressways, increasing interest is devoted to vehicle detection. As an essential task in ITS, vehicle detection aims to provide information assisting vehicle counting, vehicle speed measurement, identification of traffic accidents, traffic flow prediction and so on.

The interest in Intelligent Transportation Systems (ITS) technologies aroused about 20 years ago with respect to the problem of people and goods mobility. The field of ITS is now entering its second phase characterized by maturity in its approaches and by new technological possibilities which allow the development of the first experimental product (Bertozzi et al., 2000).

Various sensors have been utilized to collect continuous-generated traffic information. Generally, the sensors can be categorized as hardware-based sensors and software-based sensors. Hardware-based sensors (also called active sensors), mainly include Lidar (Asvadi et al., 2017), Radar (Stevenson, 2011), and inductive-loop detectors (Ki and Baik, 2006). Designed for a special purpose, these detectors transmit and receive electromagnetic signals to measure traffic parameters. Hardware-based sensors have been widely used for vehicle detection in many earlier studies. The main drawbacks of hardware equipment are high maintenance cost and environmental effects. Software sensors mainly refer to cameras, which vary in position and type. Cameras can be either mounted on a vehicle or fixed along the roadside. At the same time, camera type can be 2D or 3D (contain depth information). Comparing to hardware detectors, vision cameras are more advantageous in terms of cost and flexibility. At the same time, they provide a rich contextual information for human visualization and understanding. With the increasing coverage of traffic surveillance cameras and easier accessibility of traffic image, vision-based vehicle detection is one of the most promising techniques for driving environment understanding. In recent years, there is even a trend to fuse data from different sensors for accurate and efficient vehicle detection.

Apparently, vision-based vehicle detection takes full advantage of computer vision technologies. The British Machine Vision Associate (BMVA) defines Computer Vision as “the automatic extraction, analysis and understanding of useful information from a single image or a sequence of images”(BMVA, 2016). In recent years, more and more studies have been focusing on deep learning technologies, contributing to blooming research to vehicle classification, tracking, speed measurement and vehicle type recognition.

Vision-based vehicle detection, counting and tracking have always been a significant research topic in Intelligent Transportation Systems. In fine illumination and weather conditions, many studies have provided comprehensive reviews and solutions to detect, count and track moving vehicles. However, when it comes to adverse conditions with low illumination and visibility, existing studies mainly focus on one or two conditions and discuss corresponding approaches. Till now, a complete understanding of all kinds of driving environment is yet to be achieved using traffic surveillance cameras.

1.2 Objectives

Vehicle detection systems should be able to handle problems posed by bad weather (Nayar and Narasimhan, 1999). Current studies are found to be limited in providing solutions to one or two challenging cases, such as vehicle detection in low illumination (dusk and nighttime) or adverse weather (fog, rain, and snow) conditions. Despite very high detection rate in some studies, these detection approaches lack robustness for implementation in all kinds of challenging environmental scenarios. Till now, a universal method for all-weather vehicle detection based on vision cameras has not been provided. The objectives of this study are:

1) To propose an enhanced Three-Frame-Differencing (E-TFD) based computationally efficient method to detect, count and track vehicles within the visible road segment from images of traffic surveillance cameras under varying environments, including sunny, rainy, snowy, foggy and nighttime images;

2) To evaluate the E-TFD detection approach in terms of accuracy and efficiency. True Positive Rate, False Positive Rate, Precision and Recall are used for performance analysis, followed by a comparison with other state-of-the-art methods.

1.3 Scope of the Study

This study focuses on vehicle detection and counting in challenging environments. Different image sequences are tested, including rainy, snowy, foggy, blizzard and nighttime conditions. The tested images sequences are characterized by the presence of raindrops, snowflakes, fog and walking pedestrian. To show the robustness of the enhanced TFD detection approach, sunny images have also been tested.

To evaluate the E-TFD method, both public datasets and self-collected video sequences of different sources of the recent two decades are used. Six videos are selected from two public-available datasets - CDnet 2014 and KIT Datasets. Besides, three videos are recorded from different roads in Kowloon, Hong Kong in rainy and nighttime conditions.

A successful vehicle detection approach is always in demand to monitor different traffic parameters (Mithun et al., 2016). In this study, ROIs of moving vehicles and exact counts of vehicle numbers in each frame are two major variables to be obtained. Regions of interest (ROIs) are obtained through the enhanced TFD approach. Based on ROIs detected in each frame, the number of vehicles can be obtained in real-time.

1.4 Methodology and Significance of the Study

In this study, we aim to develop a possible solution of vehicle detection under varying environments. Given a video sequence, moving vehicles in each frame are extracted using imaging technologies from visible road segments captured by traffic surveillance cameras. The detection approach should be robust to operate in different environmental conditions, yet fast enough to be embedded into a real-time traffic surveillance system.

General vehicle detection approaches can be categorized as appearance-based method and motion-based methods. This study mainly uses a motion-based method for vehicle detection, counting and tracking. The detection framework is based on Three-Frame-Differencing (TFD). Among three consecutive frames, pre-processing, frame differencing, fast thresholding, removal of small noise regions and morphological operations are major processing steps. Regions of Interest (ROIs) are generated using the aforementioned steps. Based on detection results, the exact numbers of vehicle can be displayed on the screen.

This study has provided a computationally-efficient method for vehicle detection, counting and tracking using traffic surveillance cameras, with a special focus on adverse conditions. Based on Three-Frame-Differencing, refinements are made in low illumination and adverse weather conditions. With a fast thresholding, removal of small noise regions and morphological operations, moving vehicles in various environmental conditions can be extracted efficiently. At the same time, counts in each frame can be obtained based on the presence of ROIs. The enhanced TFD method can also track vehicles in very inclement weather conditions.

To the best of our knowledge, vehicle detection in adverse conditions has only been preliminarily explored. Current studies mainly solve one or two challenging cases, such as rain, fog, dusk or nighttime. Till now, a universal approach for all time, all-weather traffic surveillance system has not been proposed. This study provides a possible solution for all weather vehicle detection based on vision cameras and can in future be embedded into a real-time traffic surveillance system. Finally, we hope to evolve this E-TFD detection approach into commercial software.

1.5 Thesis Layout

The remainder of the thesis is organized as follows. Chapter 2 reviews the previous work of vision-based vehicle detection methods. The comprehensive review categorizes existing detection approaches as appearance-based methods and motion-based methods, followed by solutions under varying environments and different traffic surveillance objects. Chapter 3 provides a detailed explanation of the proposed enhanced Three-Frame-Differencing (E-TFD) method for vehicle detection in challenging environmental conditions. Chapter 4 elaborates experimental results using public datasets as well as self-collected video data. At the same time, the detection and counting approach is first discussed in terms of accuracy and efficiency, then compared with other state-of-the-art approaches. Chapter 5 summarize the whole thesis and give some perspectives on future work.

CHAPTER 2 A COMPREHENSIVE REVIEW ON VEHICLE DETECTION UNDER VARYING ENVIRONMENTS

This section provides a comprehensive review of vision-based vehicle detection. The general detection approaches can be categorized as appearance-based methods and motion-based methods. Solutions under varying environments are reviewed in terms of illumination and weather. Then, several traffic surveillance objectives that can be achieved based on vehicle detection are summarized, such as solutions of major on-road driving problems and derivation of vehicle parameters.

In recent decades, many studies have presented detailed literature reviews related to vision-based vehicle detection, emphasizing different aspects of Intelligent Transportation Systems. Some focus on camera settings, such as multi-camera video surveillance (Wang, 2013) and on-road vehicle detection with car-mounted cameras (Sun et al., 2006b). Meanwhile, different fields of application are investigated, such as urban traffic analysis (Buch et al., 2011), driving behavior analysis (Sivaraman and Trivedi, 2013), vehicle detection techniques for Collision Avoidance Systems (CAS) (Mukhtar et al., 2015) and vehicle detection under varying environments (Yang and Pun-Cheng, 2018). These review papers have provided valuable information for further studies.

2.1 Appearance-based Methods

The appearance of vehicles varies in size, shape and color. Based on that, appearance-based methods employ prior knowledge to segment foreground (contains objects of interest) and the background (its complementary set) (Barnich and Van Droogenbroeck, 2011).

In traffic surveillance videos, front view and rear view are the two major fields of view (FOV). Accordingly, the rectangular shape of vehicles could be very useful to extract candidates in image scenes. Symmetry, edge (horizontal/vertical) and corner are the three very important cues to identify rectangular shapes. In other fields of view, other features may be more useful, such as shadow, color, texture, etc. With the increasing installation of traffic surveillance cameras on road, more FOV is presented, illustrated in Figure 2-1.

According to Sun et al. (2006b) and Sivaraman and Trivedi (2013), appearance-based methods often follow two basic steps: 1) hypothesis generation (HG), where locations of possible vehicles are hypothesized, and 2) hypothesis verification (HV), where tests are performed to verify the presence of vehicles in the traffic scene.

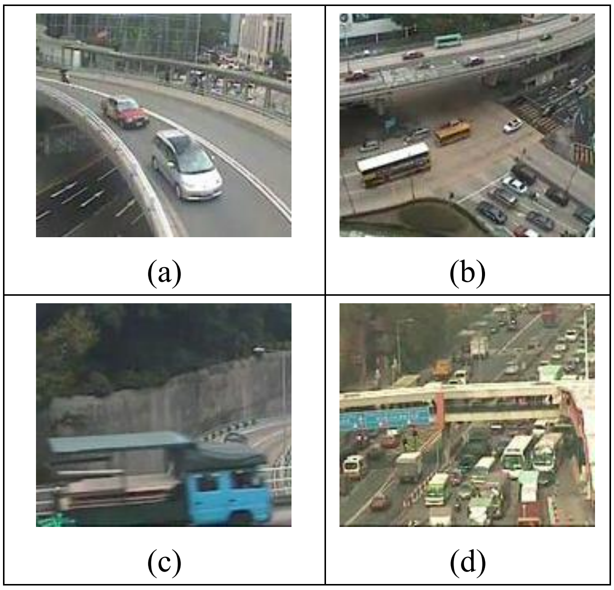


Figure 2-1. Vehicles in different field of view. (a) Front view; (b) Side view; (c) Side view; (d) Front and rear view.
Source: (Department, 2017).

2.1.1 Hypothesis Generation (HG)

In this step, possible vehicle candidates are extracted using one or several appearance features. The features can be simple global features to extract all related information from the image, or local feature descriptors to extract information on a regional level. Global features extract all related contextual information by considering all pixels in an image. These features are simple and efficient, yet of great essentiality in vehicular information extraction. In the field of vehicle detection, symmetry, edge, corner, color and shadow are mainly used and discussed as follows.

(a) *Symmetry:* Vehicles usually have the property of symmetry in front or rear view. In appearance-based methods, symmetry is often utilized by defining a

geometric model (Collado et al., 2004), finding symmetry axes (Liu et al., 2007) or center points (Teoh and Bräunl, 2012) to identify vehicle location in an image. At the same time, symmetry is often combined with other cues such as shadow and texture (Khairdoost et al., 2013), edge (Zielke et al., 1993) for appearance-based Hypothesis Generation.

The major limitation of symmetry feature lies in that symmetry information works only in front view. If the FOV changes, rectangular shape of vehicles might not be very obvious in the image. At the same time, the presence of occlusion makes the use of symmetry more difficult. Now, the symmetry property becomes a less-used feature in state-of-the-art vehicle detection methods.

(b) Edge: Horizontal and vertical edges of vehicles are important sources of contour information, especially in the front and rear view. Due to its low computational complexity, edge information can be extracted in real-time. Classic edge-based vehicle detection employ Sobel (1990), Canny (1986), or Prewitt (1970) operators to generate horizontal and vertical edge map. At the same time, Gaussian-based filter can extract edge information effectively. In related studies, edge feature can be either used as a single feature (Mu et al., 2016) or combined with other features, such as shadow (Chong et al., 2013) to locate target vehicles. Edge is an essential part of feature fusion as it provides contour information, while

other features provide contextual information for vehicle extraction.

(c) Corner: The conspicuous corners of vehicles play an important role in vison-based vehicle detection. It is a fact that a vehicle in front/rear view has four corners, the upper-left, upper-right, lower-left and lower-right. In computer vision, several corner detectors have been proposed, such as the Harris Corner Detector (Harris and Stephens, 1988). To detect rectangles in an image, Bertozzi et al. (1997) used template matching to locate the four corresponding corners. Alonso et al. (2007) separated vehicles from non-vehicles based on a simplified vehicle model characterized by corners, symmetry, and shadows.

(d) Shadow: The presence of vehicles is often followed by the shadow it casts. Shadow is defined as a dark area where light is blocked by an opaque object. As one kind of local illumination changes, the shadow can cause vehicle merging, shape distortion and losses (Prati et al., 2003), but could be very useful for vehicle identification. In related studies, shadow regions can be identified and removed by building a color model based on contrast (Asaidi et al., 2014), brightness (Horprasert et al., 1999), mean and variance of all color components (Mikic et al., 2000). Shadow can also be used as the only global feature (Yan et al., 2016) for real-time Hypothesis Generation.

(e) Color: Color provides rich information on images based on different color

space. In vehicle detection, color has been utilized to detect visual features such as vehicle lights (Chen et al., 2012) and license plates (Abolghasemi and Ahmadyfard, 2009). Traffic surveillance cameras often operate with the RGB model, but the three channels are highly correlated and individual value of red, green, blue depends on brightness strongly(Yang et al., 2011). To solve this problem, conversion from RGB color space to HSV (O'Malley et al., 2010), YCbCr (Wang et al., 2016), L*a*b* (Cabani et al., 2005) is operated to highlight red and white color and reduce the effect of illumination changes. Despite commonly-used color space, some studies even proposed some new color transformation models (Tsai et al., 2007) to identify specific vehicle color from images. Table 2-1 shows how color information is used in vehicle lights detection.

Extraction of global features is fast and efficient, but the main drawback is that any one feature cannot extract all appearance information of vehicles. Meanwhile, some unrelated information is extracted, such as moving pedestrian, waving trees and varying background. To solve this problem, one solution is to detect distinguishable features of vehicle parts, such as number plates and vehicle lights, based on with ROI is located using color, edge, blob detectors, etc. Another solution is to fuse appearance features to detect the whole vehicle. With a combination of two or more than two features, contextual and contour information of vehicles can be extracted more effectively. A brief summary of the feature fusion approaches in vehicle detection can be found in Table 2-4, Section 2.

Table 2-1. The use of color in vehicle lights detection.

Reference	Color Space	Corresponding component in RGB color space	Field of Application
Thammakaroon and Tangamchit (2009)	RGB	Red	Tail-light detection at night
O'Malley et al. (2010)	HSV	Red	Tail-light detection at night
O'Malley et al. (2008)	HSV	Red and White	Tail-light detection at night
Chen and Peng (2012)	YCbCr	Red	Tail-light detection at night
Wang et al. (2016)	YCbCr	Red	Brake-light detection at daytime
Skodras et al. (2012)	L*a*b*	Red	Tail-light detection at daytime/adverse weather conditions
Chen et al. (2016)	L*a*b*	Red	Brake light detection at daytime

In recent studies, there has been a transition from simple global features to robust local feature descriptors. In general, these descriptors are designed for a specific purpose and allow a quick search of objects in an image. Ideally, descriptors should be able to deal with various objects and robust to varying background, but also be invariant to geometric and photometric transformation (Li and Allinson, 2008). A number of appearance features have been proposed in the literature, among which two descriptors, the Histograms of Gradients (HOG) and Haar-like features, show increasing prevalence in modern vehicle detection.

a) *HOG*: Histogram of Gradients (Dalal and Triggs, 2005) is a feature descriptor designed for object recognition. By capturing the gradient structure that is very characteristic of local shape, HOG was first applied in pedestrian detection, and then expanded to the field of vehicle detection. Based on the original HOG descriptor, several adjustments have been made to suit different scenarios. Yan et al. (2016) used a combination of two HOG vectors for real-time detection and classification of the front, left, right, distant vehicles. Arróspide et al. (2013) proposed three cost-effective configurations of HOG, namely horizontal (H-HOG), vertical(V-HOG) and concentric rectangular (CR-HOG) respectively to improve the computation efficiency. In Khairdoost et al. (2013), Pyramid Histograms of Oriented Gradients (PHOG) features were extracted as basic features for front/rear vehicle detection. Principle component analysis (PCA) was then applied for dimension reduction.

b) *Haar-like features*: Haar-like features was first proposed by Viola and Jones (2001) to detect human faces. The descriptor uses integral images to represent the characteristics of ROI instead of using image pixels. The computation of Haar-like features is easy and simple, and this descriptor has been used in many vehicle detection studies for feature representation. Haar-like features have been utilized to represent vehicle edges and structures (Wen et al., 2015). Haselhoff and Kummert (2009) proposed a 2D triangle filer based on Haar-like features to detect vehicle and a general learning framework was provided by Sivaraman and Trivedi (2010). The computation of Haar-like features is fast, but the dimension of the feature vector generated from images is high (Wang et al., 2015). Therefore, an operation of dimension reduction is applied to decrease hardware storage, such as Non-negative Matrix Factorization (Wang et al., 2015).

c) Other Local Feature Descriptors: Despite the HOG and Haar-like features, some other appearance features also show remarkable performance in collecting contextual information from a regional level. The descriptors include, but are not limited to: Gabor features (Tao et al., 2007), Speed-up Robust Features (Bay et al., 2006), and Scale-Invariant Feature Transformation features(Lowe, 1999).

Gabor features have been used to extract taillights (Ming and Jo, 2011) with 5 different scales and 8 orientations. Sun et al. (2006a) used Haar wavelet for vehicle detection in the rear view. Lin et al. (2012) used SURF point detector to capture the wheel on the rear half of the vehicle. Zhang et al. (2011) used the SIFT point detector to extract interest points in an image, which help to find the bounding box of a vehicle.

2.1.2 Hypothesis Verification (HV)

After locating possible vehicle candidates from HG, the next step is to verify the correctness of the hypothesis. Hypothesis Verification (HV) in machine learning is treated as a two-class pattern classification problem: vehicle versus non-vehicle. The verification step can be categorized as templated-based methods and appearance-based methods.

(a) Template-based HV: Template-based methods extract morphological characteristics of a vehicle and compare them with a predefined model or template (Parodi and Piccioli, 1995). Some of the templates represent the vehicle in a ‘vague’ form (Ito et al., 1995), while others are more detailed.

Based on the observation that rear/front view a vehicle has a ‘U’ shape, templates with this U-shape is a commonly-used for Hypothesis Verification. Richter et al. (2008) used a U-shape contour chain to identify vehicle image. For construction vehicle detection, Ji et al. (2016) defined inverse-v feature template with specific angle ranges. To detect vehicles at daytime and nighttime, Cucchiara and Piccardi (1999) used moving edge closure as daytime templates, and headlights pairing as nighttime templates.

The main limitation of templated-based methods is the lack or robustness and universality. Due to the particularity of the predefined shape type, the template/model might not work if the FOV changes. Therefore, template-based Hypothesis Verification is not a popular method in state-of-the-art vehicle detection.

(b) Appearance-based HV: Appearance-based HV methods, on the other hand, collect positive image samples (the vehicle class) and negative image samples (the non-vehicle class) and train them with classifiers. In all training samples, each image is represented on or several appearance features. Classifiers learn the characteristics of vehicle appearance in a statistical way and draw a decision boundary between the vehicle and non-vehicle class. In order to achieve the optimum performance, huge intra-class variability should be extracted.

Appearance-based Hypothesis Verification mainly relies on local feature descriptors mentioned in the HG part. Support Vector Machine (SVM), Adaboost and Neural Networks (NN) are three representative classifiers. SVM (Cortes and Vapnik, 1995) is a discriminative classifier that constructs a hyperplane and learns the decision boundary between two classes. Adaboost (Freund and Schapire, 1997), on the other hand, is a generative weak classifier that improves the performance of a simple classifier by combining local feature descriptors (Buch et al., 2011). Neural Networks (NN) (Bishop, 1995) has been a popular classifier in the past decade, which learns highly nonlinear decision boundaries (Sun et al., 2006a).

However, Neural Networks suffers from the computation of parameter tuning and is time-consuming. With the blooming of deep learning technologies, Convolutional Neural Networks(CNN) is widely used in vision-based vehicle detection. Fast R-CNN (Fan et al., 2016) was used for vehicle detection, where a region proposal network (RPN) significantly reduced the proposed cost. Zhou et al. (2016) proposed a shallow fully convolutional network called fast vehicle proposal network (FVPN) to localize all vehicle-like objects in real-time.

Recently, some specific feature-classifier pairs have been widely used in appearance-based HG and HV. The combinations of HOG features and SVM classifiers, and Haar-like features together with Adaboost classifiers are mostly used pairs. Based on existing algorithms, comparative studies have been carried out between features, classifiers and feature-classifier combinations. Negri et al. (2008) compared HOG and Haar-like features and utilized the features to construct a cascade of boosted classifiers for rear-view vehicle detection. SVMs and NNs were compared by Sun et al. (2006a) as appearance-based HV approaches in both simple and complex traffic scene. The HOG-SVM and Haar-Adaboost combinations were studied by Sivaraman and Trivedi (2014) as an active learning framework.

Table 2-2. Representative feature-classifier combinations for appearance-based Hypothesis Verification.

Features	Classifiers	Field of view
HOG	SVM (Sivaraman and Trivedi, 2014, Cheon et al., 2012)	Rear view
	Adaboost (Yan et al., 2016)	Rear view
Haar-like	Artificial Neural Network (Naba et al., 2016)	Front/rear view
	Adaboost (Wen et al., 2015)	Multiple
Gabor	Back propagation neural network (Ming and Jo, 2011)	Rear view
	SVM (Sun et al., 2006a)	Front/rear view

It can be seen from Table 2-2 that nearly all features depict pixels in terms of orientation information, Gabor features provide the scale and orientation information, HOG features calculate gradient magnitude and orientation to construct the histogram, Haar-like features utilize rectangle filters to extract orientation information. At the same time, most combinations detect and classify vehicles based on rear view, only a small portion of detection is from the side view. Symmetry information could be very useful in the front/rear view, and edge features have been utilized in many studies to extract the symmetric appearance of moving vehicles. Both HOG-SVM and Haar-Adaboost combinations perform well in the rear-view image. Being sensitive to the edge and symmetric structures, Haar-like features, combined with Adaboost classifiers, can perform a rapid detection performance. Different from the Haar-Adaboost combination, HOG features focus more on extracting orientations of edges. Therefore, the field of view can be extended from the traditional front/rear view to multi-view vehicle detection. Yan et al. (2016) used HOG descriptors with different structures to classify left, right, distant and front vehicles. Another study (Rybski et al., 2010) used HOG features to classify four different orientations of vehicles.

2.2 Motion-based Methods

Vehicles on the road are typically in motion, introducing effects of ego and relative motion (Sivaraman and Trivedi, 2013). Without any prior knowledge, these methods mainly extract vehicles based on the inter-frame motion that differentiate from the background. Motion-based methods are often based on successive image sequence, while appearance-based methods perform either in a single image or consecutive sequences.

2.2.1 Inter-Frame-Differencing

Inter-Frame-Differencing is a very efficient frame-based method to extract moving objects in consecutive image sequences. Conventional two-frame-differencing method (Jain and Nagel, 1979) observes the difference of two successive image frames and generates first-order difference picture (FODP) by subtracting the intensity between the previous frame I_{n-1} and current frame I_n. Thresholding is then applied to segment the binary image, generating image background (do not contain vehicles) and foreground (contain vehicles). Frames selected for differencing can be consecutive (Celik and Kusetogullari, 2010) or inconsecutive (Ji et al., 2016). Inter-frame differencing can generate binary image map in a real-time manner, but the main drawbacks are: 1) the performance of the method is susceptible to illumination changes;2) the presence of holes in binary image when vehicles move slowly and 3) the ghost behind the moving vehicle (Zhang et al., 2010).

However, when vehicles move fast in video frames, contour information generated by two successive frames may overlap. In order to solve this problem, three-frame differencing (Weng et al., 2010) have been proposed. Based on three successive frames, two difference images can be obtained. Operating also in real-time, three-frame-differencing achieves better performance than two-frame methods. In recent studies, three-frame-differencing has been operated with different thresholding values and kernel sizes. Morphological operations of “difference”, “dilate”, “and” (Xia et al., 2015) and “xor” (Lan et al., 2014) have been used to fill in the holes and connect the discontinuous edges in difference images.

The Inter-Frame-Differencing method itself may have some limitations in on-road vehicle detection. Therefore, it serves as the first step in many methods. Celik and Kusetogullari (2010) performed background subtraction based on results of inter-frame-differencing to detect moving pixels. Frame differencing was used in conjunction with a special-designed feature descriptor(Ji et al., 2016) to recognize part of a vehicle, and with the Gaussian Mixture Model (Fu et al., 2016) to get a better foreground image in crowded traffic scenes.

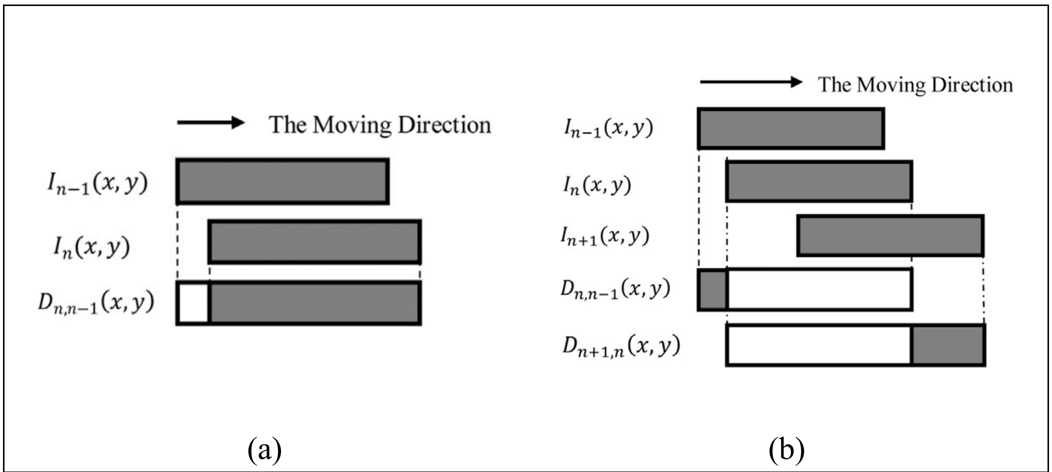


Figure 2-2. Illustration of Two-Frame-Differencing (a) and Three-Frame-Differencing (b).
Source: (Xia et al., 2015).

2.2.2 *Background Modelling*

According to Barnich and Van Droogenbroeck (2011), background modelling should be able to deal with three problems: 1) what is the model and how does it behave? 2) how is the model initialized? 3) how is the model updated over time? The general background modeling algorithms follow the procedure of background initialization, foreground detection and background maintenance.

(a) Gaussian Mixture Model: The original Gaussian Mixture Model (GMM) was proposed by Stauffer and Grimson (1999). Being one of the most widely-used background models for moving objects detection, this method assumes that all data points are generated from a mixture of finite Gaussian distributions with unknown parameters. The GMM approach models each pixel using multiple, adaptive Gaussians and uses an on-line approximation to update the model. Two parameters are required in this model, the learning rate (α) and threshold (T). In the model matching procedure, GMM requires three Gaussian possibility operations and three comparison operations. In the modeling updating procedure, GMM requires one Gaussian possibility operation, four subtraction operations and eight multiplication operations (Zhang et al., 2007).

With regard to the Gaussian Mixture Model, some improvements have been proposed in vehicle detection literature for better handling of various scenarios to make it more efficient. An adaptive improved GMM was proposed by Zivkovic (2004). By automatically selecting the necessary number of components per pixel, this new approach could reduce processing time and improve segmentation performance. Another improvement was made by Varadarajan et al. (2015). By extending traditional pixel-based mixture modelling approaches over neighborhood regions, the proposed generalized framework called Region-based Mixture of Gaussian (RMoG) could reduce false positives effectively. In Xia et al. (2016)'s study, the Gaussian Mixture Model was fused with the expectation-maximization (EM) algorithm to improve the segmentation quality of moving vehicles.

(b) Codebook: Codebook (Kim et al., 2004) is another typical background modelling approach. Each pixel is represented by a codebook. Sample background values at each pixel are quantized into codebooks which represent a compressed form of background model for a long image sequence. Five key parameters are involved in this model, four thresholds and one learning rate (Shah et al., 2015). Codebook of relevant light features was utilized together with Kalman filter to detect and track vehicle turning signals and brake lights (Almagambetov et al., 2012).

2.2.3 Optical Flow

Optical flow is a typical motion-based method for vehicle detection. Zhan and Ji (2011) proposed an algorithm using pyramidal optical flow estimation and a morphological transformation to extract vehicle targets. A tracking processed based on optical flow was applied to reduce the complexity of computing (Kuo et al., 2011). Liu et al. (2013) used optical flow to get the moving direction of the vehicle in initial frames. Then, a distant factor and an accurate feature template were used to track the on-road vehicle in low-resolution videos. Batavia et al. (1997) calculated implicit optical flow to detect overtaking on highways.

2.2.4 Other Methods

Despite the aforementioned approaches, some other motion-based models show remarkable performance in moving vehicle detection. Considering the motion property of dynamic textures in image background is quite different from moving vehicles, Zhang et al. (2010) used motion histogram to segment vehicles from the dynamic background. Jazayeri et al. (2011) built a vehicle motion model according to the scene characteristics. Parameters include average road width, distance to the target, the standard deviation of target distance, etc. Experiments were implemented in both daytime and nighttime, showing an 86.6% accuracy.

Motion-based methods not only play a significant part in detecting vehicles but also help to reduce the impact of varying backgrounds in real-world video images, such as swaying trees and flag fluttering (Zhang et al., 2007). These objects often exhibit certain stationary properties in time (Doretto et al., 2003), and sometimes change from quasi-periodic to irregular motion (Shah et al., 2015).

As the most common dynamic textures, swaying trees along the roadside are often captured by video cameras. Zhang et al. (2010) used motion histogram to select a waving tree region as the dynamic background. The swaying tree problems were also studied by Seki et al. (2003) using the correlation between neighboring image blocks.

In a word, appearance-based methods are easy and fast, but the main drawback is its high dependence on prior knowledge of vehicles, such as local/global features and appearance-based templates. Features like symmetry and edge are very limited due to the field of view. In the case of low illumination and adverse weather, modifications are needed to make the methods work. Motion-based methods could be time consuming, but less affected by varying illumination and dynamic background. Therefore, we need specific solutions to solve all-kinds of on-road vehicle detection problems and achieve satisfactory performance.

2.3 Solutions under Varying Environments

Section 2.1 and 2.2 summarized the general approaches of vehicle detection using traffic surveillance cameras. In real-world scenarios, there are many environmental factors that control the image background. This section summarizes vehicle detection methods under varying environments, where illumination and weather are two major concerns.

2.3.1 Illumination

Video cameras provide rich contextual information through measuring ambient light in real world. An image's exposure determines how light or dark an image will appear when it has been captured by the camera (Martínez et al., 2014). Still a challenging problem in vision-based vehicle detection, the illumination issue has been carried out in many studies (Cucchiara and Piccardi, 1999, O'malley et al., 2011). In urban or highway environments, traffic surveillance cameras are placed where illumination conditions vary through time (López-Rubio and López-Rubio, 2015) and the change of which can be global or local. Global illumination changes refer to scenes such as different weather, the daytime/nighttime conditions. Local illumination changes, on the other hand, refer to shadows or highlights of moving objects in the scene.

At daytime, illumination of traffic images is clear, therefore, multiple appearance-based features can be utilized to extract possible vehicles. However, illumination changes drastically during the transition of dawn and dusk. At nighttime, illumination becomes extremely low. Ambient light, as an uncontrolled environmental factor, adds extra difficulty in identifying possible vehicle candidates under low light conditions.

a) *Low-lighting Conditions:* In low-illumination images, contour and contextual information of vehicles might be visible, but exhibit low levels of saliency and contrast. To solve this problem, a simple solution would be using specific detectors for specific illumination scenarios, e.g. a day time detector focusing on texture information and a nighttime detector extracting tail-light information (Acunzo et al., 2007). It is very difficult to develop a universal approach for vehicle identification under varying lighting conditions. For this reason, many studies focus on one or two illumination conditions and propose detection approaches based on corresponding cases.

Figure 2-3. Traffic surveillance images captured during dusk (left) and night (right).
Source: Schamm et al. (2010).

The low light condition was also mentioned by Acunzo et al. (2007). Four categories of lighting conditions (daylight, low light, night and saturation) were identified using a histogram-based clustering algorithm. Classifiers trained with AdaBoost were used for low light categories. Experiments showed a considerable improvement using the context-adaptive scheme.

To deal with the low illumination issue, efforts not only focus on the arithmetic improvement but also the usage of different optical sensors. Low-light camera, as a useful tool to capture images under low illumination conditions, has been deployed in many studies for vision-based vehicle detection. According to Sun et al. (2006a), low-light camera provides a wider dynamic range than a normal camera. By setting a low static camera exposure value (O'malley et al., 2011), it can be ensured that headlamps appear in images as separate, distinct regions. Then, multiple vehicles can be detected and tracked in low-light conditions. In another study (Eum and Jung, 2013), low-exposure (LE) images were integrated with auto-exposure (AE) images for headlights and taillights detection.

Figure 2-4. Normal exposure image (left) and low exposure image (right). Source: O'malley et al. (2011).

b) Nighttime: Vehicle detection at nighttime is a very challenging task in vision-based vehicle detection, as vehicle appearance might not be very obvious in low illumination conditions. The only salient visual features are headlights, rear lights and their beams (Cucchiara et al., 2000). Therefore, vehicle lights have been used as the major local feature for nighttime vehicle identification. Based on Nakagami-m distribution, brake lights were modelled and then detected in a part-based manner by analyzing the signal in both spatial and frequency domain (Chen and Peng, 2012). Rainy images at nighttime have been tested in this study, with a 71% detection rate.

Figure 2-5. Vehicles at nighttime.

To further study illumination changes and verify the robustness of the proposed algorithms, daytime and nighttime have been tested together in many studies. Cucchiara et al. (2000) performed a difference on three consecutive frames to detect moving vehicles at daytime. Headlights were extracted as the main features for nighttime vehicle detection. Shadows underneath vehicles were used by Iwasaki and Kurogi (2007) for side-view daytime and nighttime vehicle detection.

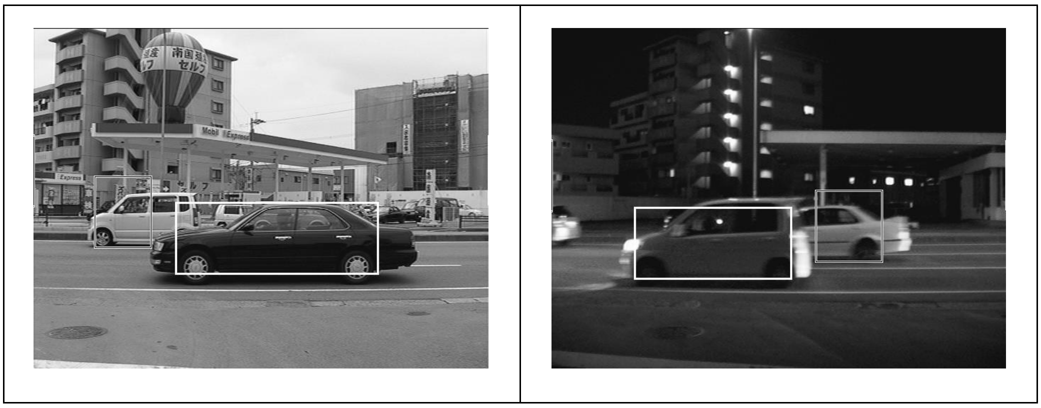


Figure 2-6. Vehicle detection using the same algorithm at daytime (left) and nighttime (right).
Source: Iwasaki and Kurogi (2007)

c) Special Illumination Conditions: Despite low-illumination and night vision, some studies carried out experiments on driving environments with special illumination, such as tunnels. Chan et al. (2007) categorized vehicle detection in tunnels as special lighting conditions. The overexposure due to high contrast makes vertical edges blurred and even disappear in images. In another study (Semertzidis et al., 2010), vehicles were detected in tunnels using stereo vision technologies. Betke et al. (2000) considered tunnels as reduced visibility condition, vehicles were detected and tracked using a combination of feature and motion information.

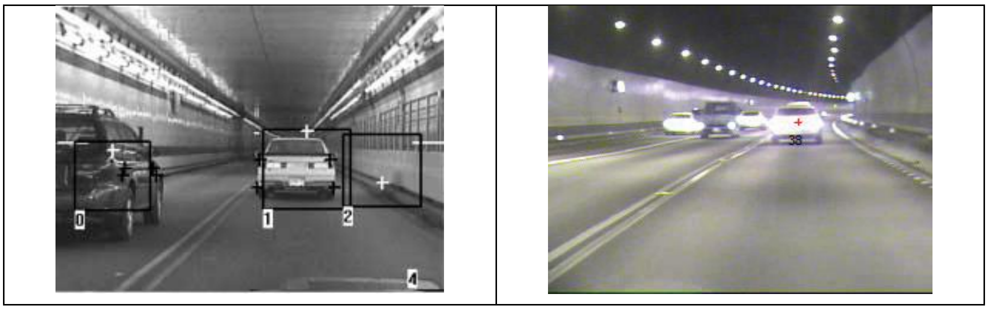


Figure 2-7. Vehicle detection results in tunnels.
Left: Betke et al. (2000); Right: Chan et al. (2007).

2.3.2 Weather

Ideally, computer vision systems are designed to perform in clear weather. However, outdoor scenes are not always perfect. In traffic surveillance videos, there are many environmental factors that control the image background. In bad weather conditions, the light reaching a camera is severely scattered by small particles in the atmosphere (Narasimhan and Nayar, 2003). Accordingly, vehicles captured by cameras exhibit different levels of vagueness, such as darkness, blurring and partial occlusion. Feature descriptors and background models that work in fine weather need to be modified to fit low illumination and varying background.

According to Padmini and Shankar (2016), weather can be categorized as static and dynamic. Static weather refers to the case of fog. Dynamic weather, on the other hand, refers to the case of rain and snow. From a brief summary of different weather and corresponding particle information, it is noticeable from Table 2-3 that raindrops have the largest radius, and haze has the largest concentration.

Table 2-3. Weather conditions and associated particle types.
Source: McCartney (1976)

Condition	**Particle Type**	**Radius (μm)**	**Concentration (cm^{-3})**
Air	Molecule	10^{-4}	10^{19}
Haze	Aerosol	10^{-2} - 1	10^{3} - 10
Fog	Water Droplet	1 - 10	100 - 10
Cloud	Water Droplet	1 - 10	300 - 10
Rain	Water Drop	10^{2} - 10^{4}	10^{-2} – 10^{-5}

An all-time, all-weather traffic surveillance system must include mechanisms that enable users to function in the presence of haze, fog, rain and blizzard (Nayar and Narasimhan, 1999). In recent decades, more studies have conducted on-road experiments under adverse weather conditions to verify the robustness of proposed algorithms. Besides, vehicle detection in inclement weather conditions has been studied as individual cases, where the removal of raindrops and fog in the image are two major concerns.

A rather simple case for bad-weather conditions is to detect vehicles on rainy days. Based on the rear view, rainy scene detection was carried out by Sun et al. (2006a) in many studies, daytime or nighttime (Chen and Peng, 2012).

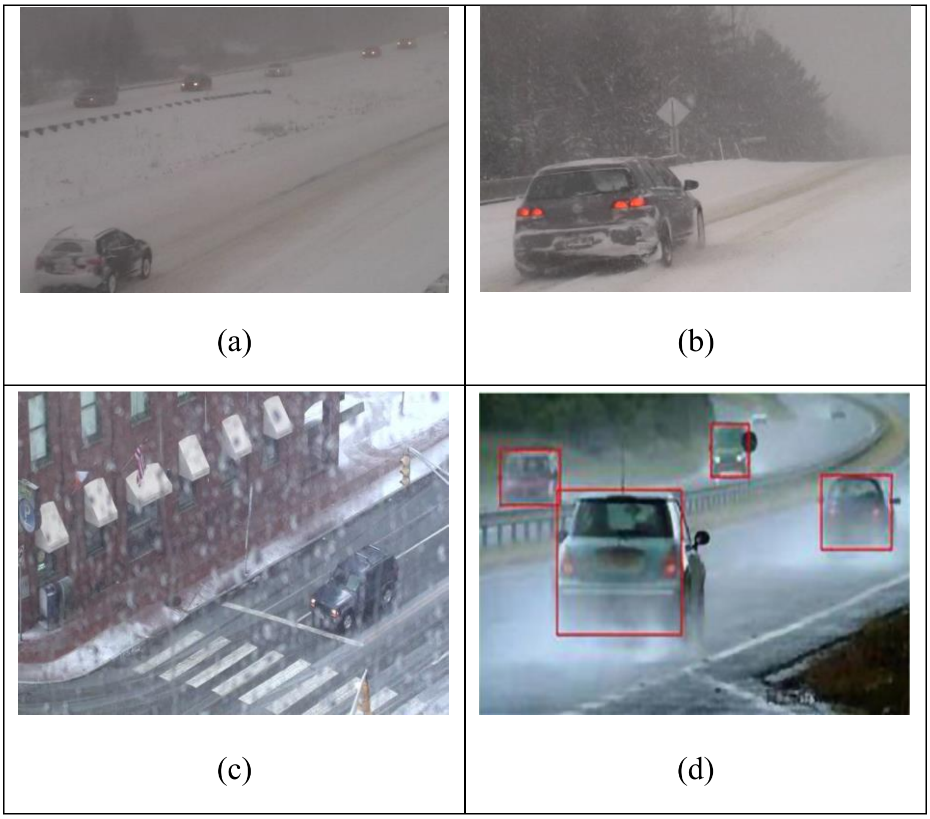


Figure 2-8. Traffic images under different weather conditions. (a) Vehicles in the blizzard; (b) Vehicle in snowfall; (c) Vehicle in wet snow; (d) Vehicles in rain. Source: (Wang et al., 2014, Jia et al., 2016)

(a) Fog: To detect vehicles in the foggy image, the first step is to reduce the impact caused by fog. The presence of fog has a global effect on images, making contrast decreases drastically. Studies on single image fog detection (He et al., 2011, Tan, 2008) have made remarkable progress. By observing the relationship between visibility and contrast, these methods can achieve very high accuracy. However, the main drawback of single image haze removal is that these algorithms need complicated computation and cannot be performed in real-time. Therefore, single image dehaze methods cannot be embedded into a real-time traffic surveillance system.

In consecutive image frames, haze removal can be operated either in the time-space domain or the time-frequency domain. In time-space domain, methods focus on extracting moving pixels based on robust global/local features, such as edge, line (Bronte et al., 2009), vehicle ego lights (Gallen et al., 2011). In the time-frequency domain, wavelet transform (Busch and Debes, 1998) was used to analyze fog availability. Pavlic et al. (2013) performed Fourier transform to extract spectral features and used a simple linear classifier to distinguish fog scenes and fog-free scenes.

(b) Rain and Snow: Rain and snow are categorized as dynamic weather. Accordingly, raindrops and snowflakes in outdoor scenes are often wrongly recognized as moving objects. To detect vehicles in the case of rain or snow, a step of raindrop/snowflake removal need to be applied before candidate extraction. In general, raindrops and snowflakes are detected using appearance-based methods.

Huang et al. (2012) utilized the matching of gray intensity feature, appearance feature and temporal feature to identify raindrops. Raindrops and snowflakes removal were studied by Kim et al. (2015) using temporal correlation and low-rank matrix completion. Xu et al. (2014) combined motion and appearance information to detect and remove raindrops in outdoor scenes.

(c) Very Inclement Weather: To the best of our knowledge, vehicle detection under

very harsh conditions (e.g., heavy snow, blizzard) has not been studied thoroughly. There may be two reasons. On one hand, it is difficult to record videos in very harsh weather, as setting up digital cameras could be dangerous and recorded data might be unstable. Traffic surveillance cameras installed along the road are likely to be damaged by strong wind and severe cold. On the other hand, only a few public datasets provide video sequences in inclement weather, leading to a lack of enough data. Vehicles in blizzard were tested by Varadarajan et al. (2015) using Region-based Gaussian Mixture Models, with no specific detection rate.

2.4 Vehicle Detection for Traffic Surveillance

In recent decades, increasing emphasis has been given to issues related to on-road traffic safety. The endeavors in solving these problems have triggered the interest towards the field of Advanced Driving Assistance Systems (ADAS), where several on-road tasks have been proposed (Bertozzi et al., 2000). These tasks include, but not limited to, vehicle counting, vehicle tracking, vehicle speed measurement, traffic flow estimation, traffic violation detection, vehicle type recognition and incident detection.

Table 2-4. A brief summary of vehicle detection approaches based on feature fusion.

Reference	Number of Features	Features	Field of View	Application Domain
Chan et al. (2007)	4	Underneath, vertical edge, symmetry	Rear	Highway scene detection
Li and Wang (2015)	4	Sketch, texture, color, flatness	Front	Complex urban traffic conditions (with occlusion)
Tsai et al. (2007)	3	Color, corner and edge	Up-down	Parking area
Wang and Cai (2015)	4	Vertical edge, symmetry, HOG and Haar-like features	Rear	Freeway
Chong et al. (2013)	2	Shadow, illumination entropy, edge	Rear	Overtaking
Sun et al. (2006a)	2	Rectangular feature and Gabor filter	Rear	Complex scenes including rainy and congested traffic conditions.
Lin et al. (2012)	2	Appearance based feature &edge-based feature	Side	Blind-spot detection
Haselhoff and Kummert (2009)	2	Haar and triangle features	Rear	Urban scene

2.4.1 Major On-road Driving Problems

Real world traffic surveillance sequences capture complex interactions between vehicles. Based on various vehicle detection approaches, multiple on-road driving tasks can be achieved using traffic surveillance cameras.

(a) Overtaking Detection: Overtaking is a common phenomenon in everyday on-road driving. Blind spot areas will occur when vehicle overtaking happens, which increase the possibilities of traffic accidents. To solve this problem, Zhu et al. (2006) detected vehicle overtaking using an integration of dynamic scene modeling, hypothesis testing and information fusion. Based on the rear view, the detection showed robust performance in various traffic environments (highways, suburban roads and city streets). In another study (Garcia et al., 2012), overtaking was detected by using two sensors, Radar and the camera. Daytime and nighttime were tested on rear view with all overtaking scenarios detected. RGB-D camera provides another solution for overtaking detection, as lost vehicular visual information can be recovered using depth information. Based on the observation that grayscale in image changes with depth value, Xia et al. (2015) used depth data to recognize poster change for overtaking detection.

To prevent the occurrence of changing lanes and overtaking, several blind-spot detection systems have been proposed. Blanc et al. (2007) used a combination of edge features and support vector machine learning for the detection of blind-spot. Ra et al. (2018) used a side-rectilinear image to detect side parts of the vehicles. Both vehicles and motorcycles were detected.

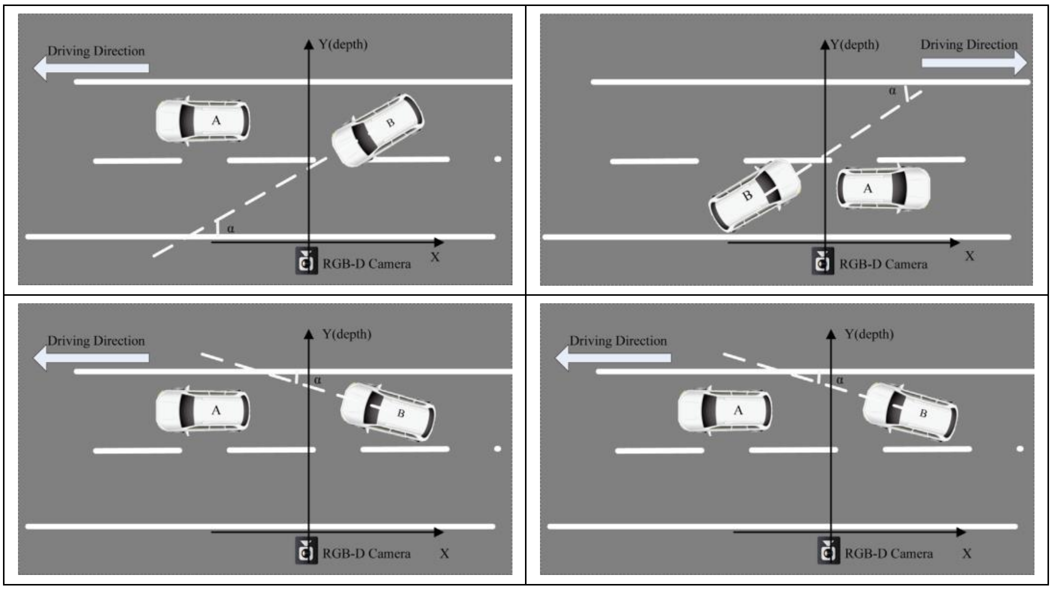


Figure 2-9. Illustration of different overtaking scenarios.
Source: Xia et al. (2015).

(b) Occlusion Handling: In real-world video sequences, interactions among vehicles result in fully or partial occlusion. When occlusion happens, vehicles can only be partly seen in images. In general, vehicle occlusion can be categorized as occlusion among vehicles, and occlusion between vehicle and multi non-vehicle objects.

Several solutions have been proposed to handle the problem of occlusion. Senior et al. (2006) modelled vehicle appearance to handle the occlusion problem. Vans with partial occlusion were tracked at intersections. Zhao et al. (2016) proposed an adaptive partial occlusion segmentation method (APPOS) for multiple vehicles tracking. Occlusions were detected occlusion by finding abnormal foreground and evaluated by the contour's optical flow. Tian et al. (2015) handled partial occlusion by designing a novel grammar model. Structure, deformation and pairwise SVM were utilized to construct the model. Then, occluded vehicles were tested using the sub-set of semantic parts.

2.4.2 Derivation of Vehicle Parameters

Another important aspect of traffic surveillance is to derive traffic parameters for the purpose of vehicle tracking, counting, classification and speed measurement. By using a combination of hardware sensors and vision cameras, various traffic parameters can be derived. In this part, derivation of three main traffic parameter are reviewed: a) traffic volume, b) trajectories of vehicles and c) average speed of moving vehicles.

(a) Vehicle Counting: Counting vehicles in traffic scenes is very helpful in evaluating traffic status. Based on the counted number of moving vehicles on the road, traffic flow can be estimated. Conventional counting methods rely on electromagnetic-based devices, such as inductive-loop detectors (Cheung et al., 2005) for counting. In computer vision, a majority of vehicle counting methods rely on clustering of points or adaptive background subtraction to count vehicles

in image sequences. Zhao and Wang (2013) proposed an approach of counting vehicles by tracking clustering feature points in the scene. In each path, average vehicle size was estimated and feature points were clustered into vehicles. Bas et al. (2007) used adaptive background subtraction and Kalman filter for vehicle detection, tracking and counting in nighttime and adverse weather conditions. Vehicle counting on highways has been studied by Liu et al. (2016) in the compressed domain.

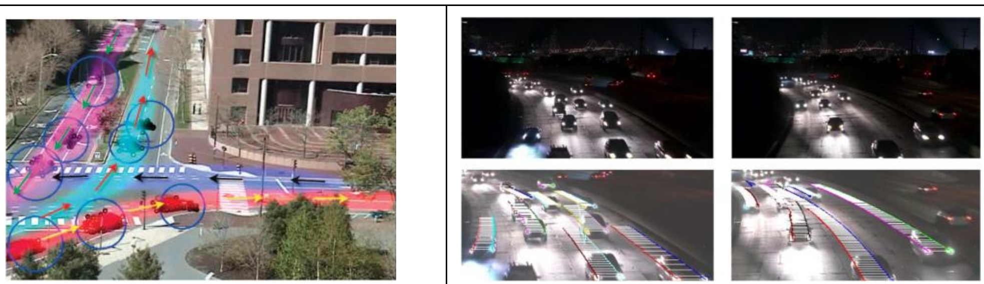
In recent decades, more methods combine hardware-based detectors with videos cameras for accurate vehicle counting. Rabbouch et al. (2017) used a virtual sensor and counted vehicle numbers based on Infinite Mixture Models. A user-defined virtual loop was used by Barcellos et al. (2015) for vehicle counting after vision-based vehicle detection. This enhancement makes vehicle counting more reliable.

(b) Vehicle Tracking: The majority of existing on-road vehicle tracking systems follow a detect-then-track scheme (Sun et al., 2006a). The tracking procedure aims to trace extracted vehicle candidates in consecutive frames. An adaptive partial occlusion segmentation method (APPOS) was proposed by Zhao et al. (2016) for multiple vehicles tracking. Kim et al. (2005) proposed a method for daytime and nighttime tracking using vision and sonar sensors. Daytime vehicles were tracked based on on-line template matching whereas nighttime vehicles were tracked using

observations of several consecutive frames. Anandhalli and Baligar (2017) implemented on-road vehicle tracking with Raspberry Pi and USB camera.

(c) Vehicle Speed Measurement: Vehicle speed is an important parameter to provide information related to traffic volume and density, for the purpose of on-road traffic analysis. Different sensors have been utilized to accurately measure the driving speed of moving vehicles, where inductive loop detectors (Ki and Baik, 2006) and vision cameras are the mainly-used equipment.

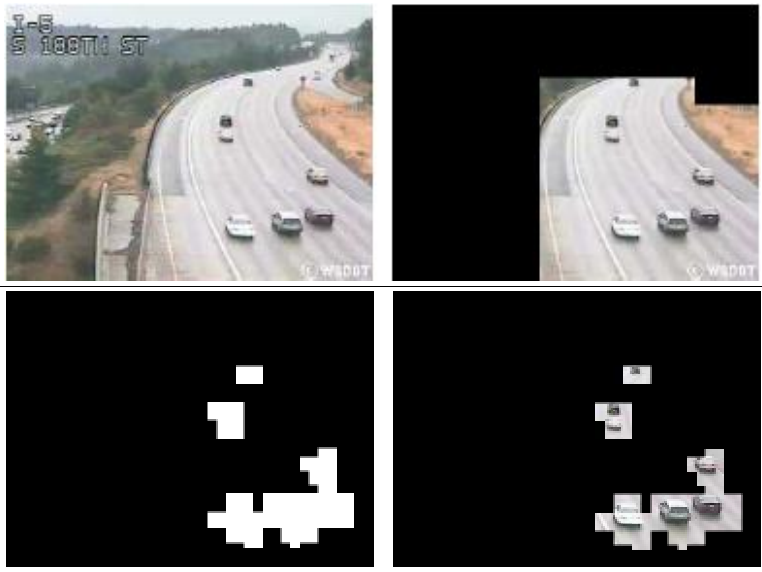
From images, the vehicle speed can be derived by calibrated (Lan et al., 2014) or uncalibrated camera (Nguyen et al., 2011). Madasu and Hanmandlu (2010) estimated vehicle speed based on motion tracking through a sequence of images. Luvizon et al. (2017) detected moving vehicle based on license plates. Vehicle speed was measured by comparing trajectories of tracked features to associated ground truth speeds obtained by an inductive loop detector.



(a) (b)

Figure 2-10 To be continued on next page

(c)



(d)

Figure 2-10. Vehicle counting and tracking in recent studies. (a) Counting vehicles on each of the intersecting paths separately (Zhao and Wang, 2013). (b) Headlight trajectories in colored lines and pairing results connected by white lines at nighttime (Zou et al., 2015). (c) Tracking Results in Congested urban environments (Liu et al., 2013). (d) Vehicle Counting on Highways (Liu et al., 2016).

2.5 Summary

To sum up, a comprehensive review of vision-based vehicle detection approaches is presented in this section. First, existing vehicle detection approaches are categorized as appearance-based methods and motion-based methods. Then, solutions under varying environments are provided, where illumination and weather are the two major concerns. Last but not least, representative traffic surveillance objectives are summarized in terms of major on-road driving problems and derivation of vehicle parameters. From the exhaustive review, we get the following conclusions:

Vision-based vehicle detection still faces many challenging problems, such as occlusion in congested urban scenes, swaying flags and leaves in the case of varying background and the presence of raindrops and snowflakes in adverse weather conditions.

With the continuous development of computer vision technologies, vision-based vehicle detection has achieved remarkable progress in expansion of FOV (from the conventional front/rear view to multiple FOVs), running time (real-time detection can be achieved even in very complex traffic scenes) and accuracy.

Despite exhaustive efforts in exploring efficient and accurate algorithms for vehicle detection, most studies are limited in providing solutions in just a few adverse conditions. There has not been a universal vehicle detection approach for all-time, all-weather traffic surveillance, especially under adverse conditions.

CHAPTER 3 METHODOLOGY ON VEHICLE DETECTION UNDER ADVERSE WEATHER CONDITIONS

In this section, an enhanced Three-Frame-Differencing (E-TFD) approach to detect and count moving vehicles in adverse environmental conditions is proposed. This method involves several innovative mechanisms and refines some classic methods for better implementation in challenging environments. First, the basic characteristics of vehicular images under varying environments are summarized, followed by a detailed methodology of vehicle detection and corresponding refinements in low illumination and adverse weather conditions. In the end, two main traffic surveillance objectives, namely, vehicle counting and tracking, are elaborated in this study.

3.1 Basic Characteristics of Traffic Images under Varying Environments

In outdoor scenes, many environmental factors may affect the quality of traffic surveillance videos. Under bad weather conditions, the light reaching a camera is severely scattered by small particles in the atmosphere (Narasimhan and Nayar, 2003). Due to the particles, the contrast of images degraded severely. The presence of raindrops and snowflakes adds difficulty in accurate detection of moving vehicles.

Traffic surveillance cameras are often operated in RGB color model, but the three color channels are highly related and individual values of red, green and blue rely on brightness strongly (Yang et al., 2011).To measure the intensity of light at each pixel, conversion from RGB color images to the grayscale image is operated. With a range from 0 to 255, histogram could be the most direct representation of pixel distributions.

3.1.1 Intensity Histogram

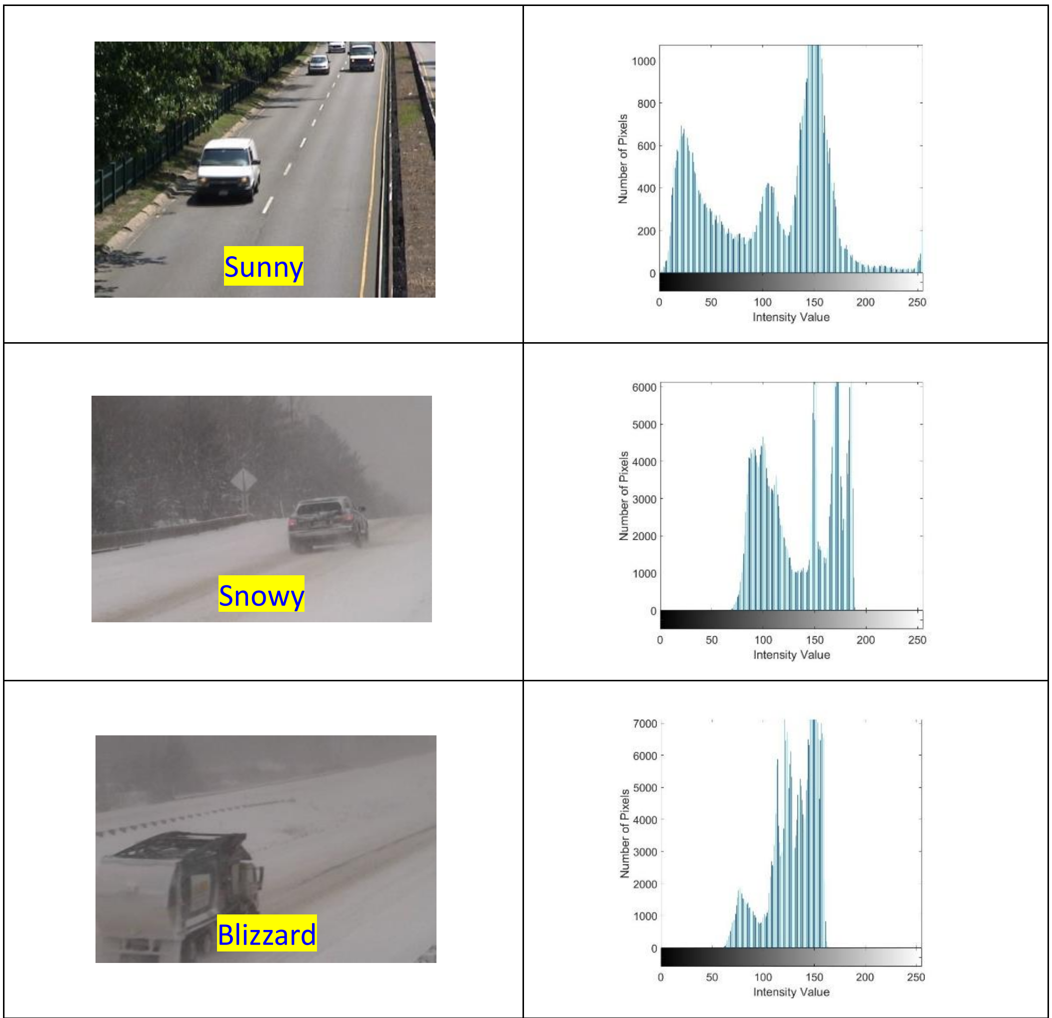


Figure 3-1 To be continued on next page

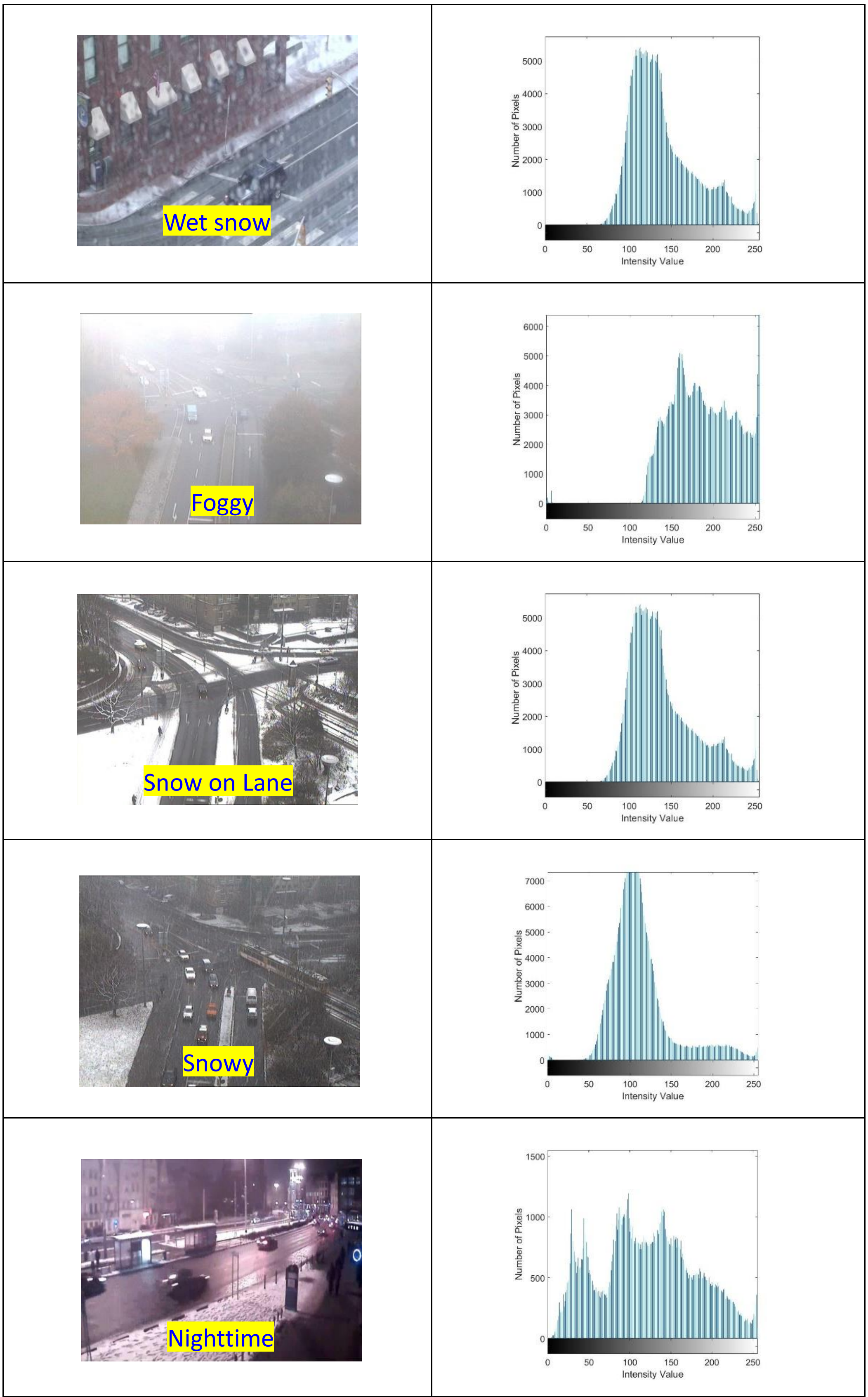


Figure 3-1 To be continued on next page

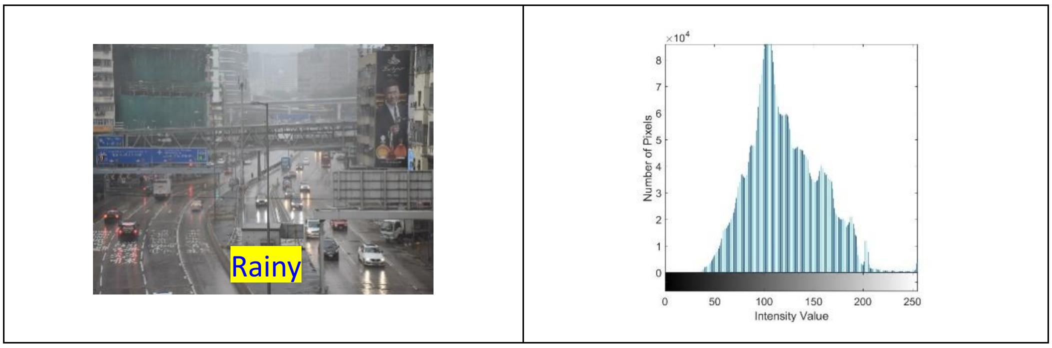


Figure 3-1. Intensity histogram of traffic images under varying environments.

From histograms shown in Figure 3-1, it is observed that histograms of bad-weather images show uneven distributions. Intensity histograms of snowy, rainy and blizzard images have a more concentrated distribution. There are barely any pixels that are smaller than 50 and larger than 200. In foggy images, pixels are more concentrated in larger intensity values. Apart from sunny and nighttime images, histograms of almost all bad weather conditions denote a lack of pixel distribution between 0 and 50. Nighttime images show no obvious lack of distribution between 0-100. In each condition, more than 100 frames are tested, based on which intensity histogram are drawn and analyzed. Therefore, uneven distribution and larger numbers of high-intensity pixels are the characteristics of bad-weather images.

3.1.2 Intensity Distribution of Specific Pixel Ranges

From Figure 3-2, the sum of pixels of a grayscale image is calculated and categorized as 0 – 100 (dark blue part) and 101 – 256 (yellow part). Different environmental conditions are calculated and shown in Figure 3-2 in the form of pie charts. Figure 3-3 shows a clearer intensity distribution in different pixel ranges. In the case of fog, image intensity between 0 – 100 takes up a very small portion of total pixel values, generally less than 1%. In other conditions, image intensity between 0 – 100 takes up more than 1% of all intensity values. This observation can be employed to categorize input image as fog or non-fog conditions, for the methods described in the later section.

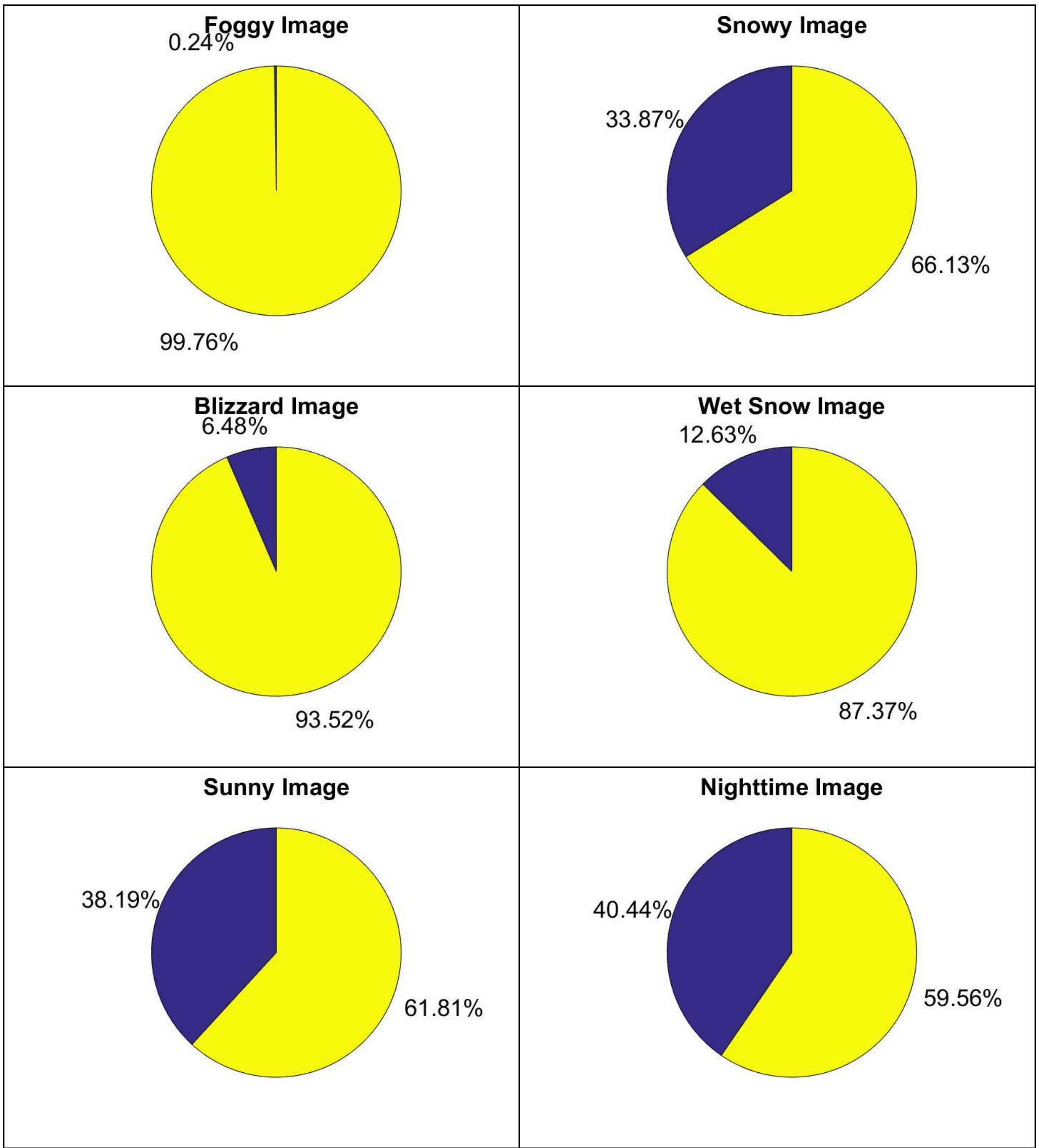


Figure 3-2 To be continued on next page

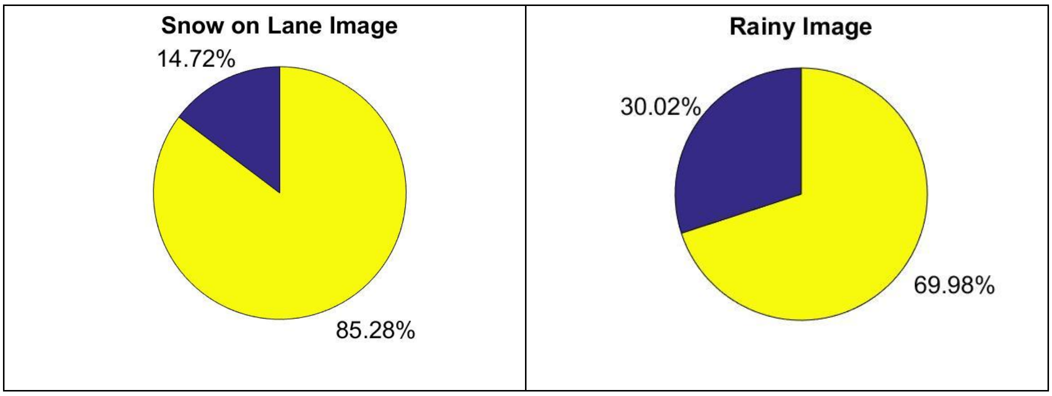


Figure 3-2. Intensity distribution of grayscale image under different weather conditions.

3.2 The Enhanced Three-Frame-Differencing Method

To deal with challenging cases of vehicle detection under adverse illumination and weather, the detection framework is proposed to implement in all conditions. Then, refinements under low illumination and adverse weather conditions are provided in Section 3.2.2 and 3.2.3, respectively. The goal of this study is to accurately detect moving vehicles in consecutive image sequences and to derive the count in real-time (Figure 3-3).

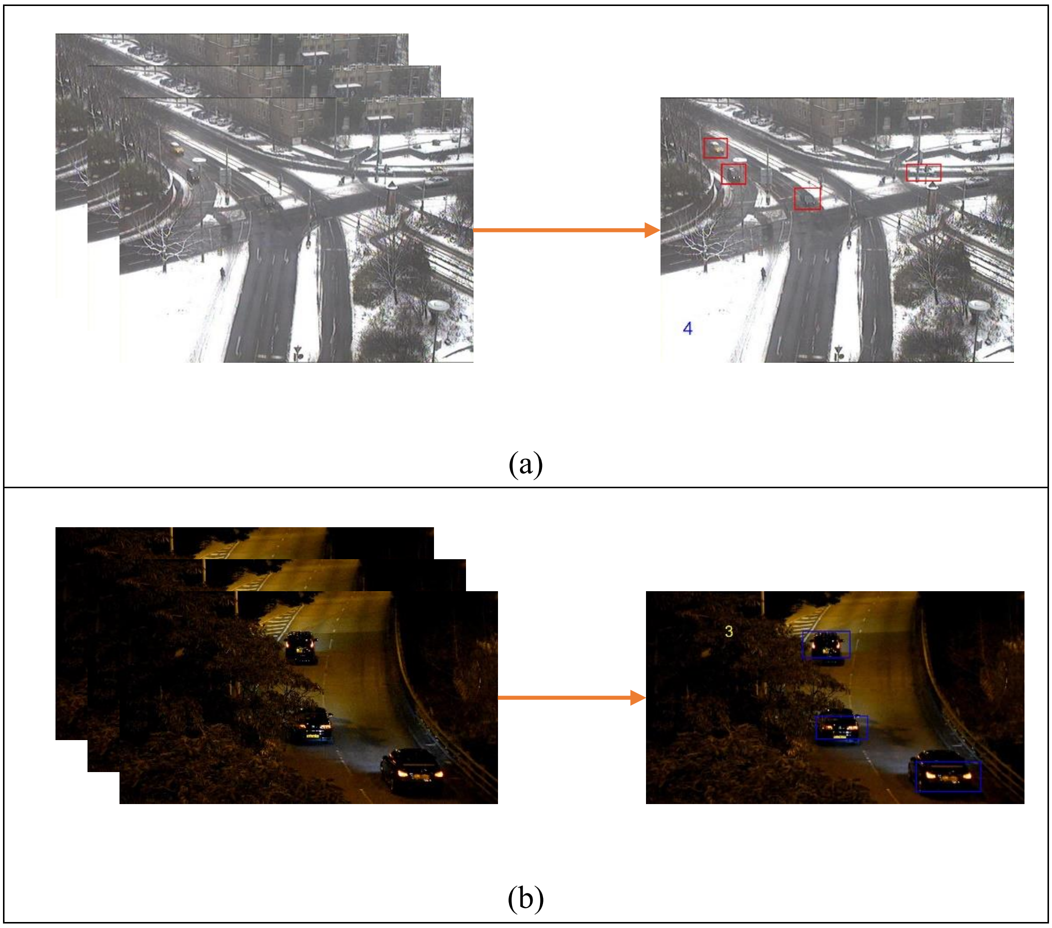


Figure 3-3. Sample illustration of enhanced TFD detection approach. Input images and detection results are shown in the case of Snow on lane (a) and Nighttime (b).

3.2.1 The Detection Framework

The detection framework of vehicle detection under varying environments is shown in Figure 3-4. In a given video sequence, every three consecutive frames are utilized to generate frame differencing images. The detection step can be divided into 6 steps, pre-processing, frame-differencing, thresholding, removal of small noise regions, morphological operations and removal of walking pedestrian. Based on the detection framework, moving vehicles can be extracted in an efficient way. Due to the challenging illumination and weather, however, different preprocessing steps need to be applied in individual scenario. At the same time, different criteria are set to remove small noise regions in different environmental conditions.

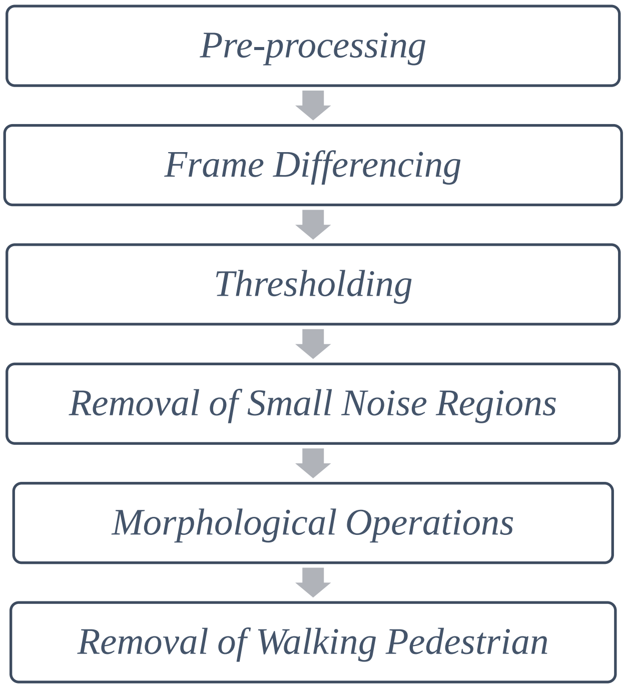


Figure 3-4. Flowchart of the enhanced Three-Frame-Differencing method.

(1) Preprocessing

To deal with the challenging issues of low illumination and bad weather, a preprocessing step is applied in each condition. At nighttime, we use a color space conversion from RGB color space to L*a*b* color space to extract the red color of vehicle taillights. In the case of fog, Histogram Equalization is applied to restore the contrast of foggy images. In self-collected videos, Gaussian Pyramid (Adelson et al., 1984) is applied to reduce image resolution and improve processing speed. Detailed elaboration of preprocessing will be discussed in Section 3.2.2. and

Section 3.2.3.

(2) Frame Differencing

Next, moving vehicles are segmented based on Three-Frame-Differencing. Traditional frame differencing is operated between two frames. Here, three consecutive frames are used to generate binary image maps. The basic idea of Three-Frame-Differencing is adopted from Weng et al., (2010), based on which several refinements have been made. First, FD_1, FD_2 and FD_3 are used to calculate difference image maps using Equation (1), (2) and (3). Based on that, two difference images, It_1 and It_2 are generated. It_1 and It_2 are then added to get the enhanced difference image, It. From a simple illustration in Figure 3-5, it is obvious that the binary images generated by three frames characterize moving vehicles with clearer outlines. In adverse conditions, three-frame-differencing can better extract moving vehicles.

$$It_1 = |FD_2 - FD_1| \tag{1}$$

$$It_2 = |FD_3 - FD_2| \tag{2}$$

$$It = It_1 + It_2 \tag{3}$$

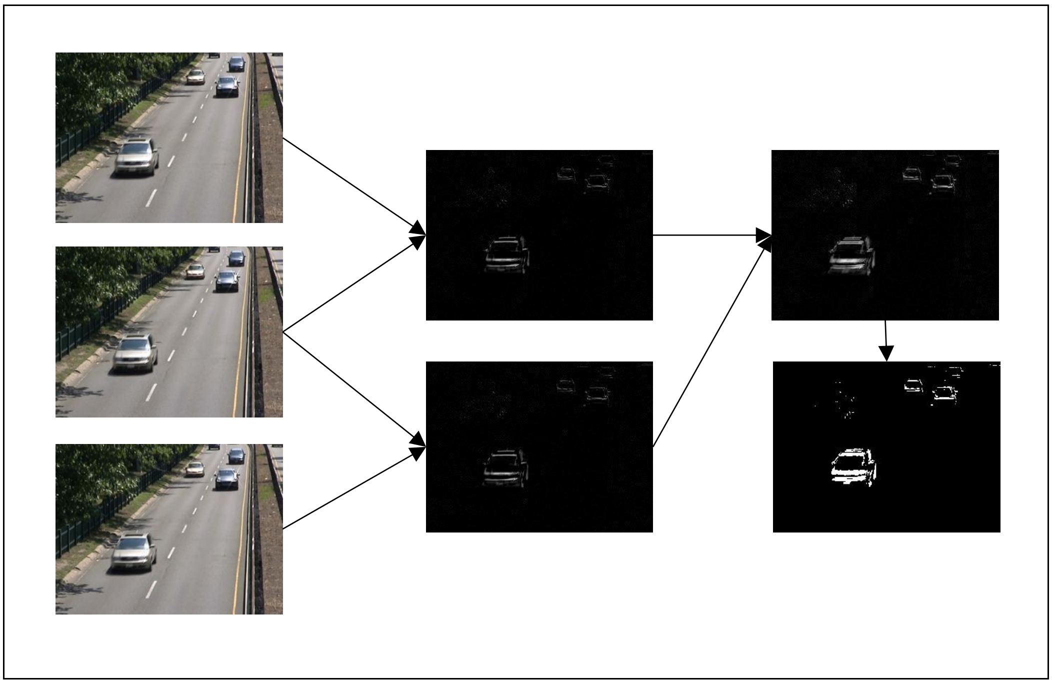


Figure 3-5. Simple illustration of the Three-Frame-Differencing approach.

(3) Thresholding

Based on the contour of difference images, thresholding is applied to segment foreground (contain vehicle candidates) and background (do not contain vehicle candidates) images to generate binary image maps. Equation (4) gives a brief criterion to determine the thresholding value of each binary image. The quality of segmentation relies highly on the cautious selection of thresholding values.

In this experiment, three different thresholding methods are used to deal with different illumination and weather conditions, which are 1) Otsu thresholding, 2) Fixed level thresholding and 3) Self-defined statistical thresholding to deal with different scenarios.

$$B_{img} = \begin{cases} 255, & if\ It > T \\ 0, & if\ It \leq T \end{cases} \tag{4}$$

Otsu Thresholding: Otsu thresholding method (Otsu, 1979) has been used been many studies for the purpose of efficient object segmentation. In sunny images, the illumination of image sequences is fine. Accordingly, intra-class variance among pixels is huge. Therefore, Otsu thresholding is suitable for foreground generation in fine illumination conditions.

Fixed Level Thresholding: In the case of fog and snow on lane, a fixed thresholding value is applied to segment moving vehicles. Based on empirical knowledge, a thresholding value of 45 is chosen in both cases. Other thresholding methods such as Otsu thresholding and thresholding based on statistics as mean and standard deviation have been tried, but none of them perform well in foggy images. The main reason may be the small intra-class variance of pixels. Therefore, knowledge-based thresholding is applied in foggy and snow-on-lane images.

Statistical Thresholding: In inclement weather conditions, however, the Otsu method might not work. Due to poor illumination and the presence of raindrops/snowflakes, the intra-class variance is not obvious in image scenes. Therefore, self-defined statistical thresholding based on statistics of intensity histograms is used to determine the threshold value. Accumulative nonzero pixels in the whole image are calculated and the value is set to 99%. The detailed steps are as follows:

Step 1: Calculate total numbers of pixels in the image based on rows and columns and set the thresholding value based on statistics.

$$N = m * n * 0.99; \tag{5}$$

Step 2: For each intensity level, compute accumulative nonzero pixels in FD

$$h_x(i) = \sum_{j=0}^{i} p_x(j),\ 0 \le i \le L \tag{6}$$

Step 3: Find the intensity level T, where accumulative pixels reaches N.

$$\mathrm{T} = \mathrm{J}\ \text{ where } h_x(j) = N \tag{7}$$

(4) Removal of Small Noise Regions

In different environmental conditions, the size of image noise could be different. On one hand, the variation could be caused by different sizes of raindrops and snowflakes. On the other hand, image noise size varies due to various FOV of traffic surveillance cameras. In the case of bad weather, the presence of raindrops and snowflakes might increase the possibility of False Positive. Accordingly, if small noise regions are not removed, vehicle counting based on the presence of ROIs might be inaccurate.

Here, instead of commonly-used operations of morphological opening and closing (Weng et al., 2010), small noise regions are removed based on region size using Equation (8). Based on a selection of specific region size in corresponding conditions, small noise regions can be removed efficiently. If the area of non-zero pixels is larger than P, the area is categorized as possible vehicle candidates, otherwise will be categorized as image noise and removed. The selection of P in different weather conditions is determined manually, shown in Table 3-1.

$$BW = \begin{cases} Possible\ Vehicle\ Candidates\ , if\ It > P \\ Image\ Noise,\ if\ It \leq P \end{cases} \quad (8)$$

Table 3-1. Parameters selection for removal of small regions.

Weather	**Dataset**	**Parameter for Small Area Removal**
Sunny	CDnet 2014	10
Wet snow	CDnet 2014	100
Blizzard	CDnet 2014	50
Snowfall	CDnet 2014	50
Fog	KIT	10
Snow on Lane	KIT	10
Rainy_DSC_0652	Self-collected data	10
Nighttime_DSC_0376	Self-collected data	10
Nighttime_DSC_0546	Self-collected data	10

In the case of wet snow, the camera is covered with the presence of large raindrops. Therefore, a large area criterion (100) is selected. In the case of sunny highways, a small threshold (10) is selected due to the small outline of moving vehicles. Figure 3-6 shows the image after region removal and some representative detection results.

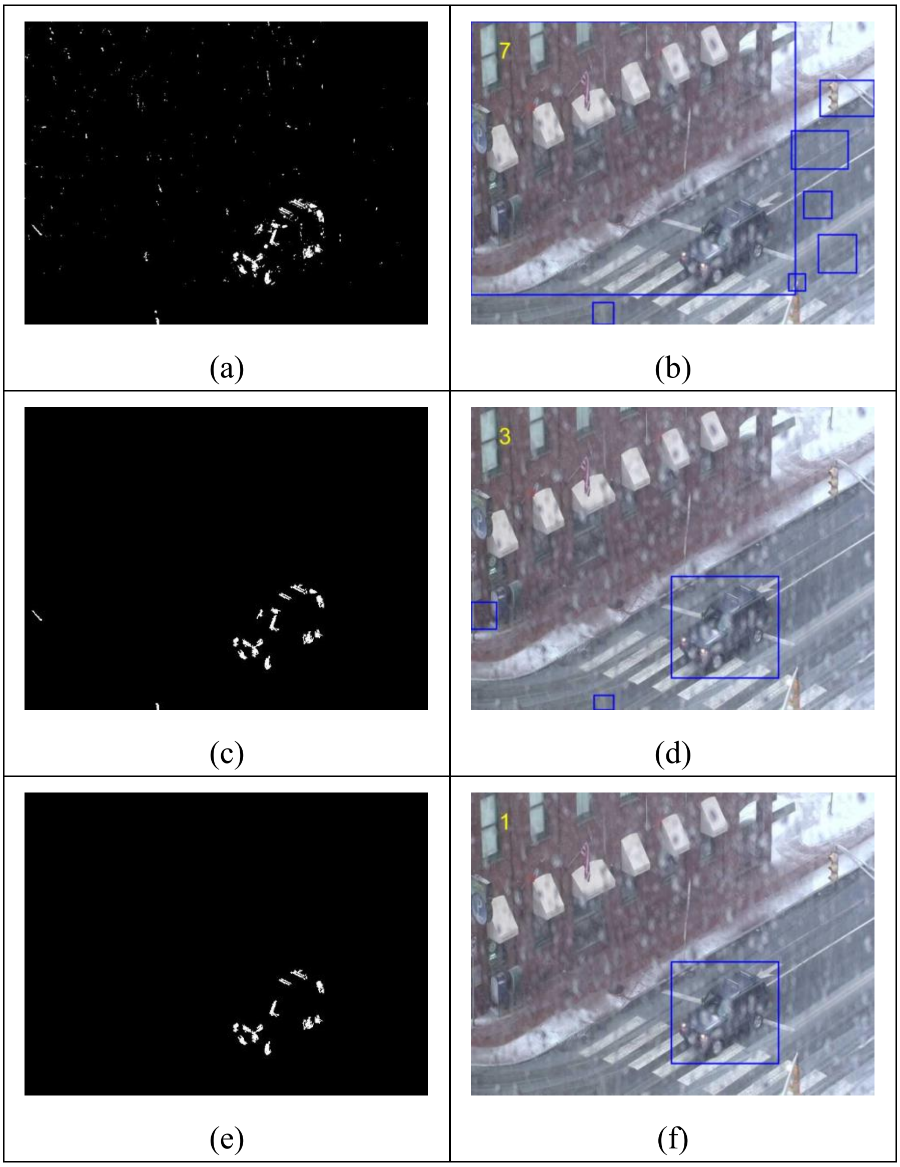


Figure 3-6. Illustrations of small region removal in the case of wet snow (frame 157,158 and 159). (a) Binary image map after statistical thresholding and removal of regions that containing fewer than 30 pixels; (b) ROI of (a); (c) Binary image map after statistical thresholding and removal of regions that containing fewer than 50 pixels; (d) ROI of (c); (e) Binary image map after statistical thresholding and removal of regions that contain fewer than 100 pixels; (f) ROI of (e).

(5) Morphological Operations

Based on binary image map *BW*, morphological operations are implemented to connect discontinuous edges. Common frame differencing algorithms use morphological open and closing (Weng et al., 2010, Kim and Kim, 2003) to eliminate small noise regions and connect small apertures. In inclement weather conditions, however, morphological open might enlarge unwanted noise region in the image. Region growing (Adams and Bischof, 1994) is also an effective method, but the main drawback is its long processing time. To meet the requirement of real-time and accurate processing, operations of closing and dilation are used to finalize ROIs of vehicle candidates. The selection of kernel size is listed in Table 3-2. For all operations, the kernel shape is considered as 'square', and kernel size is adjusted in different conditions.

Table 3-2. Parameters selection for morphological operations.

Weather	**Dataset**	**Kernel Size for Closing and Dilation**
Sunny	CDnet 2014	12
Wet snow	CDnet 2014	50
Blizzard	CDnet 2014	50
Snowfall	CDnet 2014	50
Fog	KIT	15
Snow on Lane	KIT	15
Nighttime_DSC_0376	Self-collected data	50
Nighttime_DSC_0546	Self-collected data	50
Rainy_DSC_0652	Self-collected data	50

1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1

Figure 3-7. A square-shaped structuring element with a size of 5*5.

(6) Removal of Waking Pedestrian

Through the aforementioned five steps, possible ROIs of moving vehicles can be obtained in every single image sequence. However, the presence of moving pedestrian could also be wrongly detected as vehicles. To solve this problem, the removal of small noise regions is added as the last processing step to reduce False Positives Rate and improve detection rate using Equation (9). With this operation, walking people appeared in images can be removed efficiently.

$$\mathrm{ROI} = \begin{cases} Vehicle, if\ ROI\ area <> smallest\ in\ list \\ Pedestrian, if\ ROI\ area = smallest\ in\ list \end{cases} \tag{9}$$

3.2.2 Refinements under Low Illumination Conditions

At daytime, the illumination of traffic images is good. During the transition from daytime to nighttime, illumination degrades drastically. At nighttime, vehicles are detected by a combination of enhanced TFD method and L*a*b*-color-space-based taillight detection. The detection step is illustrated in Algorithm 1.

Algorithm 1 Vehicle Detection in Low Illumination Conditions

Problem:

For a given video sequence, determining whether the ROI of vehicles exists in each frame.

Solution:

A taillight based method is used to extract vehicle candidates.

1: **for** each $Frame_i$ **do**

2: Convert orginal image from RGB color space to L*a*b* color space, get L*, a*, b* color channel.

3: Binary L* component based on statistical thresholidng (99%), get *Limg*.

4: Binary a* component based on statistical thresholding (99%), get *Aimg*.

5: Calculate the absolute difference value of $Frame_{i-1}$ & $Frame_i$ in grayscale to get difference image FD_1.

6: Calculate the absolute difference value of $Frame_i$ & $Frame_{i+1}$ in grayscale to get difference image FD_2.

7: Add FD_1 and FD_2, get *FD*.

8: Binary *FD* based on statistical thresholding, get *It*.

9: Operate logical AND between *It* and *Aimg*, get *Bimg*.

10: Remove small noise regions caused by bad weather based on setting critera of area, get BW.

11: Perform morphological dilation and closing on *BW* to connect discontinous regions, get D*img*.

12: Based on *Dimg*, determine the precense of moving vehicles.

13: **end for**

To extract the red color of vehicle taillights at nighttime, a color space conversion is operated as the first step. The L*a*b* color space is chosen as the target space, as the a* component can extract and represent red color effectively. The L*a*b* color space was first defined by Hunter (1958) and specified by CIE (Commission Internationale de I'Eclairage). Based on opponent color theory, the

L*a*b* space is a three-dimensional real number space that contains infinite representation of colors. Accordingly, the L*, a* and b* values have a specific range, shown in Table 3-3. Of all three components, the red/green opponent color can be represented using the a* axis. When b* = 0; positive values of a* correspond to red color (Cabani et al., 2005). Based on the pre-defined range of a* component, 128 is added to each pixel of this color channel to make all pixel values an integer value.

Table 3-3. A brief illustration of color components of the L*a*b* color space.

Layer	Meaning	Range
L*	Luminance	0 to 100
a*	The color variance between red and green	-128 to 128
b*	The color variance between yellow and blue	-128 to 128

Before applying the method on self-collected videos, we test the method first in a dataset containing only salient vehicles in the rear view. The dataset used here is Caltech Cars 1999 and 2001 dataset (Fe-Fei, 2003), which includes 126 images and 526 cars from the rear, respectively. The field of view is the rear view, and in each image, only one vehicle is included. Salient vehicle candidates are estimated based on L*a*b* color space and the detailed processing steps are illustrated in Algorithm 2.

$$length_{ROI} = length_{ROT} \tag{10}$$

$$height_{ROI} = height_{ROT} * 0.8 \tag{11}$$

$$Centroid_{ROI} = Centroid_{ROT} \tag{12}$$

From binary image map generated from color space conversion and Otsu thresholding, ROT of each salient vehicle can be obtained using x-axis and y-axis projection. Non-zero points are first calculated, from which the minimum and maximum coordinates in x-axis and y-axis are recorded, respectively. Using equations (10), (11) and (12), Region of Interest (ROI) can be confirmed based on the presence of Region of Taillights (ROT). Figure 3-9 shows a simple illustration of this method.

Algorithm 2 Salient Vehicle Detection based on L*a*b* Color Space

---

Problem:

For a given video sequence, determine whether the ROI of vehicles exists in each frame.

Solution:

A taillight based method is used to extract vehicle candidates.

1: **for** each $Frame_i$ **do**

2: Convert orginal image from RGB color space to L*a*b* color space, get *L*, a*, b** color channel respectively.

3: Convert the value of a* component from double to integer, get *Aimg*.

4: Segment *Aimg* using Otsu thresholding method, get *Bimg*.

5: Perform morphological opening on *Bimg* with kernel shape ‘square’ and size ‘5’, get *Oimg*.

6: Generate Region of Taillights (ROT) based on pixel projection in x-axis and y-axis.

7: ROI of vehicle candidates is finally ascertained based on the presence of ROT.

8: **end for**

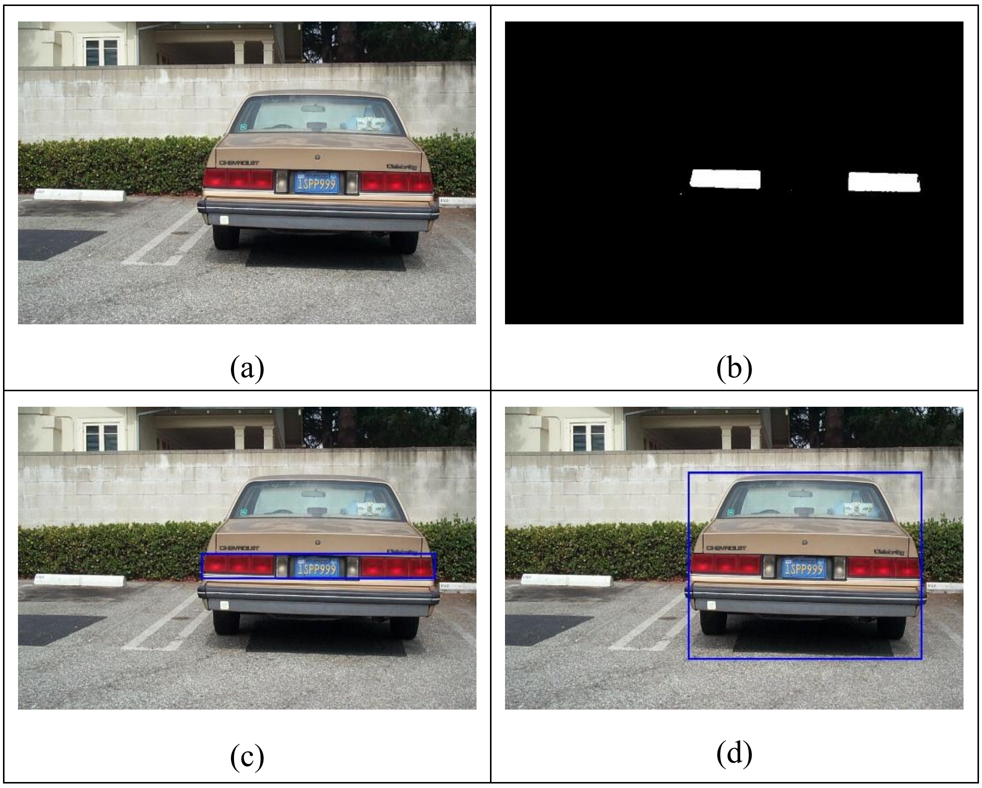


Figure 3-8. Simple illustration of taillight-based vehicle Detection. (a) Original image; (b) Binary image map of segmented taillights; (c) Region of Taillights (ROT); (d) Region of Interest (ROI).

3.2.3 Refinements under Adverse Weather Conditions

To deal with the inclement weather, vehicle detection is based on Three-frame-differencing for two public datasets (Wang et al., 2014, Karlsruhe, 1997) and self-collected videos. The weather conditions tested in the experiment include sunny, rain, fog, blizzard, snow on lane and wet snow. The detection approach is based on those described in Section 3.2.1. Algorithm 3 provides a brief summary of vehicle detection approach in challenging weather conditions.

Algorithm 3 Vehicle Detection in Adverse Weather Conditions

Problem:

For a given video sequence, determine whether the ROI of vehicles exists in each frame.

Solution:

Frame-difference algorithm is used to extract foreground regions.

1: **for** each $Frame_i$ **do**

2: For each frame, apply a quick histogram judgement to define whether it is fog.

3: In the case of fog, apply Histogram Equalization on $Frame_{i-1}$, $Frame_i$ and $Frame_{i+1}$, to restore the contrast of foggy images.

4: Calculate the absolute difference value of $Frame_{i-1}$ & $Frame_i$ in grayscale to get difference image FD_1.

5: Calculate the absolute difference value of $Frame_i$ & $Frame_{i+1}$ in grayscale to get difference image FD_2.

7: Add FD_1 and FD_2, get FD.

8: Binary FD based on statistical thresholding, get It.

9: Remove small noise regions caused by bad weather based on setting critera of area, get BW.

10: Perform morphological dilation and closing on BW to connect discontinous regions, get D*img*.

11: Based on *Dimg*, determine the precense of moving vehicles.

12: **end for**

Before Hypothesis Generation, the input video images can be categorized into fog scene and non-fog scene based on a quick judgement of grey-level histograms. From the observation in Figure 3-2 (Section 3.1.2), the criterion is described as follows: if intensity value of pixels between 0 – 100 takes up less than 1% of all pixels, the image is categorized as fog, otherwise non-fog condition.

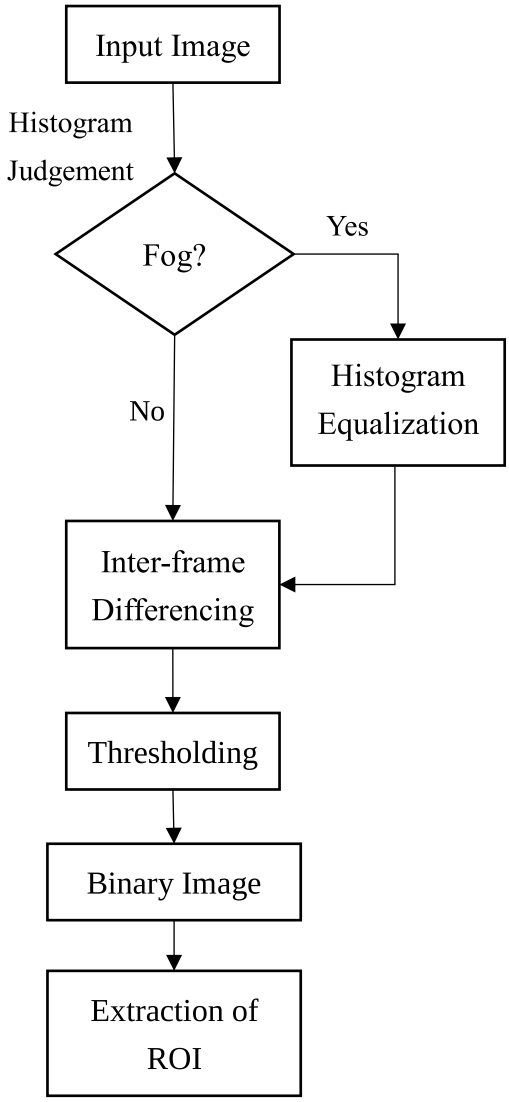


Figure 3-9. Processing flow of the histogram judgement.

In the case of fog, Histogram Equalization (HE) is applied as a pre-processing step to dehaze the image. As a widely-used operation of photometric correction, Histogram Equalization (Hummel, 1977) is a technique to enhance image contrast by adjusting image intensities. In early studies, histogram equalization was utilized in medical images for contrast enhancement. Now the method is expanded its applications to computer vision, where different types of images can be processed in an efficient manner.

In Figure 3-2, single color channel (R, G, B or grayscale), histograms of foggy images show uneven distributions. Therefore, the main objective of HE is to restore the pixel distributions, thus recovering the contrast of images in fog. The detailed description is provided below.

Step 1: Calculate the probability of the presence of a pixel in level i.

$$p_x(i) = p(x = i) = \frac{n_i}{n}, 0 \leq i \leq L \tag{13}$$

Step 2: Compute cumulative distribution function (cdf) corresponding to each pixel level.

$$cdf_x(i) = \sum_{j=0}^{i} p_x(j) \tag{14}$$

Step 3: Create a transformation to produce a new image, where k is in the range [0, L].

$$cdf_y(y') = cdf_y\big(T(k)\big) = cdf_x(k) \tag{15}$$

Step 4: Use a simple transformation to map the values back to their original range.

$$y' = y \cdot (\max\{x\} - \min\{x\}) + \min\{x\} \tag{16}$$

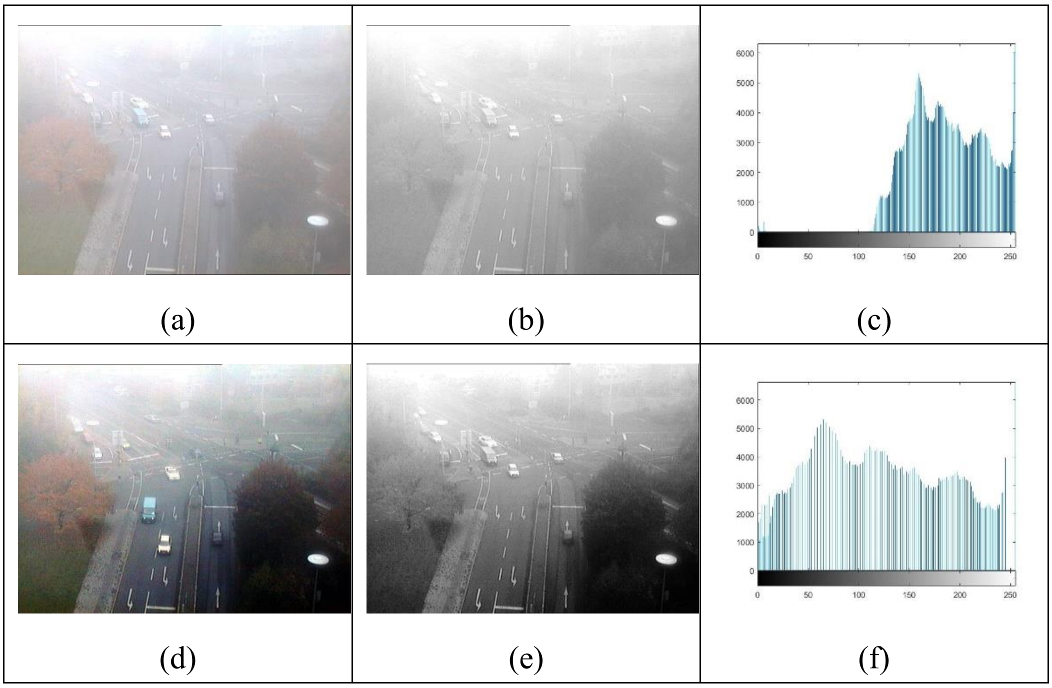


Figure 3-10. Illustrations of Histogram Equalization. (a) Original image (sequence 0013); (b) Grayscale image of (a); (c) Histogram of (b); (d) Synthesized color image of (a) after Histogram Equalization; (3) Grayscale image of (d); (f) Equalized histogram of (b).

The advantage of Histogram Equalization is that it can adjust pixel distribution of a single channel image automatically and perform in real-time. Figure 3-10 shows the implementation of Histogram Equalization of a single image captured in the Karlsruhe Institute Technology dataset. It is obvious that image after HE (Figure 3-10 (d)) is much clearer than the original image (Figure 3-10 (a)). Therefore, Histogram Equalization is a quick and efficient method for fog removal in traffic images and can be embedded into a real-time traffic surveillance system.

3.2.4 Vehicle Counting and Tracking

Based on the presence of ROIs using the enhanced Three-Frame-Differencing detection approach, the exact numbers of vehicles can be counted and displayed in each image frame. This step is implemented in MATLAB based on the summation of detected bounding boxes. Vehicle counting has been well studied in many research papers, but few of them provided exact counting in each frame. Compared to other studies, this step gives a clear value into a real-time Driver Assistance System.

3.3 Summary

In this section, an enhanced TFD approach for vehicle detection is presented, followed by refinements in low illumination and challenging weather conditions, respectively. To deal with the challenging environments, three major refinements have been made for accurate and efficient detection of moving vehicles.

First, statistical thresholding is applied after the procedure of Three-Frame-Differencing for fast segmentation of moving vehicles. This enhancement significantly improves the segmentation results by eliminating redundant non-zero pixels.

However, some small noise regions still exist in binary image maps. To remove them, an area-based image noise removal operation is used. Instead of morphological opening, a criterion based on the area of non-zero pixels is defined to remove small noise regions under varying environments.

Last but not least, morphological operations with a large kernel size is used for fast acquisition of Regions of Interest. With a square-shaped structuring element, morphological dilation and closing are applied to connect discontinuous edges, based on which ROIs of moving vehicles is finalized.

CHAPTER 4 EXPERIMENTAL RESULTS AND ANALYSIS

In the experiment, the enhanced TFD method has been tested in various traffic sequences to evaluate the accuracy and efficiency of algorithms. To start with, sample image data collected from different sources are introduced. With a special focus on varying environments, different illumination and weather conditions are tested, including sunny, rainy, foggy, snowy, blizzard, wet snow and nighttime images. The proposed statistical thresholding method works very well in even very adverse weather conditions.

4.1 Tested Datasets

In this study, video sequences containing different environmental conditions are gathered from two main sources, public datasets and self-collected videos. To the best our knowledge, many studies evaluate the performance of the proposed methods using only one category of source data, which share similar image size, resolution, illumination and field of view. Accordingly, detection results could be less convincing. A combination of tested data from multiple sources could be more challenging to test yet more robust for performance analysis, as more variations of images are presented. Here, the collected video sequences are characterized by multiple moving vehicles, non-ideal weather conditions and presence of fog, shadows, raindrops or snowflakes, some of which are with strong reflections on the ground and walking pedestrians. MATLAB R2016a is the software used to process data and run the algorithm. The processor used in this experiment is an Intel i7- 7700 CPU with a 64-bit operating system.

4.1.1 Public Datasets

Numerous datasets have been provided publicly for the implementation of vehicle detection. The main problem, however, is that not many public datasets contain traffic sequences in adverse conditions. ChangeDetection (CDnet) 2014 (Wang et al., 2014) and Karlsruhe Institute of Technology (KIT) (Karlsruhe, 1997) are the two major datasets that include traffic sequences in low illumination and bad weather conditions. In general, images from public datasets are recorded from a stationery camera, showing vehicles in different environmental conditions with various field of view, as listed in Table 4-1.

Table 4-1. A brief summary of tested data in public datasets.

Weather	Dataset	Image Resolution	Field of View	Number of Tested Frames
Fog	KIT Dataset	768 * 576	Top	338
Snow on lane	KIT Dataset	768 * 576	Top	300
Sunny	CDnet 2014	320 * 240	Front	1699
Blizzard	CDnet 2014	720 * 480	Rear	7000
Snowfall	CDnet 2014	720 * 480	Rear	6500
Wet snow	CDnet 2014	720 *540	Side	3000

4.1.2 *Self-collected Videos*

On the other hand, self-collected data of different environmental conditions are also used (Table 4-2). Nikon D5600 is the device used to record videos. The captured videos have a resolution of 1920 * 1080 and a frame rate of 60 frames per second. The captured videos are characterized by the presence of strong reflections on the road, low illumination at nighttime and swaying trees (Figure 4-1). To ensure fast computation, each original image is resized to 920 * 540 using the Gaussian Pyramid. The processing steps are as follows:

Step 1: Segment recorded videos to image sequences with a frame rate of 60 frame/second;

Step 2: Use the Gaussian pyramid to reduce the resolution of recorded images from *1920 * 1080* to *960 * 540*.

Table 4-2. A brief summary of self-collected videos.

Weather	Video Length	Image Size	Number of Tested Frames	Descriptions
Nighttime DSC_0546	10:14	1920 * 1080	93	Nighttime video captured in Kowloon, Hong Kong
Nighttime DSC_0376	5:05	1920 * 1080	328	Nighttime video captured in Kowloon, Hong Kong
Rainy DSC_0652	5:29	1920 * 1080	187	Rainy video captured in Kowloon, Hong Kong

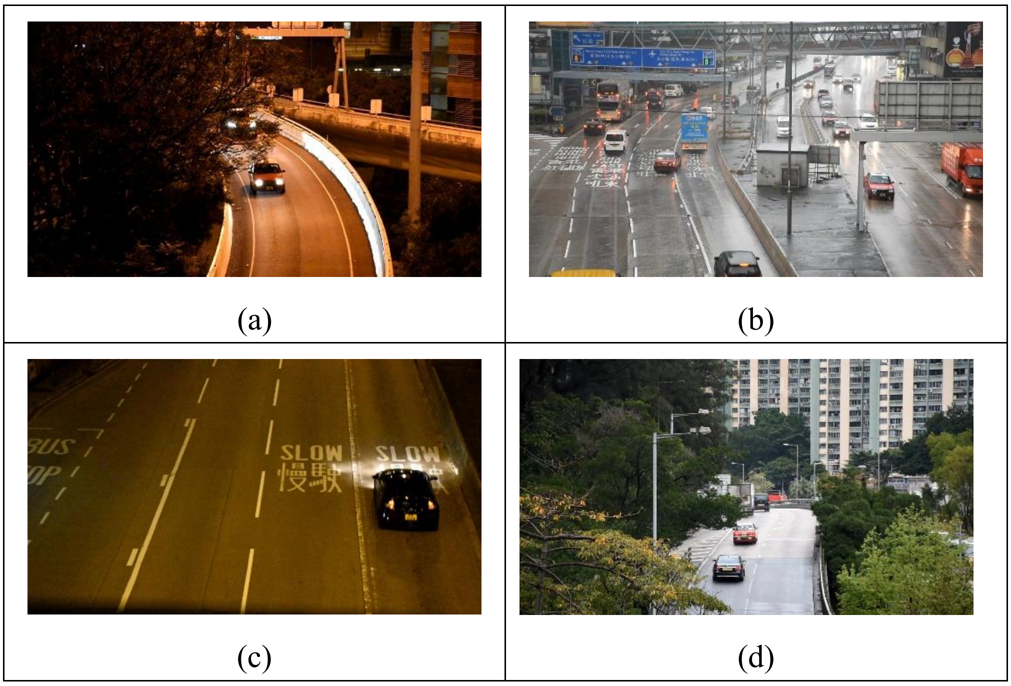


Figure 4-1. Sample of self-collected images. (a) Nighttime image in front view; (b) Rainy image with two driving lanes; (c) Nighttime image in rear view; (d) Rainy image in the rear view.

4.2 Performance of the E-TFD Method

In this section, the enhanced TFD method is evaluated in terms of several ROC parameters. ROC (Receiver Operating Characteristics) graph have long been used in vehicle detection to describe the tradeoff between hit rates and false alarm rates of classifiers (Fawcett, 2006). Four basic metrics are commonly used to evaluate the performance of the proposed method, namely, True positive rate (TPR), False Positive Rate (FPR), True Negative Rate (TNR) and False Negative Rate (FNR). A brief description of these four parameters can be found in Table 4-3. Based on these four metrics, metrics such as Precision, Recall, F Measurement (Equation (21), (22) and (23)) are further utilized to evaluate the enhanced TFD method.

Table 4-3. Performance indicators for evaluation.

Parameters	**Description**
True Positive	The number of vehicles that are correctly detected. Also known as *recall/detection rate.*
True Negative	The number of moving non-vehicles that are correctly detected.
False Positive	The number of moving non-vehicles that are incorrectly detected as moving vehicles. Also known as *false alarms.*
False Negative	The number of moving vehicles that are incorrectly detected as moving non-vehicles. Also known as *miss detections*.

$$\mathrm{TPR} = \frac{TP}{Total\ Number\ of\ Moving\ Objects} \tag{17}$$

$$\mathrm{FPR} = \frac{FP}{Total\ Number\ of\ Moving\ Objects} \tag{18}$$

$$\mathrm{TNR} = \frac{Number\ of\ moving\ non-vehicles}{Total\ Number\ of\ Moving\ Objects} \tag{19}$$

$$\mathrm{FNR} = \frac{Number\ of\ moving\ non-vehicles}{Total\ Number\ of\ Moving\ Objects} \tag{20}$$

$$Precision = \frac{TP}{TP+FP} \tag{21}$$

$$Recall = \frac{TP}{TP+FN} \tag{22}$$

$$F_{measure} = 2 * \frac{Precision*Recall}{Precision+Recall} \tag{23}$$

In this study, True Positive Rate, False Positive Rate, Precision and Recall are used for performance evaluation. Table 4-4 shows the results of the enhanced TFD method tested in 9 image sequences of different environmental conditions. With a total number of 11556 vehicles in tested sequences, the E-TFD method can successfully extract 10059 vehicles, achieving an average detection rate recall rate of 87.1% and average precision of 92%. The method performs best in self-collected nighttime image sequences DSC_0376, where a 100% FPR can be achieved. The E-TFD method has the worst performance in the case of fog, with a recall rate of 55.3% and precision of 92.0%.

Table 4-4. Performance of the enhanced TFD method in different environmental conditions.

Source Image	Number of Vehicles	True Positive	False Positive	Recall Rate	Precision	False Alarm Rate
Sunny	3765	3298	100	87.6%	97.1%	2.7%
Snowfall	755	755	209	75.5%	78.3%	27.7%
Blizzard	1184	1177	10	99.41%	99.2%	0.8%
Wet snow	557	540	59	97.3%	90.2%	10.59%
Foggy	1661	919	52	55.3%	94.6%	3.13%
Snow on lane	1900	1685	419	88.7%	80.1%	22.1%
Nighttime DSC_0376	572	527	6	92.1%	98.9%	1.0%
Nighttime DSC_0546	171	171	9	100%	95%	5.3%
Rainy DSC_0652	1064	1056	19	99.3%	98.2%	1.8%
Total	11556	10059	883	87.1%	92.0%	7.6%

4.2.1 Accuracy

It is illustrated in Figure 4-2 that the E-TFD method can accurately extract moving vehicles in both fine and low illumination, sparse and congested conditions and even with the presence of raindrops, snowflakes, walking pedestrian and swaying trees. In some very adverse conditions, e.g., blizzard, wet snow, fog, moving vehicles can be extracted in an accurate and efficient manner.

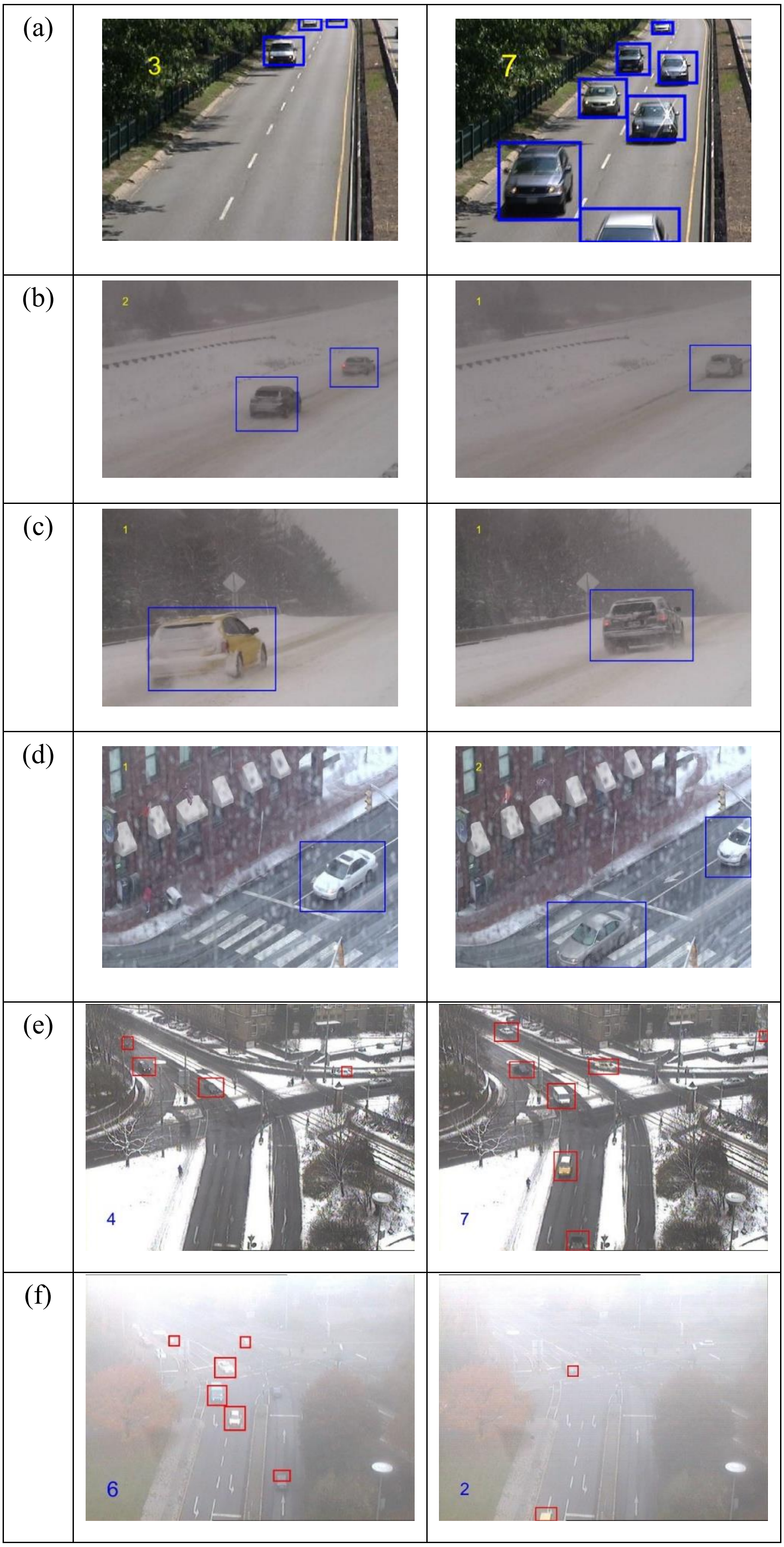


Figure 4-2. Successful detection and counting results using public datasets. (a) Highway; (b) Blizzard; (c) Snowfall; (d) Wet snow; (e) Snow on lane; (f) Fog;

Figure 4-3 shows detection results using self-collected data. The E-TFD method not only performs well at nighttime but also in the case of rain. At nighttime, vehicle candidates were extracted based on a combination of Three-Frame-Differencing and L*a*b*-color-space-based taillights extraction. On rainy days, despite the strong reflections on the ground and low illumination in video sequences, the enhanced TFD detection approach is able to extract moving vehicles with high accuracy.

Figure 4-3. Vehicle detection results using self-collected data. (a) Nighttime_DSC_0376; (b) Nighttime__DSC_0546; (c) Rainy_DSC_0652.

4.2.2 Efficiency

To be embedded in a workable traffic surveillance system, the proposed algorithm needs to be computationally efficient, preferably running in at least near real-time. Running time is an important evaluation factor for computer vision algorithms. Table 4-5 summarizes the running time of the enhanced TFD detection approach. For different environmental conditions, this system allows processing approximately 100 frames in one second for different categories of input videos.

Table 4-5. Running time in different conditions.

Test Frame	**Running Time (s)**
Sunny	0.010
Blizzard	0.009
Snowfall	0.010
Wet snow	0.008
Fog	0.011
Snow on lane	0.009
Nighttime_DSC_0376	0.016
Nighttime_DSC_0546	0.013
Rainy_DSC_0652	0.014

To make it more intuitive, the running time of each image frame is converted to the number of frames that the E-TFD method can process in each second (Figure 4-3). It is observed that the enhanced TFD approach performs best in wet snow images, in which 125 frames can be processed every second. The E-TFD method has the lowest performance in self collected sequences nighttime_DSC_0376, in which only 63 frames can be processed per second. The processing time would be affected by different resolutions of images in different datasets.

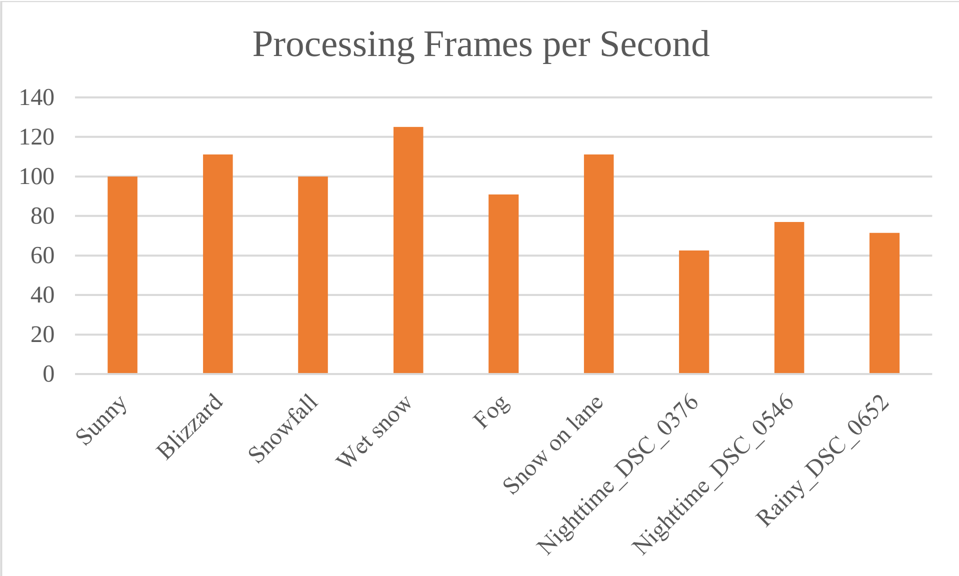


Figure 4-4. Processing speed of the E- TFD method under varying environments.

4.2.3 False Detection

False detection is inevitable in vision-based vehicle detection algorithms, especially under adverse conditions. As shown in Figure 4-5, some false detection results are caused by the presence of fog (d), raindrops, snowflakes (left image of (b)), swaying trees (left image of (e)) and flags (right image of (b)), walking pedestrian (right image of (d)) and occlusion among vehicles ((c) and right image of (e)). Figure 4-4 (a) shows a distant vehicle in the blizzard that cannot be extracted due to low illumination. In Figure 4-5 (f), the right image shows the wrong ROI caused by reflections of taillights due to rain. Reasons for false detection can vary, but typically there are three categories: global factors, local factors and occlusion.

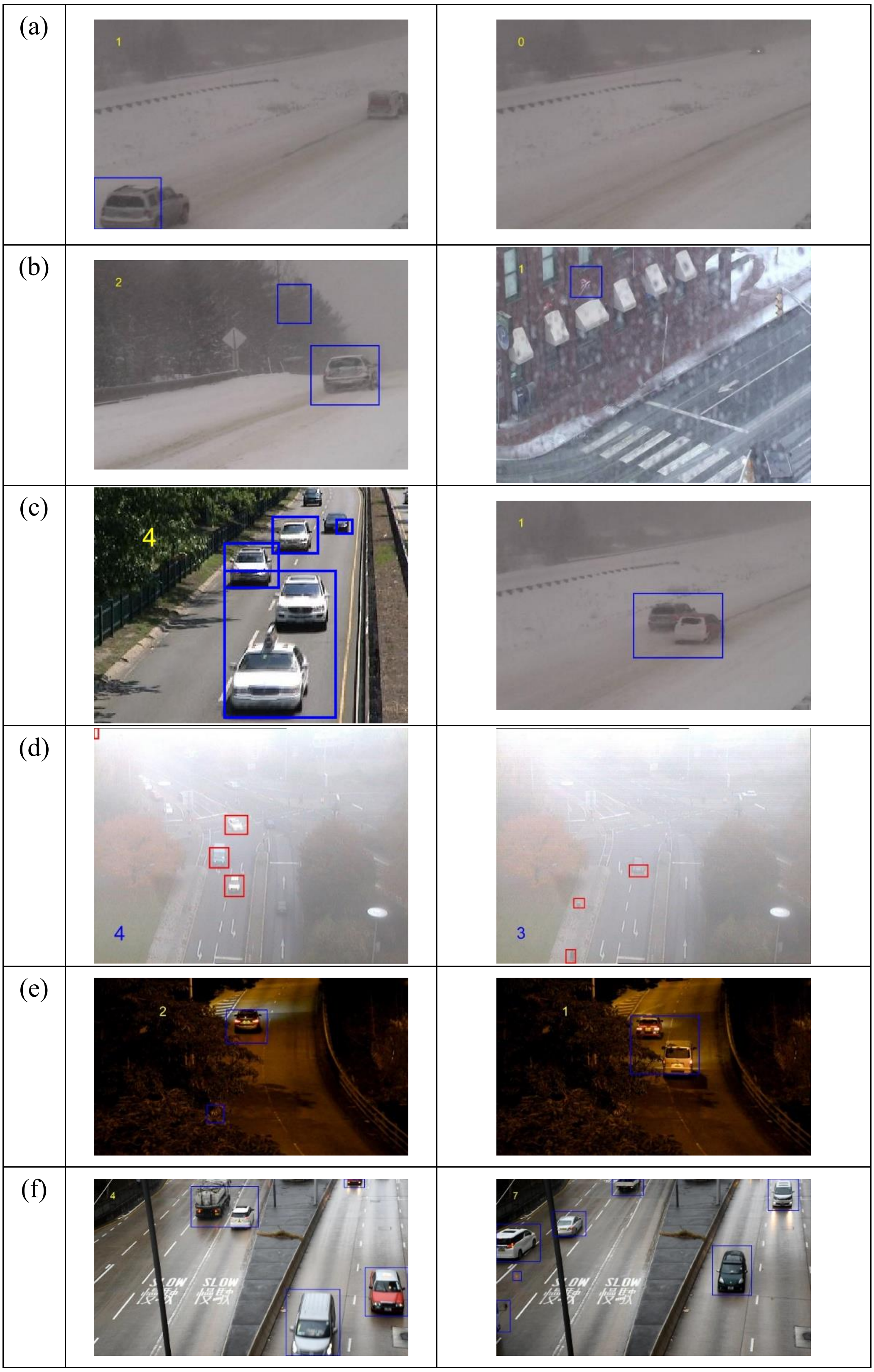


Figure 4-5. False detection under varying environments. (a) Blizzard; (b) Snowfall (left) and wet snow (right); (c) Highway (left) and blizzard (right); (d) Fog; (e) Nighttime_DSC_0376; (f) Rainy_DSC_0652;

Global Factors: In traffic surveillance videos, global factors mainly refer to the presence of fog and low illumination conditions such as dusk and nighttime. The presence of fog severely degrades the contrast of the image. After Histogram Equalization, the effect of fog can be reduced to a certain extent but not eliminated completely. At dusk or nighttime, moving vehicles cannot be detected accurately due to low visibility. Accordingly, some wrong ROIs are shown in Figure 4-5 (d). False detection caused by global factors can be reduced and eliminated by a total adjustment of image contrast or intensity. This step has been implemented in Section 3.2.3 using Histogram Equalization to restore the contrast of foggy images.

Local Factors: The presence of snowflakes, raindrops in the case of bad weather are major reasons for false detection. Accordingly, raindrops and snowflakes are wrongly identified as ROIs and counted as vehicle candidates. At nighttime images, swaying trees might also cause false detection (Figure 4-5 (e)). Impacts of local environmental factors can be reduced and eliminated by removal of small noise regions. In the case of urban conditions, walking pedestrian, falling leaves and waving flags could also be local factors that cause false detection.

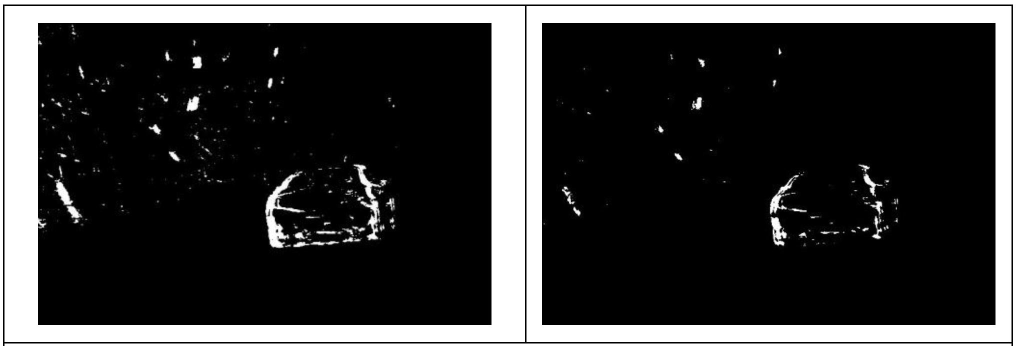
Occlusion: In urban environments, occlusion is a major reason for false detection. When two or more vehicles are driving near, the outline of vehicles might be vague due to the FOV of a fixed traffic surveillance camera. As a result, two or more vehicles are possibly identified and counted as one vehicle. Figure 4-4 (c) and (e) (right image) show three sample results caused by occlusion. At nighttime, on the highway, even in the case of the blizzard, occlusion occurred in the image sequences and resulted in false detection. Till now, occlusion is a very challenging problem, and there is no universal solution that can flexibly adapt to different conditions.

4.3 Comparison with Other Methods

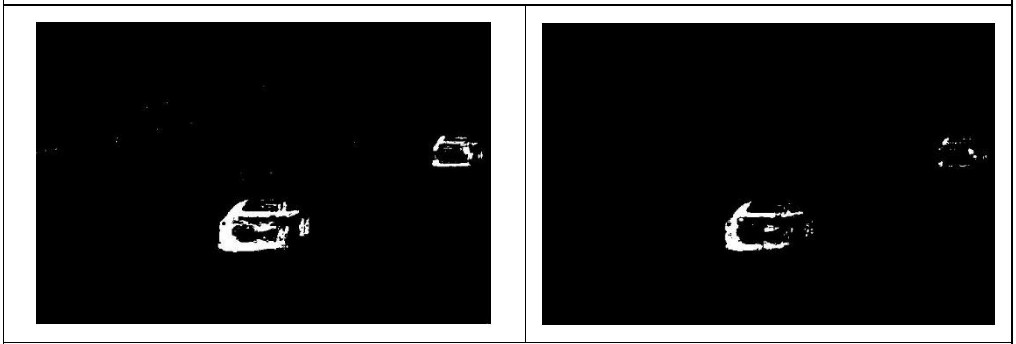
Here, the E-TFD detection and counting approach are compared with classical Two-Frame-Differencing method, then a typical Three-Frame-Differencing method. After that, the proposed statistical thresholding is compared with several state-of-the-art methods.

4.3.1 Two-Frame-Differencing versus Three-Frame-Differencing

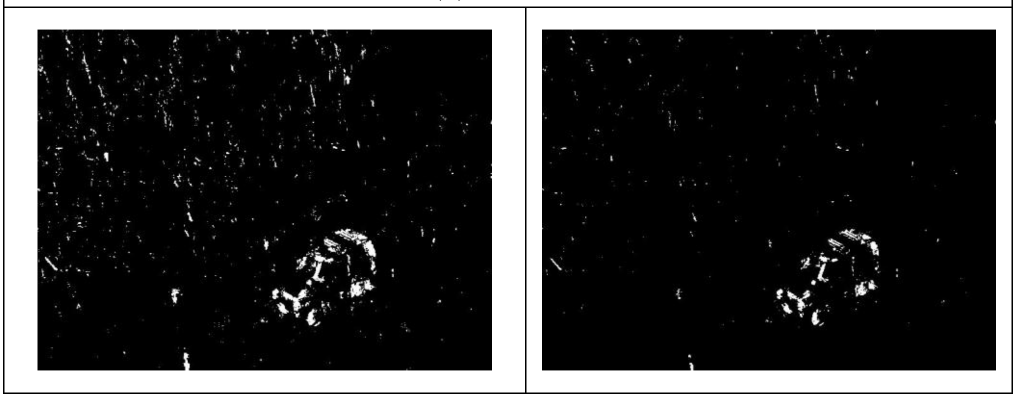
Inter-Frame-Differencing mainly differs in the choice of frame numbers. It can be two consecutive frames, three consecutive frames or three inconsecutive frames. First of all, we compare the enhanced Three-Frame-Differencing method with Two-Frame-Differencing (Jain and Nagel, 1979). Frame-Differencing images are generated using two frames and three frames as a comparison, followed by the proposed statistical thresholding method. It can be observed from Figure 4-6 that



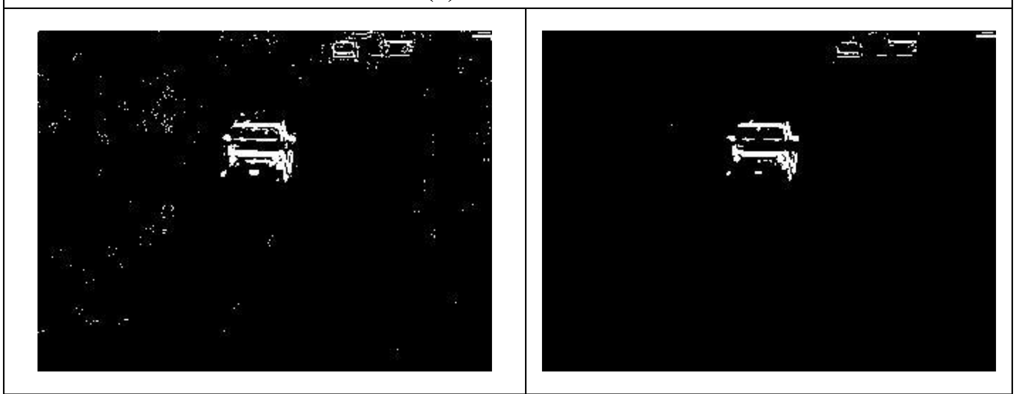
(a) Snowfall



(b) Blizzard



(c) Wet snow



(d) Highway

Figure 4-6. Frame-Differencing image generated by two frames (left columns) and three frames (right columns).

in both fine and challenging environmental conditions, three frames generate clearer outlines of moving vehicles with less image noise. The calculations in Table 4-4 also show that Three-Frame-Differencing generated less non-zero pixels in corresponding binary image maps. From the comparison of Two-Frame-Differencing and Three-Frame-Differencing, we get the conclusion that the enhanced TFD approach has better performance in extracting moving vehicles than conventional Two-Frame-Differencing methods.

Table 4-6. Calculation of pixels in Frame-Differencing.

Condition	**Number of Non-zero Pixels using 2FD**	**Number of Non-zero Pixels using 3FD**
Snowfall	7477	3033
Blizzard	6705	2969
Wet snow	8012	3414
Highway	1352	750

4.3.2 Three-Frame-Differencing versus The Enhanced Three-Frame-Differencing

As mentioned in Section 3.2.1, the enhanced Three-Frame-Differencing approach is comprised of 6 processing steps. It is based on typical Three-Frame-Differencing and some refinements are made. In this part, the E-TFD approach is compared with a typical Three-Frame-Differencing method.

As mentioned in some studies (Ji et al., 2016, Weng et al., 2010), a typical TFD detection approach follows the procedure of frame differencing, thresholding and morphological opening and closing.

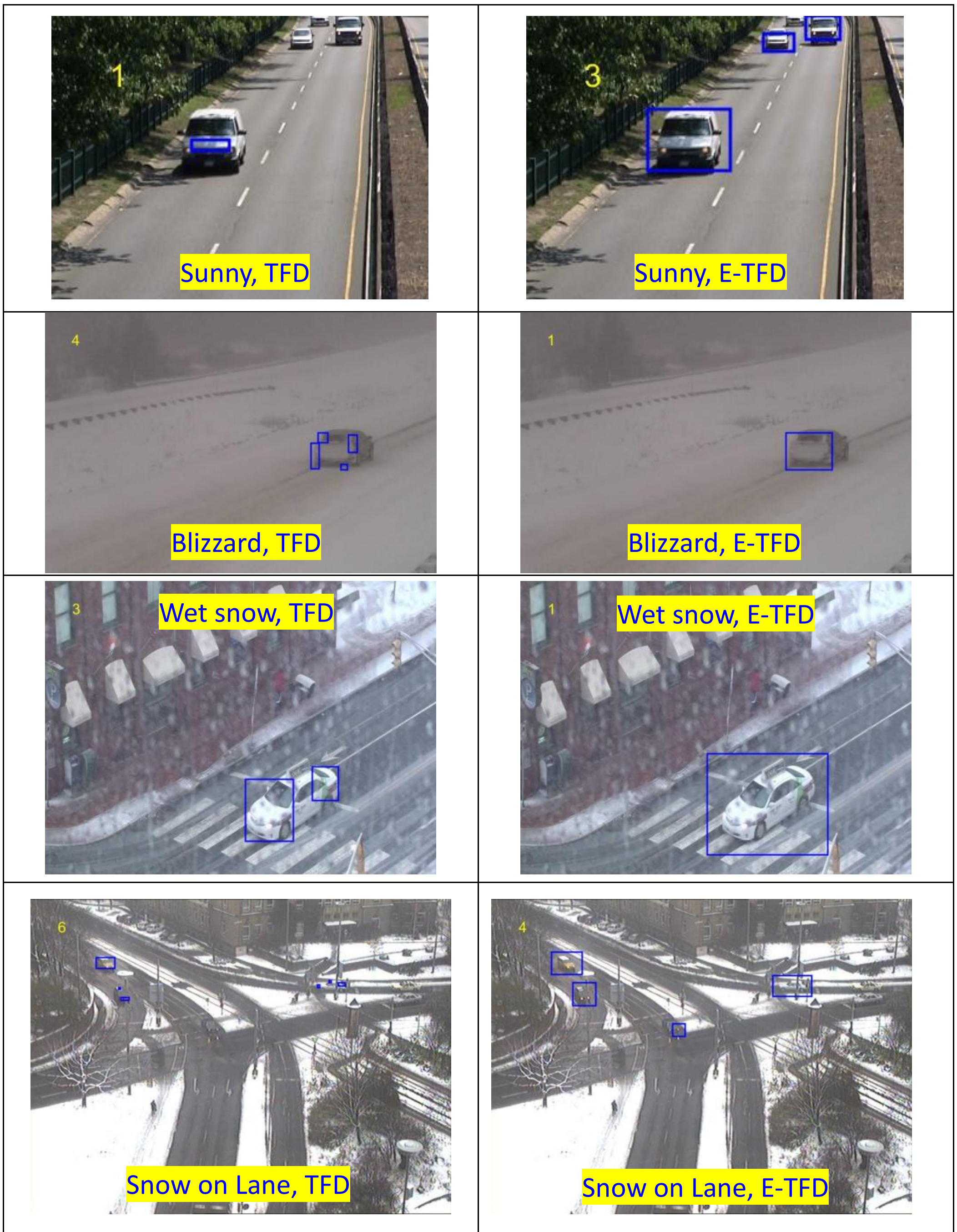


Figure 4-7 To be continued on next page

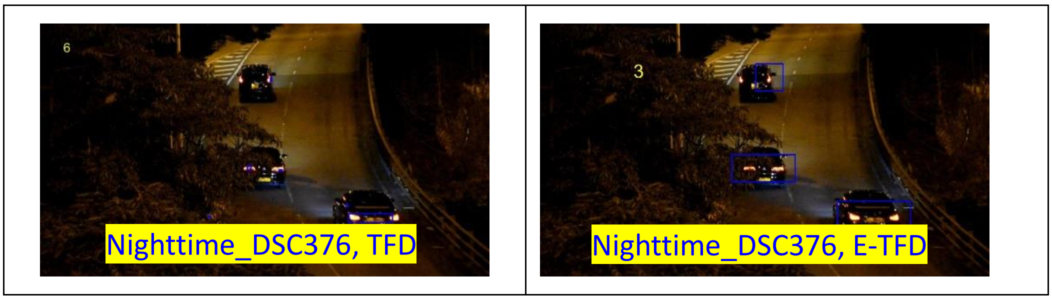


Figure 4-7. A comparison of detection and counting results generated from typical Three-Frame-Differencing (left columns) and the enhanced Three-Frame-Differencing (right columns).

From the comparison in Figure 4-7, it is obvious that the typical Three-Frame-Differencing approach does not have satisfactory performance in adverse environmental conditions. The tested results show that moving vehicles are partly detected and the numbers wrongly counted in sunny, blizzard and wet snow, snow-on-lane and nighttime images.

The main reason of the wrong generation of ROIs using typical Three-Frame-Differencing is that morphological opening cannot efficiently eliminate the noise of Frame Differencing images, comparing to the step of small noise regions removal in the E-TFD approach.

4.3.3 Thresholding Methods

On the other hand, different thresholding techniques in the existing literature are compared with the proposed statistical thresholding method. Method 1 and Method 2 utilize two statistics, the mean and standard deviation to determine the thresholding value. Method 3 is the Otsu thresholding method which calculates the optimal thresholding separating two classes (foreground and background) to make sure that intra-class variance is minimal.

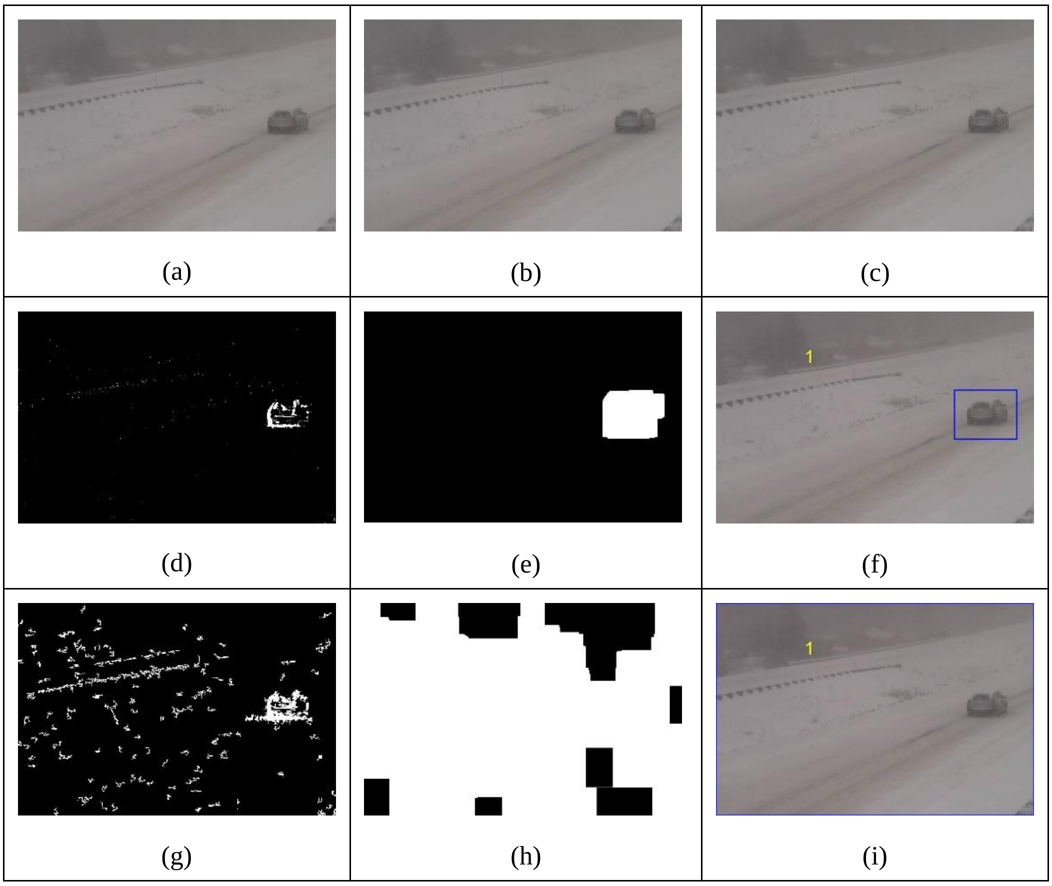


Figure 4-8. Comparison of statistical thresholding and Otsu thresholding in the case of Blizzard. From three consecutive image sequences (a), (b) and (c), binary image maps are generated using statistical thresholding (d) and Otsu thresholding (g), followed by results after morphological operations (e) and (h) and generated ROIs (f) and (i), respectively.

From a comparison of statistical thresholding and Otsu Thresholding (Figure 4-8) in blizzard images, it is observed that Otsu thresholding does not has satisfactory segmentation results. Comparing to Figure 4-8 (g), binary image map generated by statistical thresholding (Figure 4-8 (d)) segments moving vehicles with a clearer outline. Accordingly, ROI generated using statistical thresholding (Figure 4-8 (f)) captures the vehicle successfully.

Table 4-7. Comparison of different thresholding methods.

Method ID	References	Thresholding Method	Testing Results under Different Weather
Method 1	Lan et al. (2014)	$T = mean - std$	Fog: ⊗
			Snow on lane: ⊗
			Blizzard: ⊗
			Snowfall: ⊗
			Wet snow: ⊗
			Nighttime: ⊙
Method 2	Celik and Kusetogullari (2010)	$T = mean + std$	Fog: ⊗
			Snow on lane: ⊗
			Blizzard: ⊗
			Snowfall: ⊗
			Wet snow: ⊗
			Nighttime: ⊙
Method 3	Otsu (1979)	T = largest within-class variance	Fog: ⊗
			Snow on lane: ⊙
			Snow fall: ⊗
			Blizzard: ⊗
			Wet snow: ⊗
			Nighttime: ⊙
The proposed statistical thresholding	N/A	T = 99% of total intensity distributions	Fog:⊙
			Snow on lane: ⊙
			Blizzard: ⊙
			Snowfall: ⊙
			Wet snow: ⊙
			Nighttime: ⊙

⊙: Satisfactory

⊗: Unsatisfactory

Table 4-5 shows a comparison of different thresholding methods. After testing in six different conditions, the statistical thresholding performs well in segmenting vehicles of all cases, while method 1, 2 and 3 show unsatisfactory performance and cannot extract moving vehicles effectively.

Table 4-6 also summarizes different thresholding level generated by statistical thresholding and Otsu thresholding. In sunny images, statistical thresholding generates lower values for segmentation. In challenging environments, statistical thresholding (99%) generates higher value than Otsu thresholding and has better performance in foreground segmentation.

Table 4-8. Thresholding level selected by different methods.

Condition	**Statistical Thresholding**	**Otsu Thresholding**
Highway	30	65
Blizzard	7	2
Wet snow	23	9
Snowfall	24	10
Rainy_DSC_0652	78	43

4.4 Discussions and Improvements

Based on the performance of the E-TFD approach (Section 4.2) in different datasets, discussions are carried out in terms of obtained results, limitations of the study and benchmarking vehicular datasets.

4.4.1 Obtained Results

Recall Rate and Precision Rate

The performance of the E-TFD approach is evaluated in terms of accuracy and efficiency in Section 4.2. The recall rate tries to answer what proportions of extracted ROIs are correct. Based on the detection and counting results, an average recall rate of 87.1% can be achieved in all environmental conditions. The top three recall rate appeared in the case of self-collected nighttime DSC_0376 images (100%), blizzard images (99.41%) and rainy DSC_0652 images (99.3%). The E-TFD detection approach had the worst performance in foggy images, 919 ROIs were extracted out of 1661vehicles, with a detection rate of 55.3%.

Precision, as one of the ROC parameters, attempts to find what proportion of positive extractions are correct. The overall precision rate of the experiment is 92%. The top three precision rate in tested scenarios are 99.2% in blizzard images, 98.9% in nighttime (DSC_0376) images and 98.2% in rainy (DSC_0652) images. Despite the low detection rate, foggy images can achieve a precision of 94.6%.

In addition to the ROC parameters evaluation and running time calculation presented in Section 4.2, removal of walking pedestrian and vehicle tracking are discussed here as part of obtained results.

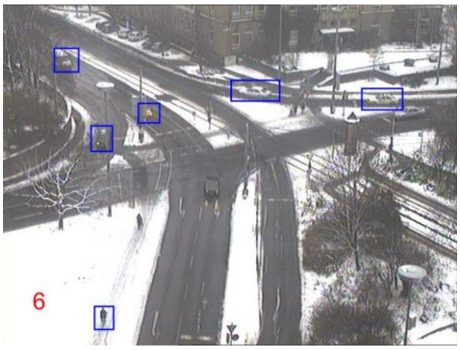
Removal of Walking Pedestrian

In many datasets, walking pedestrians were recorded together with moving vehicles. The presence of walking people in traffic scenes could be a great challenge in vision-based vehicle detection, as moving people might be wrongly detected as vehicles. To deal with this issue, we propose a simple method to eliminate the impact of moving pedestrian, thus improving the detection rate. Based on existing bounding boxes, we observe that the area of moving pedestrian is obviously smaller than moving vehicles. Therefore, a criterion is set to remove the walking pedestrian. Each image according to its scale is identified a smallest ROI for vehicle detection, less than which is considered as walking people and such ROI will be deleted from the bounding box list. For example, in wet snow images, if the area of the smallest ROI is less than 5050 pixels, the ROI will be deleted from the bounding box list. In the case of snow on lane, the threshold of small ROI is set as 1050 pixels.

Figure 4-9 shows an improved result of detection and counting after small noise region removal. After applying Equation (9) in Section 3.3.1, walking pedestrian detection in Figure 4-9 (a) and (c) can be eliminated effectively. At the same time, wrong counting of vehicles is corrected, shown in Figure 4-9 (b) and (d).

(a)

(b)



(c)

(d)

Figure 4-9. Detection results before and after removal of the walking pedestrian. Counting results with the walking pedestrian in wet snow and snow-on-lane image are shown in (a) and (c), with corrected counting results in (b) and (d).

Vehicle Tracking under Inclement Weather Conditions

The enhanced TFD method can not only efficiently detect moving vehicles but also track them in even very inclement weather conditions. Based on existing ROIs, centroids of vehicles can be extracted as well. Figure 4-10 and 4-11 show sample results of successful tracking in the blizzard and wet snow. Of all tested data in our experiments, the method can track the vehicle in very adverse conditions for up to 56 frames.

Figure 4-10. Successive tracking of a vehicle in blizzard.

Figure 4-11. Successive tracking of vehicles in wet snow image sequences.

4.4.2 Limitations of the Study

In Section 4.2.3, global factors, local factors and occlusion are considered as three major causes of false detection. Despite false detection, the main limitation of the E-TFD approach lies in that the average detection rate is not very high. Studies that utilize HOG features and SVM classifiers combinations show better performance in vehicle detection, where a detection rate of 98.6 % can be achieved

(Wei et al., 2018). Meanwhile, the combinations of Haar-like features and Adaboost classifiers can achieve a detection rate of 95.47 %, as stated by Wen et al. (2015). However, the two-stage vehicle detection approach suffered from very long training time and the time length can be more than 100 hours (Wen et al., 2015).

4.4.3 Benchmarking Vehicular Datasets

Benchmarking numerous vehicular data is an essential part of vision-based vehicle detection. DAS-related studies rely on Ground Truth (total number of vehicles in an image) to evaluate the performance of an algorithm. Appendix II and III provide a summary of publicly-available datasets that have been widely used in recent studies. Of all public-available datasets, only a few of them provide vehicular images in adverse conditions. At the same time, many studies still prefer to collect their own video data to train classifiers and test proposed algorithms. The fact that past work often reported results on nonpublic sequences made a fair comparison an elusive goal (Wang et al., 2014). Therefore, a systematic benchmark is required to further conduct research on transport monitoring under varying environments, especially in challenging conditions.

4.5 Summary

This section provides detection and counting results of the enhanced TFD method. First, the performance of the E-TFD detection approach is analyzed using ROC parameters. Extensive quantitative evaluation and related comparisons demonstrate that the enhanced TFD method can detect and count moving vehicles in real-time, yet able to achieve an average recall rate of 87.1%. Meanwhile, the E-TFD approach can also track vehicles in successive image frames.

Then, the problems that vehicle detection may encounter in challenging environments are discussed with corresponding solutions. Vehicle detection in adverse conditions suffers from some common challenges, such as the presence of fog, raindrops, snowflakes, walking pedestrian and falling leaves. In low illuminations conditions (rainy or nighttime images), strong reflections on the ground and the presence of taillights often lead to false detection. In congested urban traffic conditions or at intersections, occlusion might also cause false detection.

CHAPTER 5 CONCLUSIONS AND FUTURE WORK

In conclusion, we would like to highlight some recurring themes that cropped up repeatedly throughout the whole thesis. First, we summarized the contributions that have been made in this study. Then, we pointed out the deficiencies in this work that can be improved in the near future.

5.1 Contributions

First and foremost, we have made the following contributions in advancing the study of vehicle detection, counting and tracking under varying environments.

(1) We have provided a comprehensive review of vision-based vehicle detection approaches. The detection approaches are categorized as appearance-based methods and motion-based methods, based on which solutions under varying environments are provided (Yang and Pun-Cheng, 2018). Meanwhile, different traffic surveillance objectives that can be achieved base on vehicle detection are discussed in terms of major driving problems and derivation of traffic parameters.

(2) We have proposed an enhanced TFD approach for vehicle detection, counting and tracking in adverse conditions. Based on Three-Frame-Differencing (TFD), fast thresholding, removal of small noise regions and morphological operations, vehicle candidates in different environmental conditions can be extracted in an

efficient and accurate manner. The exact number of vehicles in each image can be counted and displayed on each image sequence. By and large, this study has provided a possible solution of vehicle detection, counting and tracking in all-weather conditions, which can be embedded into a real-time traffic surveillance system.

(3) We have tested the enhanced TFD detection approach in both public-available datasets and self-collected videos. Different weather conditions are tested, including sunny, rainy, foggy, snowy, wet snow, blizzard and nighttime images. In the experiment, we have tested over 11556 vehicles. Comparing to other studies that containing adverse conditions (Mu et al., 2016, Ershadi et al., 2018), the image frame numbers that we have tested are huge.

(4) From the quantitative evaluation of the E-TFD method, it is shown that the proposed detection approach can achieve an average detection rate of 87.1% of all cases. The highest detection rate was achieved in self-collected nighttime video DSC_0546, where a 100% detection rate can be achieved. We have provided the exact value of True Positive and False Positive, Precision and Recall in every condition, which has not been provided in other studies, especially in adverse conditions. The enhanced TFD method is computationally efficient. It can process approximately 100 frames per second on a desktop computer, which meets the

requirement of real-time processing. At the same time, the enhanced TFD detection approach can track successive frames in adverse weather conditions for up to 56 frames.

5.2 Future Work

5.2.1 Algorithmic Improvements

With the continuous development of computer vision technologies, significant enhancement in algorithms has been made. Classic approaches rely on robust local/ global features and efficient background models to extract driving-related information from traffic environments. In recent years, deep learning algorithms have been widely used in the field of Driver Assistance Systems, where car license plate recognition (Li et al., 2018), vehicle type recognition (Hu et al., 2017) and background subtraction (Babaee et al., 2017) have been explored using Neural Networks.

To evaluate an algorithm, accuracy is a significant factor. The common fact, however, is that false detection is inevitable in all studies. A negative result occurs when the outcome of an experiment or a model is not as what is expected (Borji, 2017). Obviously, computer vision cannot outperform the sensing ability of human eyes. In very inclement conditions, vision-based detection approaches may not work very accurately (e.g., vehicle detection, counting and tracking in foggy images), but still helps to understand different driving environments and prevent traffic accidents.

Comparing to the state-of-the-art methods, this study focuses on vehicle detection under challenging environments, which comprises a small part of the Driver Assistance System and Intelligent Transportation Systems. Based on existing detection and counting results, ROIs extracted in different environmental conditions can be cropped and used for training. In the future, we hope to integrate existing extracted ROIs with robust local descriptors such as HOG features, Haar-like features for further analysis and make the current detection architecture more robust. Meanwhile, the enhanced TFD approach can be combined with other background models to improve detection accuracy.

5.2.2 Implementation of Various Traffic Surveillance Objectives

The promising results of on-road experiments in this study demonstrate that a full implementation of all-time, all-weather traffic surveillance based on traffic surveillance cameras is totally feasible. With the auxiliary of robust feature descriptors and efficient models, real-time identification of vehicles can be achieved, at the same time multiple traffic surveillance objectives can be achieved, such as tracking, counting and classification of vehicles, derivation of driving speed, traffic flow analysis, etc.

In this thesis, we mainly focus on the implementation of vehicle detection in adverse conditions based on traffic surveillance cameras. Vehicle counting and tracking are also studied and tested using different datasets. In the future, we hope to extend the achieved objectives to a total implementation of on-road Driver Assistance Systems (DAS) with multiple functionalities, such as collision avoidance systems, brake assistance systems, adaptive cruise control and lane departure systems.

5.2.3 Ultimate Goal: All-time, All-weather Traffic Surveillance System

Since the 1980s, a long-term evolution of research in Driver Assistance Systems (DAS) has been witnessed. It took three decades for DAS to find the way from research to production (Bengler et al., 2014). But sadly, few examples of camera-based driver assistance systems have entered the automotive market (Ranft and Stiller, 2016). Meanwhile, despite very mature computer vision algorithms and a wide variety of applications, the automotive industry is facing the new challenge of developing a universal method for all-time, all-weather vehicle detection, especially in poor visibility conditions, such as fog, rain, snow and nighttime (Pinchon et al., 2016).

In this work, we have provided a possible solution of all-weather vehicle detection, counting and tracking that can be embedded into a traffic surveillance system. Several adverse conditions have been tested and analyzed. The results are found promising in providing recall rate in the case of fog, which has not been investigated in other studies.

As the compensation for human error, the ultimate goal of Intelligent Transportation System, in all, is to achieve the best possible performance of detection, yet able to inform road users and traffic managers of real-time driving status under different conditions, especially challenging environments.

Glossary

Abbreviation	Full Name
ADAS	Automotive Driver Assistance System
CDnet	Change Detection
E-TFD	Enhanced Three-Frame-Differencing
FN	False Negative
FOV	Field of View
FP	False Positive
GMM	Gaussian Mixture Model
GT	Ground Truth
HE	Histogram Equalization
HG	Hypothesis Generation
HOG	Histogram of Gradients
HV	Hypothesis Verification
ITS	Intelligent Transportation System
KIT	Kalsruhe Institute of Technology
NN	Neural Networks
ROC	Receiver Operating Characteristics
ROI	Region of Interest

ROT	Region of Taillights
SIFT	Scale-Invariant Feature Transformation
SURF	Speeded Up Robust Features
SVM	Support Vector Machine
TN	True Negative
TP	Ture Positive

Appendixes

Appendix I. Selected Work on vehicle detection in adverse weather conditions.

Reference	Methodology	Weather conditions	Detection rate included?	Number of tested samples provided?
Tsai et al. (2007)	Appearance-based	Cloudy, rainy	Yes	Yes
Mu et al. (2016)	Appearance-based	Cloudy, rainy, foggy, snowy	Yes	Yes
Jia et al. (2016)	Appearance-based	Rainy, foggy	No	No
Sun et al. (2006a)	Appearance-based	Rainy	No	No
Shen (2007)	Appearance-based	Cloudy, rainy, misty	Yes	Yes
Chen and Peng (2012)	Appearance-based	Rainy (nighttime)	Yes	Yes
Rabbouch et al. (2017)	Motion-based	Snowy, foggy	No	No
Varadarajan et al. (2015)	Motion-based	Blizzard	No	No

Appendix II. Vehicle Detection Benchmarks with Special Focus on Different Weather Conditions.

Weather conditions	Reference	Dataset	Category		Number of frames	Field of View	Description
Fine	Goyette et al. (2012)	Change Detection 2012	Baseline	highway	1700	Front	The sequences present vehicle-related scenes in fine weather scenarios.
			Camera Jitter	boulevard	2500	Side	
				Traffic	1570	Side rear	
			Dynamic background	fall	4000	Side	
				boats	7999	Side	
			Intermittent object motion	Abandoned box	4500	Multi-view	
				Street light	3200	Side	
				Tram stop	3200	Side	
			Shadow	bungalows	1700	Front/rear	
	Wang et al. (2014)	Change Detection 2014	Low Frame rate	Tram crossroad	900	Intersection	
			PTZ	Intermittent pan	3500	Side	
Rain	Wang et al. (2014)	Change Detection 2014	PTZ	Two position	2300	Multi-view, intersection	This category provides vehicle-related images

				PTZ camera			on rainy days.
Fog	Karlsruhe (1997)	Karlsruhe Institute of Technology	Dtneu_nebel	Heavy fog	290	Up-down, intersections	This category includes image sequences taken on foggy days. The camera is mounted high, so vehicles look tiny.
Snow	Goyette et al. (2012)	Change Detection 2012	Intermittent object motion	Winter driveway	2500	Front/side	The 'bad weather' and 'intermittent object motion' category provides sequences taken on snowy days. Due to the weather, backgrounds with white color.
	Wang et al. (2014)	Change Detection 2014	Bad weather	Blizzard	7000	Side rear	
				Snowfall	6500	Side rear	
				Wet snow	3500	Side front	
	Karlsruhe (1997)	Karlsruhe Institute of Technology	dtneu_schnee	Heavy snowfall	300	Up-down, intersection	The two categories show scenes at intersections on snowy winter days.
			dtneu_winter	Snow on lanes	300	Up-down, intersection	

Appendix III. Vehicle Detection Benchmarks with Special Focus on Different Illumination Conditions.

Illumination conditions	Reference	Dataset	Category		Number of frames	Field of View	Description
Low	Wang et al. (2014)	Change Detection 2014	Low Frame rate	Tunnel exists	4000	Front	The camera captured vehicles moving out of a tunnel with low illumination
Nighttime	Wang et al. (2014)	Change Detection 2014	Night videos	Bridge entry	2499	Multi-view	The 'night videos' category provides on-road night scene. Due to the low illumination at nighttime, image sequences show different levels of vagueness.
				Busy boulvard	2760	Intersection	
				Fluid highway	1364	Front/rear	
				Street corner at night	5200	Side front/rear	
				Tram station	3000	Side	
				Winter street	1785	Multi-view	
	Chen (2014)	SYSU Vehicle Detection dataset	Nighttime		5575	Rear	This category contains nighttime vehicles driving on the urban road at nighttime.
Special illumination	Jensen et al. (2016)	LISA Traffic Light Dataset	Vid 2		492	Rear	The image sequences contain videos with bright background, which can be considered as special illumination.
			Vid 5		540	Rear	

Bibliography

Abolghasemi, V. & Ahmadyfard, A. 2009. An edge-based color-aided method for license plate detection. *Image and Vision Computing,* 27**,** 1134-1142.

Acunzo, D., Zhu, Y., Xie, B. & Baratoff, G. 2007. Context-adaptive approach for vehicle detection under varying lighting conditions. In *2007 Intelligent Transportation Systems Conference*. Bellevue: IEEE, 654-660.

Adams, R. & Bischof, L. 1994. Seeded region growing. *IEEE Transactions on pattern analysis and machine intelligence,* 16**,** 641-647.

Adelson, E. H., Anderson, C. H., Bergen, J. R., Burt, P. J. & Ogden, J. M. 1984. Pyramid methods in image processing. *RCA engineer,* 29**,** 33-41.

Almagambetov, A., Casares, M. & Velipasalar, S. 2012. Autonomous tracking of vehicle rear lights and detection of brakes and turn signals. In *2012 IEEE Symposium on Computational Intelligence for Security and Defence Applications (CISDA)*. Ottawa: IEEE, 1-7.

Alonso, D., Salgado, L. & Nieto, M. 2007. Robust vehicle detection through multidimensional classification for on board video based systems. In *IEEE International Conference on Image Processing*. San Antonio: IEEE, 321-324.

Anandhalli, M. & Baligar, V. P. 2017. A novel approach in real-time vehicle detection and tracking using Raspberry Pi. *Alexandria Engineering*

Journal, 57**,** 1597-1607.

Arróspide, J., Salgado, L. & Camplani, M. 2013. Image-based on-road vehicle detection using cost-effective histograms of oriented gradients. *Journal of Visual Communication and Image Representation,* 24**,** 1182-1190.

Asaidi, H., Aarab, A. & Bellouki, M. 2014. Shadow elimination and vehicles classification approaches in traffic video surveillance context. *Journal of Visual Languages & Computing,* 25**,** 333-345.

Asvadi, A., Garrote, L., Premebida, C., Peixoto, P. & Nunes, U. J. 2017. Multimodal vehicle detection: fusing 3D-LIDAR and color camera data. *Pattern Recognition Letters***,** 20-29.

Babaee, M., Dinh, D. T. & Rigoll, G. 2017. A deep convolutional neural network for background subtraction. *arXiv preprint arXiv:1702.01731*.

Barcellos, P., Bouvié, C., Escouto, F. L. & Scharcanski, J. 2015. A novel video based system for detecting and counting vehicles at user-defined virtual loops. *Expert Systems with Applications,* 42**,** 1845-1856.

Barnich, O. & Van Droogenbroeck, M. 2011. ViBe: A universal background subtraction algorithm for video sequences. *IEEE Transactions on Image processing,* 20**,** 1709-1724.

Bas, E., Tekalp, A. M. & Salman, F. S. 2007. Automatic vehicle counting from video for traffic flow analysis. In *2007 IEEE Intelligent Vehicles*

Symposium. Istanbul: IEEE, 392-397.

Batavia, P. H., Pomerleau, D. & Thorpe, C. E. 1997. Overtaking vehicle detection using implicit optical flow. In *Proceedings of Conference on Intelligent Transportation Systems*. Boston: IEEE, 729-734.

Bay, H., Tuytelaars, T. & Van Gool, L. 2006. Surf: Speeded up robust features. In *European conference on computer vision*. Graz: Springer, 404-417.

Bengler, K., Dietmayer, K., Farber, B., Maurer, M., Stiller, C. & Winner, H. 2014. Three decades of driver assistance systems: Review and future perspectives. *IEEE Intelligent Transportation Systems Magazine,* 6**,** 6-22.

Bertozzi, M., Broggi, A. & Castelluccio, S. 1997. A real-time oriented system for vehicle detection. *Journal of Systems Architecture,* 43**,** 317-325.

Bertozzi, M., Broggi, A. & Fascioli, A. 2000. Vision-based intelligent vehicles: State of the art and perspectives. *Robotics and Autonomous systems,* 32**,** 1-16.

Betke, M., Haritaoglu, E. & Davis, L. S. 2000. Real-time multiple vehicle detection and tracking from a moving vehicle. *Machine vision and applications,* 12**,** 69-83.

Bishop, C. M. 1995. *Neural networks for pattern recognition*, Oxford university press.

Blanc, N., Steux, B. & Hinz, T. 2007. LaRASideCam: A fast and robust vision-based blindspot detection system. In *2007 IEEEIntelligent Vehicles Symposium*. Istanbul: IEEE, 480-485.

BMVA. 2016. *What is Computer vision?* [Online]. Available: http://www.bmva.org/visionoverview [Accessed 2018].

Borji, A. 2017. Negative Results in Computer Vision: A Perspective. *arXiv preprint arXiv:1705.04402*.

Bronte, S., Bergasa, L. M. & Alcantarilla, P. F. 2009. Fog detection system based on computer vision techniques. In *12th International IEEE Conference on Intelligent Transportation Systems*. St. Louis: IEEE, 30-35.

Buch, N., Velastin, S. A. & Orwell, J. 2011. A review of computer vision techniques for the analysis of urban traffic. *IEEE Transactions on Intelligent Transportation Systems,* 12**,** 920-939.

Busch, C. & Debes, E. 1998. Wavelet transform for analyzing fog visibility. *IEEE Intelligent Systems and Their Applications,* 13**,** 66-71.

Cabani, I., Toulminet, G. & Bensrhair, A. 2005. Color-based detection of vehicle lights. In *2005 IEEE Intelligent Vehicles Symposium Proceedings*. Las Vegas: IEEE, 278-283.

Canny, J. 1986. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence***,** 679-698.

Celik, T. & Kusetogullari, H. 2010. Solar-powered automated road surveillance system for speed violation detection. *IEEE Transactions on Industrial Electronics,* 57**,** 3216-3227.

Chan, Y.-M., Huang, S.-S., Fu, L.-C. & Hsiao, P.-Y. 2007. Vehicle detection under various lighting conditions by incorporating particle filter. In *2007 IEEE Intelligent Transportation Systems Conference*. Washington: IEEE, 534-539.

Chen, D.-Y., Lin, Y.-H. & Peng, Y.-J. 2012. Nighttime brake-light detection by Nakagami imaging. *IEEE Transactions on Intelligent Transportation Systems,* 13**,** 1627-1637.

Chen, D.-Y. & Peng, Y.-J. 2012. Frequency-tuned taillight-based nighttime vehicle braking warning system. *IEEE Sensors Journal,* 12**,** 3285-3292.

Chen, H.-T., Wu, Y.-C. & Hsu, C.-C. 2016. Daytime preceding vehicle brake light detection using monocular vision. *IEEE Sensors Journal,* 16**,** 120-131.

Chen, L. 2014. *Sun Yat-sen University Night-time Vehicle Dataset* [Online]. Available: http://www.carlib.net/?page_id=35 [Accessed].

Cheon, M., Lee, W., Yoon, C. & Park, M. 2012. Vision-based vehicle detection system with consideration of the detecting location. *IEEE transactions on intelligent transportation systems,* 13**,** 1243-1252.

Cheung, S., Coleri, S., Dundar, B., Ganesh, S., Tan, C.-W. & Varaiya, P. 2005. Traffic measurement and vehicle classification with single magnetic sensor.

Transportation research record: journal of the transportation research board, 1917(1)**,** 173-181.

Chong, Y., Chen, W., Li, Z., Lam, W. H., Zheng, C. & Li, Q. 2013. Integrated real-time vision-based preceding vehicle detection in urban roads. *Neurocomputing,* 116**,** 144-149.

Collado, J. M., Hilario, C., De la Escalera, A. & Armingol, J. M. 2004. Model based vehicle detection for intelligent vehicles. In *2004 IEEE Intelligent Vehicles Symposium*. Parma: IEEE, 572-577.

Cortes, C. & Vapnik, V. 1995. Support-vector networks. *Machine learning,* 20**,** 273-297.

Cucchiara, R. & Piccardi, M. 1999. Vehicle Detection under Day and Night Illumination. In *Proceedings of the 3rd International ICSC Symposium on Intelligent Industrial Automation*. Genova.

Cucchiara, R., Piccardi, M. & Mello, P. 2000. Image analysis and rule-based reasoning for a traffic monitoring system. *IEEE Transactions on Intelligent Transportation Systems,* 1**,** 119-130.

Dalal, N. & Triggs, B. 2005. Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference onComputer Vision and Pattern (CVPR'05)*. San Diego: IEEE, 886-893.

Department, T. 2017. *Live webcast of traffic condition* [Online]. Available:

http://traffic.td.gov.hk/ [Accessed].

Doretto, G., Chiuso, A., Wu, Y. N. & Soatto, S. 2003. Dynamic textures. *International Journal of Computer Vision,* 51, 91-109.

Ershadi, N. Y., Menéndez, J. M. & Jiménez, D. 2018. Robust vehicle detection in different weather conditions: Using MIPM. *PloS one,* 13, e0191355.

Eum, S. & Jung, H. G. 2013. Enhancing light blob detection for intelligent headlight control using lane detection. *IEEE transactions on intelligent transportation systems,* 14, 1003-1011.

Fan, Q., Brown, L. & Smith, J. 2016. A closer look at Faster R-CNN for vehicle detection. In *2016 IEEE Intelligent Vehicles Symposium (IV)*. Gothenburg: IEEE, 124-129.

Fawcett, T. 2006. An introduction to ROC analysis. *Pattern recognition letters,* 27, 861-874.

Fe-Fei, L. 2003. A Bayesian approach to unsupervised one-shot learning of object categories. In *Ninth IEEE International Conference on Computer Vision*. Nice: IEEE, 1134-1141.

Freund, Y. & Schapire, R. E. 1997. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences,* 55, 119-139.

Fu, H., Ma, H., Liu, Y. & Lu, D. 2016. A vehicle classification system based on hierarchical multi-SVMs in crowded traffic scenes. *Neurocomputing,* 211, 182-190.

Gallen, R., Cord, A., Hautière, N. & Aubert, D. 2011. Towards night fog detection through use of in-vehicle multipurpose cameras. In *2011 IEEE Intelligent Vehicles Symposium (IV)*. Baden-Baden: IEEE, 399-404.

Garcia, F., Cerri, P., Broggi, A., de la Escalera, A. & Armingol, J. M. 2012. Data fusion for overtaking vehicle detection based on radar and optical flow. In *2012 IEEE Intelligent Vehicles Symposium (IV)*. Alcalá de Henares: IEEE, 494-499.

Goyette, N., Jodoin, P.-M., Porikli, F., Konrad, J. & Ishwar, P. 2012. Changedetection. net: A new change detection benchmark dataset. In *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Providence: IEEE, 1-8.

Harris, C. & Stephens, M. 1988. A combined corner and edge detector. In *Alvey vision conference*. Manchester: University of Sheffield Printing Office, 10.5244.

Haselhoff, A. & Kummert, A. 2009. A vehicle detection system based on haar and triangle features. In *2009 IEEE Intelligent Vehicles Symposium*. Xi'an: IEEE, 261-266.

He, K., Sun, J. & Tang, X. 2011. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence,* 33**,** 2341-2353.

Horprasert, T., Harwood, D. & Davis, L. S. 1999. A statistical approach for real-time robust background subtraction and shadow detection. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*. Kerkyra: Citeseer, 1-19.

Hu, B., Lai, J.-H. & Guo, C.-C. 2017. Location-aware fine-grained vehicle type recognition using multi-task deep networks. *Neurocomputing,* 243**,** 60-68.

Huang, D.-Y., Chen, C.-H., Hu, W.-C. & Su, S.-S. 2012. Reliable moving vehicle detection based on the filtering of swinging tree leaves and raindrops. *Journal of Visual Communication and Image Representation,* 23**,** 648-664.

Hummel, R. 1977. Image enhancement by histogram transformation. *Computer graphics and image processing,* 6**,** 184-195.

Hunter, R. S. 1958. Photoelectric color difference meter. *Josa,* 48**,** 985-995.

Ito, T., Yamada, K. & Nishioka, K. 1995. Understanding driving situations using a network model. In *Proceedings of the Intelligent Vehicles' 95 Symposium*. Detroit: IEEE, 48-53.

Iwasaki, Y. & Kurogi, Y. 2007. Real-time robust vehicle detection through the same algorithm both day and night. In *International Conference on Wavelet Analysis and Pattern Recognition*. Beijing: IEEE, 1008-1014.

Jain, R. & Nagel, H.-H. 1979. On the analysis of accumulative difference pictures from image sequences of real world scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 206-214.

Jazayeri, A., Cai, H., Zheng, J. Y. & Tuceryan, M. 2011. Vehicle detection and tracking in car video based on motion model. *IEEE Transactions on Intelligent Transportation Systems,* 12, 583-595.

Jensen, M. B., Philipsen, M. P., Møgelmose, A., Moeslund, T. B. & Trivedi, M. M. 2016. Vision for looking at traffic lights: Issues, survey, and perspectives. *IEEE Transactions on Intelligent Transportation Systems,* 17, 1800-1815.

Ji, W., Tang, L., Li, D., Yang, W. & Liao, Q. 2016. Video-based construction vehicles detection and its application in intelligent monitoring system. *CAAI Transactions on Intelligence Technology,* 1, 162-172.

Jia, D., Zhu, H., Zou, S. & Huang, K. 2016. Recognition method based on Green Associative Mechanism for weak contrast vehicle targets. *Neurocomputing,* 203, 1-11.

Karlsruhe, U. o. 1997. *Image sequence server* [Online]. Available: http://i21www.ira.uka.de/image_sequences/ [Accessed].

Khairdoost, N., Monadjemi, S. A. & Jamshidi, K. 2013. Front and rear vehicle detection using hypothesis generation and verification. *Signal & Image Processing,* 4, 31.

Ki, Y.-K. & Baik, D.-K. 2006. Model for accurate speed measurement using double-loop detectors. *IEEE Transactions on Vehicular Technology,* 55**,** 1094-1101.

Kim, J.-H., Sim, J.-Y. & Kim, C.-S. 2015. Video deraining and desnowing using temporal correlation and low-rank matrix completion. *IEEE Transactions on Image Processing,* 24**,** 2658-2670.

Kim, J. B. & Kim, H. J. 2003. Efficient region-based motion segmentation for a video monitoring system. *Pattern Recognition Letters,* 24**,** 113-128.

Kim, K., Chalidabhongse, T. H., Harwood, D. & Davis, L. 2004. Background modeling and subtraction by codebook construction. In *International Conference onImage Processing*. Singapore: IEEE, 3061-3064.

Kim, S., Oh, S.-Y., Kang, J., Ryu, Y., Kim, K., Park, S.-C. & Park, K. 2005. Front and rear vehicle detection and tracking in the day and night times using vision and sonar sensor fusion. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Edmonton: IEEE, 2173-2178.

Kuo, Y.-C., Pai, N.-S. & Li, Y.-F. 2011. Vision-based vehicle detection for a driver assistance system. *Computers & Mathematics with Applications,* 61**,** 2096-2100.

Lan, J., Li, J., Hu, G., Ran, B. & Wang, L. 2014. Vehicle speed measurement based on gray constraint optical flow algorithm. *Optik-International Journal for Light and Electron Optics,* 125**,** 289-295.

Li, H., Wang, P., You, M. & Shen, C. 2018. Reading Car License Plates Using Deep Neural Networks. *Image and Vision Computing,* 72**,** 14-23.

Li, J. & Allinson, N. M. 2008. A comprehensive review of current local features for computer vision. *Neurocomputing,* 71**,** 1771-1787.

Li, Y. & Wang, F.-Y. 2015. Vehicle detection based on And–Or Graph and Hybrid Image Templates for complex urban traffic conditions. *Transportation Research Part C: Emerging Technologies,* 51**,** 19-28.

Lin, B.-F., Chan, Y.-M., Fu, L.-C., Hsiao, P.-Y., Chuang, L.-A., Huang, S.-S. & Lo, M.-F. 2012. Integrating appearance and edge features for sedan vehicle detection in the blind-spot area. *IEEE Transactions on Intelligent Transportation Systems,* 13**,** 737-747.

Liu, W., Wen, X., Duan, B., Yuan, H. & Wang, N. 2007. Rear vehicle detection and tracking for lane change assist. In *2007 IEEE Intelligent Vehicles Symposium*. Istanbul: IEEE, 252-257.

Liu, X., Wang, Z., Feng, J. & Xi, H. 2016. Highway vehicle counting in compressed domain. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, 3016-3024.

Liu, Y., Lu, Y., Shi, Q. & Ding, J. 2013. Optical flow based urban road vehicle tracking. In *Ninth International Conference onComputational Intelligence and Security (CIS)*. Leshan: IEEE, 391-395.

López-Rubio, F. J. & López-Rubio, E. 2015. Local color transformation analysis for sudden illumination change detection. *Image and Vision Computing,* 37**,** 31-47.

Lowe, D. G. 1999. Object recognition from local scale-invariant features. In *The proceedings of the seventh IEEE international conference on Computer vision*. Kerkyra: IEEE, 1150-1157.

Luvizon, D. C., Nassu, B. T. & Minetto, R. 2017. A video-based system for vehicle speed measurement in urban roadways. *IEEE Transactions on Intelligent Transportation Systems,* 18**,** 1393-1404.

Madasu, V. K. & Hanmandlu, M. 2010. Estimation of vehicle speed by motion tracking on image sequences. In *2010 IEEE Intelligent Vehicles Symposium (IV)*. San Diego: IEEE, 185-190.

Martínez, J., Santos, M. & López, V. 2014. Photograph's Exposure Control by Fuzzy Logic. *Practical Applications of Intelligent Systems.* Springer.

McCartney, E. J. 1976. Optics of the atmosphere: scattering by molecules and particles. *New York, John Wiley and Sons, Inc.*

Mikic, I., Cosman, P. C., Kogut, G. T. & Trivedi, M. M. 2000. Moving shadow and object detection in traffic scenes. In *Proceedings 15th International*

Conference on Pattern Recognition. Barcelona: IEEE, 321-324.

Ming, Q. & Jo, K.-H. 2011. Vehicle detection using tail light segmentation. In *Proceedings of the 6th International Forum on Strategic Technology (IFOST)*. Harbin: IEEE, 729-732.

Mithun, N. C., Howlader, T. & Rahman, S. M. 2016. Video-based tracking of vehicles using multiple time-spatial images. *Expert Systems with Applications,* 62**,** 17-31.

Mu, K., Hui, F., Zhao, X. & Prehofer, C. 2016. Multiscale edge fusion for vehicle detection based on difference of Gaussian. *Optik-International Journal for Light and Electron Optics,* 127**,** 4794-4798.

Mukhtar, A., Xia, L. & Tang, T. B. 2015. Vehicle detection techniques for collision avoidance systems: A review. *IEEE Transactions on Intelligent Transportation Systems,* 16**,** 2318-2338.

Naba, A., Pratama, B. M., Nadhir, A. & Harsono, H. 2016. Haar-like feature based real-time neuro car detection system. In *International Seminar on Sensors, Instrumentation, Measurement and Metrology (ISSIMM)*. Malang: IEEE, 67-70.

Narasimhan, S. G. & Nayar, S. K. 2003. Contrast restoration of weather degraded images. *IEEE transactions on pattern analysis and machine intelligence,* 25**,** 713-724.

Nayar, S. K. & Narasimhan, S. G. 1999. Vision in bad weather. In *The Proceedings of the Seventh IEEE International Conference on Computer Vision*. Kerkyra: IEEE, 820-827.

Negri, P., Clady, X., Hanif, S. M. & Prevost, L. 2008. A cascade of boosted generative and discriminative classifiers for vehicle detection. *EURASIP Journal on Advances in Signal Processing,* 2008**,** 136.

Nguyen, T. T., Dai Pham, X., Song, J. H., Jin, S., Kim, D. & Jeon, J. W. 2011. Compensating background for noise due to camera vibration in uncalibrated-camera-based vehicle speed measurement system. *IEEE Transactions on Vehicular Technology,* 60**,** 30-43.

O'malley, R., Glavin, M. & Jones, E. 2011. Vision-based detection and tracking of vehicles to the rear with perspective correction in low-light conditions. *IET Intelligent Transport Systems,* 5**,** 1-10.

O'Malley, R., Jones, E. & Glavin, M. 2010. Rear-lamp vehicle detection and tracking in low-exposure color video for night conditions. *IEEE Transactions on Intelligent Transportation Systems,* 11**,** 453-462.

O’Malley, R., Glavin, M. & Jones, E. 2008. Vehicle detection at night based on tail-light detection. In *1st international symposium on vehicular computing systems*. Dublin.

Otsu, N. 1979. A threshold selection method from gray-level histograms. *IEEE* *transactions on systems, man, and cybernetics,* 9, 62-66.

Padmini, T. & Shankar, T. 2016. A Review On Visibility Restoration Of Degraded Images Under Inclement Weather Conditions. *APRN Journal of Engineering and Applied Sciences,* 11, 12858-12866.

PARLIAMENT, T. & UNION, T. 2010. Directive 2010/40/eu of the european parliament and of the council. *Official Journal of the European Union,* 50, 207.

Parodi, P. & Piccioli, G. 1995. A feature-based recognition scheme for traffic scenes. In *Proceedings of the Intelligent Vehicles' 95 Symposium*. Detroit: IEEE, 229-234.

Pavlic, M., Rigoll, G. & Ilic, S. 2013. Classification of images in fog and fog-free scenes for use in vehicles. In *2013 IEEE Intelligent Vehicles Symposium (IV)*. Gold Coast: IEEE, 481-486.

Pinchon, N., Ibn-Khedher, M., Cassignol, O., Nicolas, A., Bernardin, F., Leduc, P., Tarel, J. P., Bremond, R., Bercier, E. & Julien, G. 2016. All-weather vision for automotive safety: which spectral band? In *SIA Vision 2016-International Conference Night Drive Tests and Exhibition*. Paris: SIA, 7p.

Prati, A., Mikic, I., Trivedi, M. M. & Cucchiara, R. 2003. Detecting moving shadows: algorithms and evaluation. *IEEE transactions on pattern analysis and machine intelligence,* 25, 918-923.

Prewitt, J. M. 1970. Object enhancement and extraction. *Picture processing and Psychopictorics,* 10**,** 15-19.

Ra, M., Jung, H. G., Suhr, J. K. & Kim, W.-Y. 2018. Part-based Vehicle Detection in Side-rectilinear Images for Blind-Spot Detection. *Expert Systems with Applications*.

Rabbouch, H., Saâdaoui, F. & Mraihi, R. 2017. Unsupervised video summarization using cluster analysis for automatic vehicles counting and recognizing. *Neurocomputing***,** 157-173.

Ranft, B. & Stiller, C. 2016. The role of machine vision for intelligent vehicles. *IEEE Transactions on Intelligent Vehicles,* 1**,** 8-19.

Richter, E., Schubert, R. & Wanielik, G. 2008. Radar and vision based data fusion-advanced filtering techniques for a multi object vehicle tracking system. In *Intelligent Vehicles Symposium, 2008 IEEE*. Eindhoven: IEEE, 120-125.

Rybski, P. E., Huber, D., Morris, D. D. & Hoffman, R. 2010. Visual classification of coarse vehicle orientation using histogram of oriented gradients features. In *2010 IEEE Intelligent Vehicles Symposium*. San Diego: IEEE, 921-928.

Schamm, T., von Carlowitz, C. & Zöllner, J. M. 2010. On-road vehicle detection during dusk and at night. In *2010 IEEE Intelligent Vehicles Symposium*. San Diego: IEEE, 418-423.

Seki, M., Wada, T., Fujiwara, H. & Sumi, K. 2003. Background subtraction based on cooccurrence of image variations. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Madison: IEEE, II-II.

Semertzidis, T., Dimitropoulos, K., Koutsia, A. & Grammalidis, N. 2010. Video sensor network for real-time traffic monitoring and surveillance. *IET intelligent transport systems,* 4**,** 103-112.

Senior, A., Hampapur, A., Tian, Y.-L., Brown, L., Pankanti, S. & Bolle, R. 2006. Appearance models for occlusion handling. *Image and Vision Computing,* 24**,** 1233-1243.

Shah, M., Deng, J. D. & Woodford, B. J. 2015. A Self-adaptive CodeBook (SACB) model for real-time background subtraction. *Image and Vision Computing,* 38**,** 52-64.

Shen, X.-L. 2007. *Vision Detection of Lanes and Vehicles for Advanced Safety Vehicles.* Master, National Central University.

Sivaraman, S. & Trivedi, M. M. 2010. A general active-learning framework for on-road vehicle recognition and tracking. *IEEE Transactions on Intelligent Transportation Systems,* 11**,** 267-276.

Sivaraman, S. & Trivedi, M. M. 2013. Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis. *IEEE Transactions on Intelligent Transportation Systems,* 14**,** 1773-1795.

Sivaraman, S. & Trivedi, M. M. 2014. Active learning for on-road vehicle detection: A comparative study. *Machine vision and applications*, 1-13.

Skodras, E., Siogkas, G., Dermatas, E. & Fakotakis, N. 2012. Rear lights vehicle detection for collision avoidance. In *2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP)*. Vienna: IEEE, 134-137.

Sobel, I. 1990. An isotropic 3× 3 image gradient operator. *Machine vision for three-dimensional scenes*, 376-379.

Stauffer, C. & Grimson, W. E. L. 1999. Adaptive background mixture models for real-time tracking. In *1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Fort Collins: IEEE, 246-252.

Stevenson, R. 2011. A driver's sixth sense. *IEEE Spectrum,* 48, 50-55.

Sun, Z., Bebis, G. & Miller, R. 2006a. Monocular precrash vehicle detection: Features and classifiers. *IEEE transactions on image processing,* 15, 2019-2034.

Sun, Z., Bebis, G. & Miller, R. 2006b. On-road vehicle detection: A review. *IEEE transactions on pattern analysis and machine intelligence,* 28, 694-711.

Tan, R. T. 2008. Visibility in bad weather from a single image. In *2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Anchorage: IEEE, 1-8.

Tao, D., Li, X., Wu, X. & Maybank, S. J. 2007. General tensor discriminant analysis and gabor features for gait recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 29.

Teoh, S. S. & Bräunl, T. 2012. Symmetry-based monocular vehicle detection system. *Machine Vision and Applications,* 23**,** 831-842.

Thammakaroon, P. & Tangamchit, P. 2009. Predictive brake warning at night using taillight characteristic. In *2009 IEEE International Symposium on Industrial Electronics (ISIE)*. Seoul: IEEE, 217-221.

Tian, B., Tang, M. & Wang, F.-Y. 2015. Vehicle detection grammars with partial occlusion handling for traffic surveillance. *Transportation Research Part C: Emerging Technologies,* 56**,** 80-93.

Tsai, L.-W., Hsieh, J.-W. & Fan, K.-C. 2007. Vehicle detection using normalized color and edge map. *IEEE transactions on Image Processing,* 16**,** 850-864.

Varadarajan, S., Miller, P. & Zhou, H. 2015. Region-based mixture of gaussians modelling for foreground detection in dynamic scenes. *Pattern Recognition,* 48**,** 3488-3503.

Viola, P. & Jones, M. 2001. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*. Kauai: IEEE, 511-518.

Wang, H. & Cai, Y. 2015. Monocular based road vehicle detection with feature fusion and cascaded Adaboost algorithm. *Optik-International Journal for Light and Electron Optics,* 126**,** 3329-3334.

Wang, H., Yuan, C. & Cai, Y. 2015. Smart road vehicle sensing system based on monocular vision. *Optik-International Journal for Light and Electron Optics,* 126**,** 386-390.

Wang, X. 2013. Intelligent multi-camera video surveillance: A review. *Pattern recognition letters,* 34**,** 3-19.

Wang, X., Tang, J., Niu, J. & Zhao, X. 2016. Vision-based two-step brake detection method for vehicle collision avoidance. *Neurocomputing,* 173**,** 450-461.

Wang, Y., Jodoin, P.-M., Porikli, F., Konrad, J., Benezeth, Y. & Ishwar, P. 2014. CDnet 2014: an expanded change detection benchmark dataset. In *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Columbus, 387-394.

Wei, Y., Tian, Q., Guo, J., Huang, W. & Cao, J. 2018. Multi-vehicle detection algorithm through combining Harr and HOG features. *Mathematics and Computers in Simulation,* 155**,** 130-145.

Wen, X., Shao, L., Fang, W. & Xue, Y. 2015. Efficient feature selection and classification for vehicle detection. *IEEE Transactions on Circuits and Systems for Video Technology,* 25**,** 508-517.

Weng, M., Huang, G. & Da, X. 2010. A new interframe difference algorithm for moving target detection. In *3rd International Congress on Image and Signal Processing (CISP)*. Yantai: IEEE, 285-289.

Xia, Y., Shi, X., Song, G., Geng, Q. & Liu, Y. 2016. Towards improving quality of video-based vehicle counting method for traffic flow estimation. *Signal Processing,* 120**,** 672-681.

Xia, Y., Wang, C., Shi, X. & Zhang, L. 2015. Vehicles overtaking detection using RGB-D data. *Signal Processing,* 112**,** 98-109.

Xu, G., Xu, J., Wang, B., Tian, Y., Ye, Y. & Shah, S. G. 2014. Removal of rain in video based on motion and shape characteristics of raindrops. *Optik-International Journal for Light and Electron Optics,* 125**,** 3926-3930.

Yan, G., Yu, M., Yu, Y. & Fan, L. 2016. Real-time vehicle detection using histograms of oriented gradients and AdaBoost classification. *Optik-International Journal for Light and Electron Optics,* 127**,** 7941-7951.

Yang, Y., Gao, X. & Yang, G. 2011. Study the method of vehicle license locating based on color segmentation. *Procedia Engineering,* 15**,** 1324-1329.

Yang, Z. & Pun-Cheng, L. S. 2018. Vehicle Detection in Intelligent Transportation Systems and its Applications under Varying Environments: A Review. *Image and Vision Computing,* 69**,** 143-154.

Zhan, W. & Ji, X. 2011. Algorithm research on moving vehicles detection.

Procedia Engineering, 15, 5483-5487.

Zhang, W., Fang, X., Yang, X. & Wu, Q. J. 2007. Spatiotemporal Gaussian mixture model to detect moving objects in dynamic scenes. *Journal of Electronic Imaging,* 16, 023013-023013-6.

Zhang, W., Wu, Q. J. & bing Yin, H. 2010. Moving vehicles detection based on adaptive motion histogram. *Digital Signal Processing,* 20, 793-805.

Zhang, X., Zheng, N., He, Y. & Wang, F. 2011. Vehicle detection using an extended hidden random field model. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. Washington: IEEE, 1555-1559.

Zhao, N., Xia, Y., Xu, C., Shi, X. & Liu, Y. 2016. APPOS: An adaptive partial occlusion segmentation method for multiple vehicles tracking. *Journal of Visual Communication and Image Representation,* 37, 25-31.

Zhao, R. & Wang, X. 2013. Counting vehicles from semantic regions. *IEEE Transactions on Intelligent Transportation Systems,* 14, 1016-1022.

Zhou, Y., Liu, L., Shao, L. & Mellor, M. 2016. DAVE: A Unified Framework for Fast Vehicle Detection and Annotation. In *European Conference on Computer Vision*. Amsterdam: Springer, 278-293.

Zhu, Y., Comaniciu, D., Pellkofer, M. & Koehler, T. 2006. Reliable detection of overtaking vehicles using robust information fusion. *IEEE Transactions on*

Intelligent Transportation Systems, 7**,** 401-414.

Zielke, T., Brauckmann, M. & Vonseelen, W. 1993. Intensity and edge-based symmetry detection with an application to car-following. *CVGIP: Image Understanding,* 58**,** 177-190.

Zivkovic, Z. 2004. Improved adaptive Gaussian mixture model for background subtraction. *International Conference on Pattern Recognition,* 2**,** 28-31.

Zou, Q., Ling, H., Luo, S., Huang, Y. & Tian, M. 2015. Robust nighttime vehicle detection by tracking and grouping headlights. *IEEE Transactions on Intelligent Transportation Systems,* 16**,** 2838-2849.