International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 13 August 2023 **Revised:** 19 September 2023 **Accepted:** 26 October 2023*

Comprehensive Survey and Analysis of Techniques, Advancements, and Challenges in Video-Based Traffic Surveillance Systems

Ms. Joshi Rakhi Madhukaro, D. S. Rao
bhardwaj.rakhi78@ klh.edu.in ,dsrao@klh.edu.in
Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation, Hyderabad-500075,
Telangana, India

Abstract
The challenges inherent in video surveillance are compounded by a several factors, like dynamic lighting conditions, the coordination of object matching, diverse environmental scenarios, the tracking of heterogeneous objects, and coping with fluctuations in object poses, occlusions, and motion blur. This research endeavor aims to undertake a rigorous and in-depth analysis of deep learning- oriented models utilized for object identification and tracking. Emphasizing the development of effective model design methodologies, this study intends to furnish a exhaustive and in-depth analysis of object tracking and identification models within the specific domain of video surveillance.

Keywords: Deep Learning, Object detection, Object classification, Traffic Analysis, Video surveillance

I. INTRODUCTION

The increasing demands of security necessitate the advancement of intelligent technologies for video monitoring, as human-based surveillance systems are unable to cope adequately [15, 22, 26, 30, and 75]. Object detection and tracking are essential features of video surveillance systems, enabling effective security measures. Surveillance involves analyzing and observing the behavior of suspicious objects to ensure public safety. Automatic video monitoring is a cost-effective alternative to human-dependent security systems.

Computer vision involves segmenting moving objects to extract dynamic foreground content while eliminating stationary background [34, 36, 86]. Preprocessing is crucial for complex tasks like traffic monitoring, target identification, and object tracking. As visual data increases and processing resources become constrained in domains like self-driving cars and robotics, interpreting video data becomes increasingly important.

Traffic video analytics involves stages such as camera calibration, object detection, tracking, region selection, and incident detection [31, 38, 39, 53]. Object detection is the foundation for most processes, and researchers employ various methods to locate objects in traffic scenes. Motion-based methods leverage temporal information, while appearance-based systems extract significant elements from every input picture. The technique of choice is determined by the particular application, camera location, video quality, and processing capacity that is available [15, 33, and 74].

Video surveillance systems face challenges such as illumination variations, occlusion, and motion blur, tracking diverse objects, coordinate matching issues, environmental factors, and object tracking concerns. Proper camera placement is essential to ensure tracking capabilities are maximized. The paper outlines object detection, classification, tracking techniques, challenges in identifying and classifying vehicles, and available datasets in sections II, III, IV, V, and VI.

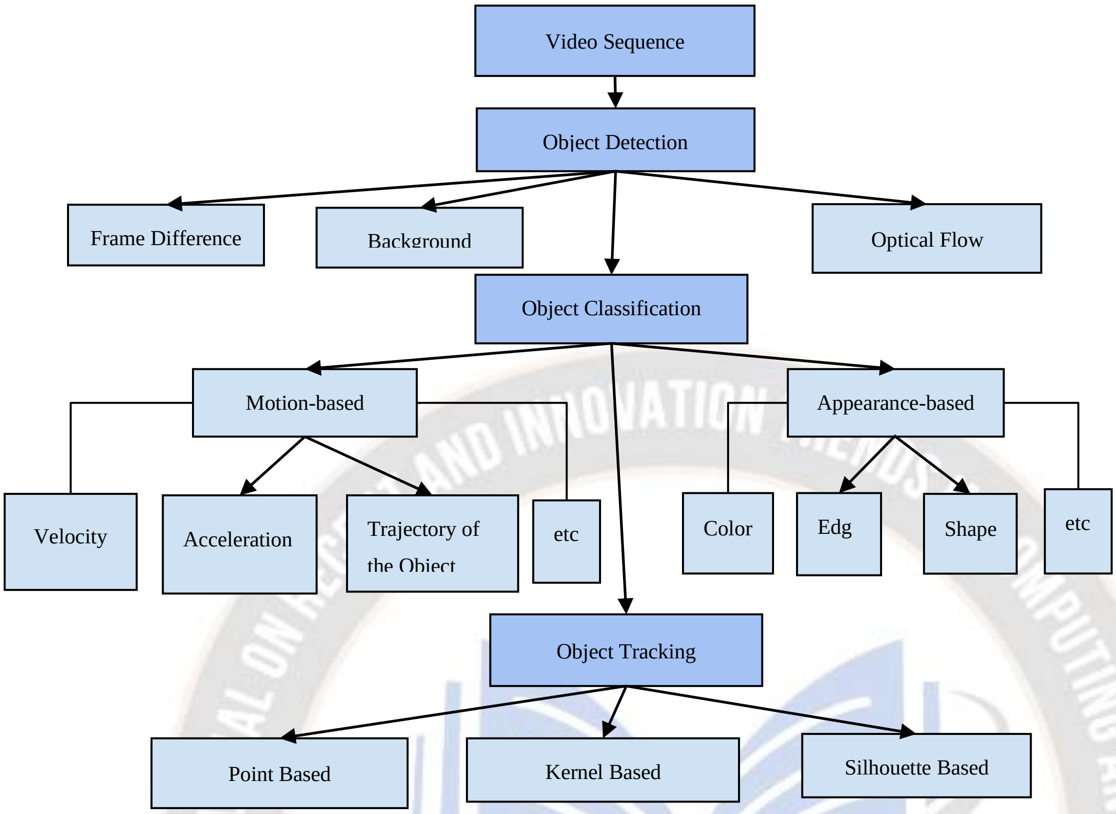

Figure 1: Generalized Traffic analysis techniques flow graph

II. SURVEY ON OBJECT DETECTION

Motion-based object detection is a technique used in video analysis to identify moving objects. It involves separating moving objects from still background objects, using techniques such as background removal, optical flow, and frame differencing [11, 32, and 65]. Background removal is a common technique used in image series from stationary cameras, where foreground pixels are identified based on a threshold. This method is straightforward and easy to implement, but can be inaccurate in recognizing non-stationary objects. To address this, a "moving body" approach is introduced, which classifies foreground pixels as foreground and background pixels as background.

Frame difference is another method used to identify moving objects by comparing the differences between two consecutive images. This method is simple and easy to apply, but can compromise the accuracy of identifying non-stationary objects due to the difficulty in obtaining a complete outline of a moving object. Usually, the identification procedure is comparing a preset threshold to the absolute difference between frame i-1 and frame i. If the absolute difference exceeds the threshold, it indicates the presence of a moving object [80].

The optical flow field technique is used to detect and move background objects, providing comprehensive information through grayscale conversion, proper threshold setting, and morphological operations. However, it is slow due to numerous calculations and is unreliable for real-time applications due to calculating complexity and inadequate anti-noise performance [134].

Background subtraction applications include object tracking, human-computer interaction applications, traffic monitoring, video surveillance, and content-based video encoding. Studies have employed motion segmentation to identify objects in traffic video recordings. Overall, motion-based and appearance-based object detection methods offer different approaches to identifying and monitoring moving objects in video surveillance systems [54, 58, and 96,107,120,136]. Object recognition techniques, such as convolution neural networks or customized characteristics (CNNs), are essential for detecting objects in various applications. These techniques include Template Matching

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 13 August 2023 **Revised:** 19 September 2023 **Accepted:** 26 October 2023*

and Feature Extraction, which are commonly used in traffic video analytics to identify objects of interest.

Template matching is the simplest detection method, using a portion of the image known as the template. A collection of templates can be used for greater accuracy. The system must be trained with a large dataset of a specific object, such as vehicles, to make it as resilient as possible [107,120].

Feature extraction is a process that identifies information patterns in feature vectors, which are then used in feature selection to simplify calculations and speed up detection. Techniques include Haar-like, gradient, projection histograms, unitary image transforms, deformable templates, SIFT, Gabor features, contour profiles, SURF, and Zernike moments. Template matching is effective when features are tracked and extracted from the template. Local Binary Pattern (LBP) features are used for extracting local features from textures. Object detection approaches involve hypothesis generation and verification using machine learning techniques like SVM, eSVM, and AdaBoost [1, 12, 19, 57, and 88,122].

Figure 1 illustrates the visualization of two distinct types of object detection: motion-based and appearance-based. Motion-based detection relies on parameters such as velocity, acceleration, and direction to identify objects. On the other hand, appearance-based techniques focus on static characteristics like size, color, shape, and other visual features of objects.

Table 1: Survey on Object detection using motion-based segmentation

Ref.	Year	Data	Objective
M Betkeet al. [9]	2000	Highway traffic	Road boundaries, markings and vehicles detection
C Zhanget al. [53]	2003	Road traffic	Traffic video analysis
G Wang et al. [87]	2008	Road and highway traffic	Vehicle detection review
Petrovskaya et al. [62]	2009	Urban traffic	Urban vehicle detection
KH Lim et al. [52]	2009	Highway traffic	Land and Vehicle detection in moving camera
AmiraliJazayerli et al. [39]	2011	Road and highway traffic	Real-time Vehicle detection dealing with changes in environment and illumination
Cao et al. [13-14]	2011	Highway traffic	vision-based vehicle detection
S Sivaraman et al. [76]	2012	Highway traffic	Real-time Vehicle detection
Hsieh et al. [33]	2013	Crossroads (USA)	Non–moving vehicles identification in a shifting environment
Bao et al. [7]	2014	Cityroad traffic	Embedded monitoring in live feed
M Rezaeiet al. [27]	2015	Highway traffic	Vehicle detection under challenging light conditions
Aqel et al. [5]	2015	CDnet dataset	Rapid dynamic vehicle identification
Z Charouh et al. [16]	2016	Highway traffic	Moving Vehicle detection
S Kamkar et al[42]	2016	Highway traffic	Vehicle detection and counting
M Ibarra-Arenado [36]	2017	Highway traffic	Vehicle detection in Urban traffic

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 13 August 2023 **Revised:** 19 September 2023 **Accepted:** 26 October 2023*

Hadi et al. [31]	2017	CDnet dataset	Detecting and tracking vehicle motion in shadowy conditions
S Li et al. [50]	2018	Road traffic	Single stage Vehicle detection
Cheng et al. [17]	2018	Expressway/Highway traffic	identification of congestion during adverse weather
Liu et al. [53]	2018	Expressway/Highway traffic	Real-time Vehicle detection
Garg et al. [24]	2019	Challenge dataset with background modelling	vehicle identification
Li et al. [49]	2019	AI City Challenge dataset [41]	ROI estimation and anomaly recognition
C.Premachandra et al. [66]	2019	Crossroads	identification of vehicles in a 360-degree perspective
Jaiswal et al. [38]	2020	National Highway traffic	Vehicle detection using two ameras
Jin et al. [40]	2021	State Highway traffic	Real time traffic flow monitoring
Ghahremannezhad et al. [74]	2021	National Highway traffic	Analyzing anomalies
Liu et al. [54]	2021	National Highway traffic	Edge-cloud traffic analyst solution
S. A. Alhuthali et al. [4]	2022	City Road traffic	Finding and enumerating of vehicles.
Ghahremannezhad et al. [25-26]	2022	ATON dataset	shadow removal method to detect vehicles
Zaidi et al. [100]	2022	Highway and Road traffic	Real-time Vehicle detection using Deep learning
J Wang et al. [42]	2023	Highway and city traffic	High precision method for vehicle classification
Duan et al. [19]	2023	Highway and Road traffic	Real-time Vehicle detection using Deep learning
Nugroho et al. [60]	2023	Expressway and Highway traffic	Deep learning based multi-vehicle detection.
W Shao et al. [88]	2023	Highway traffic	Failure detection in autonomous vehicle operation

Challenges that motion-based techniques must contend with

Motion segmentation methods address challenges like illumination, camera jitter, multimodal backgrounds, small object detection, and low frame rate in traffic surveillance recordings. However, they face performance drops due to stopped objects, weather, illumination, and occlusions. Studies use motion and appearance-based methods, such as faster R-CNN or YOLO, to detect stopped vehicles but face performance drops in object occlusions.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *13 August 2023* ***Revised:*** *19 September 2023* ***Accepted:*** *26 October 2023*

Table 2: Survey on Object detection using appearance-based segmentation

Ref.	Year	Feature	Data	Observations
Grimson et al.[57]	2005	SIFT	Traffic in metropolitan areas	Vehicle identification and characterization
A Khammari et al.[43]	2005	HOG	City traffic	Vehicle identification and characterization
J Zhou, D Gaoet al. [56]	2007	HOG	City traffic	Tracking and locating vehicles
P Negri, X Clady[67]	2008	Haar-like	City traffic	Tracking and locating vehicles
PE Rybski[69]	2010	HOG	Expressway and Highway traffic	Vehicle identification and characterization
X Cao et al[13]	2011	HOG	Expressway and Highway traffic	Vehicle identification and characterization
Z. Qia Guerrero-Gómez n et al.[30]	2013	SIFT	City traffic	Tracking and locating vehicles
S Tuermer et al.[13]	2013	HOG	Expressway and Highway traffic	Vehicle identification and tracking in urban areas
J.W Hsieh et al. [33]	2014	HOG, SURF	Traffic in metropolitan areas	Vehicle Model and Make Identification
A Mukhtar et al.[58]	2015	HOG	Urban traffic	Motorcycle detection and classification
K Kovačić, et al.[45]	2015	SURF, HOG and SIFT	City traffic	Tracking and locating vehicles
J. Lan et al. [47]	2016	HOG	City traffic	Identifying obstacles in real time
Y Xu, G Yu [94]	2016	(HOG + SVM)	Expressway and Highway traffic	Identification of vehicles from UAV images
D Neumann, T Langner, et al.[59]	2017	Haar-like, LBP, HOG	Expressway and Highway traffic	Recognition and categorization of vehicles
M. I. Ramadhani et al.[3]	2018	Haar-like	City traffic	Characterizing and identifying vehicles
Yu et al. [99]	2019	SIFT	Urban traffic	Preventing accidents
D Prasad et al.[65]	2019	HOG, LBP	Expressway and Highway traffic	Tracking and locating vehicles
J Caoet al. [12]	2020	Haar-like	Expressway and Highway	Automatic vehicle identification and

			traffic	classification
J Hu et al. [34]	2021	Haar-like	City and Highway traffic	Vehicle identification and characterization
Z Charouh et al.[16]	2022	Haar-lik	Expressway and Highway traffic	Moving vehicle detection and tracking
J Duanet al. [19]	2023	HAAR, LBP and HOG	Expressway and Highway traffic	Implicit vehicle categorization and identification
P Vandeth et al.[85]	2023	Haar-like	Expressway and Highway traffic	Motorcycle detection and tracking at nighttime
J Wang et al. [88]	2023	Haar-like	Expressway and Highway traffic	Tracking and locating vehicles

CNN-driven methods

Artificial neural networks (DCNNs) have grown in acceptance in 2012 due to advancements in object performance evaluation detection methods, computational speed, and hardware capacity. Deep convolution neural networks (DCNNs) have improved computer vision tasks like object detection using two- and one-stage techniques. Binary classifiers in the final layer enable deep learning object recognition in multi-class detection, widely used in traffic videos.

Two-stage object detection methods include Region-based Convolution Neural Network (R-CNN), Spatial Pyramid Pooling (SPP-Net), fast R-CNN, Mask R-CNN, and Single Shot Multi-book Detector (SSD). R-CNN research is significant in object identification, focusing on region-based classification. SPP-Net processes images of any size using Selective Search for 2000 object candidates are subsequently subjugated to propagate via classifiers. However, SPP-Net is too sluggish for real-time applications like traffic surveillance [22, 34, 35, and 80, 128]

Fast R-CNN improves image and object recommendations by initiating the process with the input image, eliminating the bottleneck and accelerating the inference process. Mask R-CNN segmentation technique is used for real-time object detection in traffic surveillance videos. Faster R-CNN is extended, extracting data on traffic flow and complete vehicle information. Traffic-related apps have also used image-based techniques YOLACT and YOLACT++ [22, 27, 35, and 60].

Single-stage object detection methods include You Only Look Once (YOLO), Non-Maximum Suppression (NMS), Single Shot Multi-book Detector (SSD), Center Net, Efficient-Det, and lightweight networks such as Squeeze-Net, Mobile-Net, and Shuffle-Net. SSD surpasses YOLO in performance and is faster than R-CNN but faces challenges in accurately recognizing small objects. Efficient-Det incorporates a pyramid network of weighted, bidirectional attributes (BiFPN) to fuse input features of varying sizes, while lightweight networks like Squeeze-Net, Mobile-Net, and Shuffle-Net can be utilized. The use of potent AI models is made possible by edge devices' capacity to identify objects [4, 8, 67, 73 and 93, 121].

Table 3: on Object detection using CNN driven methods

Reference	Year	CNN driven Method	2D/3D model	Regional Data	Observation
Wang S et al. [122]	2016	Fast-RCNN	2D	Urban area traffic	Multiple perspectives on vehicle classification.
Li et al. [65]	2016	FCN	3D	Road Traffic	3D Lidar vehicle detection using a fully convolutional

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *13 August 2023* ***Revised:*** *19 September 2023* ***Accepted:*** *26 October 2023*

					network.
Espinosaet al.[23]	2017	Faster R-CNN	3D	City traffic	Estimating vehicle speed.
Zhang et al. [74]	2017	R-CNN	2D	Road Traffic	Tracking and detecting vehicles in real time
Chang et al.[15]	2018	Faster R-CNN	3D	City traffic	Detecting anomalies in data.
Yu et al.[126]	2018	Mask R-CNN	3D	Vehicle Speed dataset	Traffic risk identification system.
Namet al.[78]	2018	CNN	2D	Road traffic	Vehicle classification captured by Thermal camera images
Asha et al.[4]	2018	YOLO	2D	City traffic	Vehicle statistics
Redmon,et al.[91]	2018	YOLO v3	2D	City-Road Traffic	Improved vehicle detection, tracking accuracy
Clausse et al.[16]	2019	R-CNN	2D	Urban area traffic	Autonomous vehicles challenges in urban environments
Wang et al. [114]	2019	YOLO	2D	City-Road Traffic	Detecting anomalies in data.
Bochkovskiyet al.[8]	2020	YOLOv4	2D	City traffic	YOLOV4 enhanced vehicle detection
Ghahremannezhad et al.[30]	2020	Mask R-CNN, Faster CNN, R-CNN	2D	Survey	Survey of CNN driven methods
Hu et al. [44]	2021	Mask R-CNN, Faster CNN, R-CNN	2D	Survey	Survey of CNN driven methods
Jana et al.[49]	2021	Computer Vision	2D	Road Traffic	Automated computer vision classification of vehicle movements at crossings
H. Perreault et al.[84]	2021	FFAVOD	2D	UA-DETRAC and UAVDT datasets used for classification.	Traffic surveillance uses video object detection.
Stadler al.[104]	2021	Cascade R-CNN	2D	AI City Challenge	Vehicle tracking using multiple cameras

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *13 August 2023* ***Revised:*** *19 September 2023* ***Accepted:*** *26 October 2023*

Luo et al. [69]	2021	YOLOv3	2D	UAV vehicle dataset	Crossroads vehicle counts
Yu et al. [125]	2021	R-CNN	2D	Highway traffic	Monocular 3D traffic surveillance system without training.
M. Rezaei et al. [93]	2021	YOLOv5	3D	Data analysis using UA-DETRAC, MIO-TCD, and GRAM.	Detection and tracking of vehicles
D. N.-N. Tran et al. [20]	2022	R-CNN scaled-YOLOv4, YOLOv5, and YOLOR Mask	2D	AI City Challenge	Vehicle tracking using multiple cameras
Z Charouh et al. [14]	2022	Faster R-CNN	2D	Highway traffic	Moving vehicle detection
Dinh, et al. [17]	2022	Faster R-CNN	2D	Faster R-CNN	Detecting and counting vehicles.
F. Li et al. [63]	2022	YOLOv5	2D	AI City Challenge	Vehicle tracking using multiple cameras
Liang et al. [67]	2022	YOLO	2D	Road Traffic	Edge based Vehicle detection
Liu, Zhang et al. [72]	2022	YOLOv7	2D	Urban area traffic	Vehicle tracking with multiple cameras utilizing inter-vehicle and occlusion-aware data
Duan et al. [19]	2023	Deep Cascade AdaBoost	2D	Highway traffic	Autonomous Vehicles detection
Farid et al. [25]	2023	Faster R-CNN	2D	Highway traffic	Acoustic vehicle detection for Unconstrained Environments
Herzog et al. [41]	2023	Faster R-CNN	2D	Highway traffic	Occlusion-aware and inter-vehicle facts for multi-camera vehicle monitoring
Ugrohoet al. [82]	2023	Mask R-CNN	2D	Road Traffic	Multi-camera parking slots empty detection.
Tanget al. [106]	2023	Faster R-CNN	3D	City and Highway Traffic	3D vehicle localization network
Vandethet al. [113]	2023	ANN, SVM	2D	City and Highway Traffic	Nighttime Motorcycle Detection

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 13 August 2023 **Revised:** 19 September 2023 **Accepted:** 26 October 2023*

Challenges that appearance-based techniques must contend with

I. Deep neural networks have revolutionized computer vision, but appearance-based object detection methods still face limitations such as performance, false positives, and time complexity. The speed of object detection algorithms has been improved via research, but not to an acceptable level.

II. Due to the numerous background samples and minority groups, object detection in traffic surveillance applications is challenging. Focused loss is used in deep learning models like Retina-Net to overcome this issue and lessen the effects of class imbalance. Occlusion and viewing perspective continue to be challenges, nevertheless.

III. The non-maximum suppression (NMS) in the Faster R-CNN approach has been substituted with a gentle NMS algorithm in several areas of research. Another study included a part-aware region proposal network. Other studies used 3D bounding boxes or multi-camera tracking.

III. CLASSIFICATION OF THE OBJECT

Object classification is a technique used to categorize objects into classes or groups, enabling detailed descriptions of events, increased detection accuracy, and high-level metadata [49, 94].
Static object classification uses intensity distance, similarities, and probabilities, while moving object classification uses a confidence map to count frames sequentially.

Object classification extracts distinct visual appearances from video scenes using features like size, shape, and compactness [55, 57, 60, and 76]. Supervised learning methods perform well in similar or similar target scenes, but performance decreases with small differences in view angle. Object classification follows object detection, analyzing shape, size, color, texture, and motion.

Shape-based classification uses boxes, blobs, and points to gather data about moving objects, while VASM classifies blobs into different classes and uses temporal consistency constraints.

Color-based histogram is accurate and requires less computational time due to its non-relationship with image size, orientation, and scale [96,103,107,112].

Texture-based classification reveals directionality, contrast, smoothness, brightness, roughness, and coarseness through the surface of an object, using methods like steerable pyramids, wavelet treatments, and Gabor wavelets [15, 33, 40, 56, and 121].
Motion-based classification categorizes objects based on acceleration, velocity, and position changes [89, 92].

IV. OBJECT TRACKING

Object tracking is a method used to determine an object's route and trajectory in the image plane by determining its position. It involves various techniques such as forward tracking, backtracking, and point tracking. Point tracking identifies points in next frames after object detection using position or feature extraction. Deterministic and statistical or probabilistic methods are the two categories under which point tracking is classified.

Deterministic methods use parameters to establish a relationship between previous and current frames, such as maximum velocity, smooth motion, common motion, proximity, proximity uniformity, and rigidity. These parameters can be used in deterministic and statistical methods and can be extended to multiple frames for better tracking results.

Statistical or probabilistic methods account for uncertainty while tracking, using the object's position, velocity, and acceleration to determine or forecast its next state. Single- or multiple-object tracking can be categorized as single- or multiple-object tracking [39, 41, 64, and 93,102,104].

Single Object Tracking can be done with any of these methods: Kalman Filter (KF), particle filtering, Multiple Object Tracking, Kernel Tracking, Template Tracking, Mean Shift Method, Support Vector Machines (SVM), Layering-Based-Matching, and Silhouette-Based Object Tracking [52,66].

Point tracking identifies points in next frames after object detection using position or feature extraction. It is challenging to employ in occlusion situations because once the object is totally obscured, it may lead to erroneous detection or miss-detection. The Global Nearest Neighbor (GNN) approach-based conventional approaches are efficient for single-target findings but may result in false identification when many objects are adjacent to each other. Joint Probability Data Association Filter (JPDAF) has a formation issue that makes it challenging to create precise

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 13 August 2023 **Revised:** 19 September 2023 **Accepted:** 26 October 2023*

correspondence amongst objects that are close together or in proximity [68].

Multiple Hypothesis Tracking (MHT) is a preferred method for multiple object tracking, forming alternate data association hypotheses in case of conflict. MHT can be classified into techniques that are both track- and hypothesis-oriented [102,104,108,123].

Kernel Tracking is the core of non-rigid object tracking, localizing the object in an image frame is owed to the isotropic kernel being spatially masked with the target to produce a similarity function. The Bhattacharya coefficient is frequently used to determine correlation scores, which provide a measure of how closely target models and candidate objects resemble each other and enable quick object location.

Template Tracking is a brute force method for locating object templates in larger images or video frames. It involves pixel by pixel matching, sliding pixel matrix at various image frame positions, and higher cross correlations are achieved at the pixel values where the template matches with the image frame.

The Mean Shift Method is a non-parametric robust statistical method introduced by Fukuunaga and Hostetler in 1975 for color-based object tracking. Improvements include the Continuously **Adaptive Mean Shift algorithm** (CAM Shift), which has a dynamic search window and improved performance by quantizing the 1-D histogram representing pixel probability values.

Support Vector Machines (SVM) classifier maximizes the separation of linear patterns at the decision boundary and utilizes regression and outlier detection using training systems with positive and negative values. Patterns not separable linearly are mapped to a new space using the kernel function, which selects the hyper plane with the highest margin and no classification error [12, 88,122].

Layering-based-Matching computes motion and condense scene representation, with segmentation and motion computation done iteratively. Dynamic layer tracking incorporates appearance data, global segmentation, and temporal consistency for challenging interactions like pass/stop.

V. RESEARCH GAP IN TRAFFIC MONITORING SYSTEMS TECHNIQUES AND FUTURE DIRECTIONS

The crucial necessity for building precise, reliable, and effective algorithms for finding objects of interest in traffic surveillance films offers up opportunities and research study possibilities. In this section, the primary difficulties, and potential areas for future research in intelligent traffic video analytics systems are briefly covered.

I. For object tracking, categorization, and event detection, traffic monitoring systems rely on motion segmentation algorithms. However, their broad use is hampered by antiquated infrastructures, computing constraints, and scant training data. Combining could potentially increase detection performance of appearance-based techniques with motion segmentation.

II. Traffic surveillance videos capture images from various locations, causing variations in lighting. Video quality can be influenced by factors such as resolution, viewing angle, distance, frame rate, and weather conditions. Detecting objects in traffic videos is challenging due to variations. Video analytics systems must adapt to weather, illumination, and perspectives. Integration of thermal cameras and LiDAR is limited due to costs.

III. Intelligent traffic video analytics systems must process video frames in real-time while considering aspects like resolution, frame rate, processing power, and cost-efficiency guidelines. Real-time needs must be configured for object detection methods. On edge computing platforms, improving application efficiency for traffic surveillance is still difficult.

IV. Real-time traffic monitoring using AI and ML algorithms for efficient detection, management, and response to traffic incidents, congestion, and abnormal patterns.

V. Predictive analytics for traffic flow involves using historical and real-time data to forecast traffic conditions, enabling proactive route planning, optimizing traffic signal timings, and improving overall traffic management strategies.

VI. Multi-modal traffic analysis involves analyzing vehicles, pedestrians, cyclists, and public transportation using advanced computer vision algorithms for accurate tracking and interaction analysis.

Research gaps can improve traffic surveillance and analysis, leading to more accurate, efficient, and intelligent systems, enhancing transportation efficiency, and reducing congestion and catastrophic situations.

VI. AVAILABLE DATASETS AND FRAMEWORKS

PASCAL VOC and COCO are common object detection datasets for traffic surveillance, with several openly accessible datasets and object detection-optimized video analytics systems available. The BGS Library, an online C++ framework, offers 43 background subtraction algorithms for motion segmentation in video analytics applications supporting traffic surveillance. The LRS Library is a MATLAB framework for detecting moving objects in videos.

The NVIDIA Deep Stream SDK is an expedited framework for creating real-time intelligent video analytics pipelines, scalable, cross-platform, and Transport Layer Security-encrypted. It supports modern object identification and instance segmentation models, as well as cutting-edge methods for tracking multiple objects [30, 37, 51, 54, and 56,101,112,119].

The Intel Deep Learning Streamer is a freely available framework designed for constructing streaming video analytics pipelines on edge or cloud servers. It uses pre-trained Open VINO models for Intel hardware platforms and supports object detection models for vehicle and pedestrian tracking. Azure Percept AI platform analyzes real-time video and audio using the COCO dataset, with other trained models like YOLO and Faster R-CNN being proficient at finding cars and people on foot.

Although not created expressly for surveillance applications, PASCAL VOC and COCO are two common object detection datasets for traffic surveillance. Table V provides a selection of openly accessible datasets and object detection-optimized video analytics systems [30, 37, 51, 54, and 56,101,112,119].
Table IV includes publicly available datasets from traffic surveillance videos that have been collected for tasks and research involving traffic surveillance, facilitating analysis in this area.

Table IV Survey on Available dataset for vehicle detection and Classification

Dataset Name	Year of The Dataset	Size of Dataset	Resolution	Observations
CDnet[119]	2014	160,000 frames	vary from 320×240 to 720×486	Motion and change detection
BIT-Vehicle [37]	2015	9,850 images	high resolution images (1600 × 1200 px and 1920 × 1080 px).	classification of vehicle types
Bharadwaj et al [128]	2016	66,591 images	1920 × 1080	Classification of vehicles and pedestrians
AI City Challenge [63]	2017 -2021	129,018 frames	Diverse	Enhanced vehicle tracking, re-identification, and traffic anomaly detection.
CADP [37]	2018	1,416 videos	Diverse	Analyzing accidents
TCP [56]	2018	234 videos	1280 × 720	Analyzing traffic behaviors
Vis Drone [135]	2018 - 2021	179,264 frames	Diverse	Tracking and identifying

				moving objects and people.
Fedorov et al. [26]	2019	982 frames	1920 × 1080	A projection of traffic flow.
AU-AIR [51]	2020	32,823 frames	1920 × 1080	Identifying and classifying travelers
DAWN [56]	2020	1,000 images	Diverse	road users in severe weather to identify.
MTID [51]	2020	6,200 frames	1920 × 1080	vehicle recognition using intersectional multiple views.
UA-DETRAC [123]	2020	140,000 frames	960 × 540	Object detection and tracking
Jin et al. [53]	2021	4,400 images	2288 × 1712	Vehicle identification with variable illumination
Neupane et al. [80]	2022	9,335 frames	1920 × 1080	Object detection
AI-TOD [30,128]	2023	28,036 IMAGES	Diverse	Vehicle detection and classification

VII. CONCLUSION

This paper discusses object detection, classification, and tracking techniques, focusing on their working principles and challenges. The importance of precise and effective tracking algorithms for video surveillance applications is emphasized. Survey examines computer vision techniques in monocular traffic surveillance.

Popular datasets are presented for training and validating these methods. While appearance-based deep learning algorithms are appropriate for monitoring urban crossings, motion segmentation-based methods are appropriate for monitoring highway and road traffic.

Traffic surveillance has been significantly affected by the development of deep learning, with more systems adapting to deep convolution network training for precise object detection and tracking. Further research is needed to improve performance, generalization, and efficiency in current algorithms.

REFERENCES

[1.] Agus, E. (2020). *by Eko Agus*. https://eprints.umm.ac.id/60746/19/Similarity - Ramadhani Minarno Cahyono - Haar-like Feature AdaBoost Cascade Classifier Vehicle Detection Digital Image Processing.pdf

[2.] Alhuthali, S. A. H., Zia, M. Y. I., & Rashid, M. (2022). A Simplified Traffic Flow Monitoring System Using Computer Vision Techniques. *Proceedings of 2022 2nd International Conference on Computing and Information Technology, ICCIT 2022*, *January*, 167–170. https://doi.org/10.1109/ICCIT52419.2022.9711550

[3.] Aqel, S., Sabri, A., & Aarab, A. (2015). Background modeling algorithm based on transitions intensities. *International Review on Computers and Software*, *10*(4), 387–392. https://doi.org/10.15866/irecos.v10i4.5432

[4.] Asha, C. S., & Narasimhadhan, A. V. (2018, October 4). Vehicle Counting for Traffic Management System using YOLO and Correlation Filter. *2018 IEEE International Conference on Electronics, Computing and Communication Technologies, CONECCT 2018*. https://doi.org/10.1109/CONECCT.2018.8482380

[5.] Bao, X., Javanbakhti, S., Zinger, S., Wijnhoven, R., & De With, P. H. N. (2014). Context-based object-of-interest detection for a generic traffic surveillance analysis system. *11th IEEE International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2014*, 136–141. https://doi.org/10.1109/AVSS.2014.6918657

[6.] Barnich, O., & Droogenbroeck, M. Van. (2010). Optimal insulin treatment in type 2 diabetes (The New England Journal of Medicine (2009) 361, (1801-1803)). *New England Journal of Medicine*, *363*(21), 2078.

[7.] Betke, M., Haritaoglu, E., & Davis, L. S. (2000). Real-time multiple vehicle detection and tracking from a moving vehicle. *Machine Vision and Applications*, *12*(2), 69–83. https://doi.org/10.1007/s001380050126

[8.] Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). *YOLOv4: Optimal Speed and Accuracy of Object Detection*. http://arxiv.org/abs/2004.10934

[9.] Bui, K.-H. N., Yi, H., & Cho, J. (n.d.). *A Vehicle Counts by Class Framework using Distinguished Regions Tracking at Multiple Intersections*.

[10.] Cao, J., Song, C., Song, S., Peng, S., Wang, D., Shao, Y., & Xiao, F. (2020). Front vehicle detection algorithm for smart car based on improved SSD model. *Sensors (Switzerland)*, *20*(16), 1–21. https://doi.org/10.3390/s20164646

[11.] Cao, X., Wu, C., Lan, J., Yan, P., & Li, X. (2011). Vehicle detection and motion analysis in low-altitude airborne video under urban environment. *IEEE Transactions on Circuits and Systems for Video Technology*, *21*(10), 1522–1533. https://doi.org/10.1109/TCSVT.2011.2162274

[12.] Cao, X., Wu, C., Yan, P., & Li, X. (2011). Linear Svm Classification Using Boosting Hog Features For Vehicle Detection In Low-Altitude Airborne Videos Center for OPTical Imagery Analysis and Learning (OPTIMIMAL), State Key Laboratory of Transient Optics and Photonics , Xi ' an Institute of Optic. *Ieee International Conference On Image Processing*, 2469–2472. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=46e785039d10088e391f3ba92f8e2a261e6c4aec

[13.] Chang, M.-C., Wei, Y., Song, N., & Lyu, S. (n.d.). *Video Analytics in Smart Transportation for the AIC'18 Challenge*. http://detrac-db.rit.albany.

[14.] Charouh, Z., Ezzouhri, A., Ghogho, M., & Guennoun, Z. (2022). A Resource-Efficient CNN-Based Method for Moving Vehicle Detection. *Sensors*, *22*(3), 1–13. https://doi.org/10.3390/s22031193

[15.] Cheng, J., Liu, B., & Tang, X. (2018). An automatic traffic-congestion detection method for bad weather based on traffic video. *International Journal of High Performance Computing and Networking*, *11*(3), 251–259. https://doi.org/10.1504/IJHPCN.2018.091909

[16.] Clausse, A., Benslimane, S., & De La Fortelle, A. (2019). Large-scale extraction of accurate vehicle trajectories for driving behavior learning. *IEEE Intelligent Vehicles Symposium, Proceedings*, *2019-June*, 2391–2396. https://doi.org/10.1109/IVS.2019.8814095

[17.] Dinh, D. L., Nguyen, H. N., Thai, H. T., & Le, K. H. (2021). Towards AI-Based Traffic Counting System with Edge Computing. *Journal of Advanced Transportation*, *2021*. https://doi.org/10.1155/2021/5551976

[18.] Dong, Z., Pei, M., He, Y., Liu, T., Dong, Y., & Jia, Y. (2014). Vehicle type classification using unsupervised convolutional neural network. *Proceedings - International Conference on Pattern Recognition*, 172–177. https://doi.org/10.1109/ICPR.2014.39

[19.] Duan, J., Ye, H., Zhao, H., & Li , Z. (2023). Deep Cascade AdaBoost with Unsupervised Clustering in Autonomous Vehicles. *Electronics (Switzerland)*, *12*(1), 1–17. https://doi.org/10.3390/electronics12010044

[20.] Duong, N.-N., Tran, L. H., Pham, H.-J., Jeon, H.-H., Nguyen, H.-M., Jeon, T., Tran, H.-P., & Jeon, J. W. (n.d.). *A Robust Traffic-Aware City-Scale Multi-Camera Vehicle Tracking Of Vehicles*.

[21.] Eleuch, S., Khouja, N., Milani, S., Erseghe, T., & Tlili, F. (2022). A Study on the Impact of Multiview Distributed Feature Coding on a Multicamera Vehicle Tracking System at Roundabouts. *IEEE Access*, *10*, 39502–39517. https://doi.org/10.1109/ACCESS.2022.3164681

[22.] Espinosa, J. E., Velastin, S. A., & Branch, J. W. (2021). Detection of Motorcycles in Urban Traffic Using Video Analysis: A Review. *IEEE Transactions on Intelligent Transportation Systems*, *22*(10), 6115–6130. https://doi.org/10.1109/TITS.2020.2997084

[23.] Espinosa, J. E., Velastin, S. A., & Branch, J. W. (2017). Vehicle detection using alex net and faster R-CNN deep learning models: A comparative study. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10645 LNCS*, 3–15. https://doi.org/10.1007/978-3-319-70010-6_1

[24.] Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J., & Zisserman, A. (2012). The PASCAL Visual Object Classes Challenge 2007 (VOC2007) Development Kit. *Pattern Analysis, Statistical Modelling and Computational Learning, Tech. Rep*, *2007*, 1–45. http://host.robots.ox.ac.uk/pascal/VOC/voc2007/devkit_doc_07-Jun-2007.pdf

[25.] Farid, A., Hussain, F., Khan, K., Shahzad, M., Khan, U., & Mahmood, Z. (2023). A Fast and Accurate Real-Time Vehicle Detection Method Using Deep Learning for Unconstrained Environments. *Applied Sciences (Switzerland)*, *13*(5). https://doi.org/10.3390/app13053059

[26.] Fedorov, A., Nikolskaia, K., Ivanov, S., Shepelev, V., & Minbaleev, A. (2019). Traffic flow estimation with data from a video surveillance camera. *Journal of Big Data*, *6*(1). https://doi.org/10.1186/s40537-019-0234-z

[27.] Fu, L., Yu, H., Juefei-Xu, F., Li, J., Guo, Q., & Wang, S. (2022). Let There Be Light: Improved Traffic Surveillance via Detail Preserving Night-to-Day Transfer. *IEEE Transactions on Circuits and Systems for Video Technology*, *32*(12), 8217–8226. https://doi.org/10.1109/TCSVT.2021.3081999

[28.] Garg, K., Ramakrishnan, N., Prakash, A., & Srikanthan, T. (2020). Rapid and Robust Background Modeling Technique for Low-Cost Road Traffic Surveillance Systems. *IEEE Transactions on Intelligent Transportation Systems*, *21*(5), 2204–2215. https://doi.org/10.1109/TITS.2019.2917560

[29.] Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. (2021). *YOLOX: Exceeding YOLO Series in 2021*. 1–7. http://arxiv.org/abs/2107.08430

[30.] Ghahremannezhad, H., Liu, C., & Shi, H. (2022). *Traffic Surveillance Video Analytics: A Concise Survey*. *July*. https://www.researchgate.net/publication/362366364

[31.] Ghahremannezhad, H., Shi, H., & Liu, C. (2020a). A New Adaptive Bidirectional Region-of-Interest Detection Method for Intelligent Traffic Video Analysis. *Proceedings - 2020 IEEE 3rd International Conference on Artificial Intelligence and Knowledge Engineering, AIKE 2020*, 17–24. https://doi.org/10.1109/AIKE48582.2020.00012

[32.] Ghahremannezhad, H., Shi, H., & Liu, C. (2020b). Automatic road detection in traffic videos. *Proceedings - 2020 IEEE International Symposium on Parallel and Distributed Processing with Applications, 2020 IEEE International Conference on Big Data and Cloud Computing, 2020 IEEE International Symposium on Social Computing and Networking and 2020 IEEE International Conference on Sustainable Computing and Communications, ISPA-BDCloud-SocialCom-SustainCom 2020*, *i*, 777–784. https://doi.org/10.1109/ISPA-BDCloud-SocialCom-SustainCom51426.2020.00123

[33.] Ghahremannezhad, H., Shi, H., & Liu, C. (2020c). Robust Road Region Extraction in Video under Various Illumination and Weather Conditions. *4th International Conference on Image Processing, Applications and Systems, IPAS 2020*, *ii*, 186–191. https://doi.org/10.1109/IPAS50080.2020.9334959

[34.] Ghahremannezhad, H., Shi, H., & Liu, C. (2022). Illumination-Aware Image Segmentation for Real-Time Moving Cast Shadow Suppression. *IST 2022 - IEEE International Conference on Imaging Systems and Techniques, Proceedings*, *June*. https://doi.org/10.1109/IST55454.2022.9827738

[35.] Girshick, R. (2015). Fast R-CNN. *Proceedings of the IEEE International Conference on Computer Vision*, *2015 Inter*, 1440–1448. https://doi.org/10.1109/ICCV.2015.169

[36.] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 580–587. https://doi.org/10.1109/CVPR.2014.81

[37.] Goyette, N., Jodoin, P.-M., Porikli, F., Konrad, J., & Ishwar, P. (2012). A new change detection benchmark dataset. *IEEE Workshop on Change Detection (CDW'12) at CVPR'12*, 1–8. https://openresearch-repository.anu.edu.au/bitstream/1885/29680/2/01_Goyette_Changedetection.net%3A_A_new_201

2.pdf

[38.] Guerrero-Gómez-Olmedo, R., López-Sastre, R. J., Maldonado-Bascón, S., & Fernández-Caballero, A. (2013). Vehicle tracking by simultaneous detection and viewpoint estimation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7931 LNCS*(PART 2), 306–316. https://doi.org/10.1007/978-3-642-38622-0_32

[39.] Hadi, R. A., George, L. E., & Mohammed, M. J. (2017). A Computationally Economic Novel Approach for Real-Time Moving Multi-vehicle Detection and Tracking toward Efficient Traffic Surveillance. *Arabian Journal for Science and Engineering*, *42*(2), 817–831. https://doi.org/10.1007/s13369-016-2351-8

[40.] Han, Y., & Hu, D. (2020). Multispectral fusion approach for traffic target detection in bad weather. *Algorithms*, *13*(11), 1–13. https://doi.org/10.3390/a13110271

[41.] Herzog, F., Chen, J., Teepe, T., Gilg, J., Hormann, S., & Rigoll, G. (2023). Synthehicle: Multi-Vehicle Multi-Camera Tracking in Virtual Cities. *Proceedings - 2023 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops, WACVW 2023*, 1–11. https://doi.org/10.1109/WACVW58289.2023.00005

[42.] Hofmann, M., Tiefenbacher, P., & Rigoll, G. (2012). Background segmentation with feedback: The pixel-based adaptive segmenter. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 38–43. https://doi.org/10.1109/CVPRW.2012.6238925

[43.] Hsieh, J. W., Chen, L. C., & Chen, D. Y. (2014). Symmetrical SURF and Its applications to vehicle detection and vehicle make and model recognition. *IEEE Transactions on Intelligent Transportation Systems*, *15*(1), 6–20. https://doi.org/10.1109/TITS.2013.2294646

[44.] Hu, J., Sun, Y., & Xiong, S. (2021). Research on the cascade vehicle detection method based on CNN. *Electronics (Switzerland)*, *10*(4), 1–20. https://doi.org/10.3390/electronics10040481

[45.] Huang, X., He, P., Rangarajan, A., & Ranka, S. (2022). Machine-Learning-Based Real-Time Multi-Camera Vehicle Tracking and Travel-Time Estimation. *Journal of Imaging*, *8*(4). https://doi.org/10.3390/jimaging8040101

[46.] Ibarra-Arenado, M., Tjahjadi, T., Pérez-Oria, J., Robla-Gómez, S., & Jiménez-Avello, A. (2017). Shadow-based vehicle detection in urban traffic. *Sensors (Switzerland)*, *17*(5), 1–19. https://doi.org/10.3390/s17050975

[47.] Ijjina, E. P., Chand, D., Gupta, S., & K, G. (2019). *Computer Vision-based Accident Detection in Traffic Surveillance*. https://doi.org/10.1109/ICCCNT45670.2019.8944469

[48.] Jaiswal, D., & Kumar, P. (2020). Real-time implementation of moving object detection in UAV videos using GPUs. *Journal of Real-Time Image Processing*, *17*(5), 1301–1317. https://doi.org/10.1007/s11554-019-00888-5

[49.] Jana, U., & Ness, D. (2021). *Automated Approach for Computer Vision-based Vehicle Movement Classification at Traffic Intersections*. 1–17.

[50.] 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*(2), 583–595. https://doi.org/10.1109/TITS.2011.2113340

[51.] Jensen, M. B., Møgelmose, A., & Moeslund, T. B. (2020). Presenting the Multi-view Traffic Intersection Dataset (MTID): A Detailed Traffic-Surveillance Dataset. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITSC 2020*. https://doi.org/10.1109/ITSC45102.2020.9294694

[52.] Jiao, J., & Wang, H. (2022). Traffic Behavior Recognition from Traffic Videos under Occlusion Condition: A Kalman Filter Approach. *Transportation Research Record*, *2676*(7), 55–65. https://doi.org/10.1177/03611981221076426

[53.] Jin, S., Gao, Y., Jing, S., Hui, F., Zhao, X., & Liu, J. (2021). Traffic Flow Parameters Collection under Variable Illumination Based on Data Fusion. *Journal of Advanced Transportation*, *2021*. https://doi.org/10.1155/2021/4592124

[54.] Kalsotra, R., & Arora, S. (2019). A Comprehensive Survey of Video Datasets for Background Subtraction. *IEEE Access*, *7*, 59143–59171. https://doi.org/10.1109/ACCESS.2019.2914961

[55.] Kamkar, S., & Safabakhsh, R. (2016). Vehicle detection, counting and classification in various conditions. *IET Intelligent Transport Systems*, *10*(6), 406–413. https://doi.org/10.1049/iet-its.2015.0157

[56.] Kenk, M. A., & Hassaballah, M. (2020). *DAWN: Vehicle Detection in Adverse Weather Nature Dataset*. https://doi.org/10.17632/766ygrbt8y.3

[57.] Khammari, A., Nashashibi, F., Abramson, Y., & Laurgeau, C. (2005). Vehicle detection combining gradient analysis and AdaBoost classification. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, *2005*, 66–71. https://doi.org/10.1109/ITSC.2005.1520202

[58.] Kim, K., Chalidabhongse, T. H., Harwood, D., & Davis, L. (2004). Background modeling and subtraction by codebook construction. *Proceedings - International Conference on Image Processing, ICIP*, *2*(November), 3061–3064. https://doi.org/10.1109/icip.2004.1421759

[59.] Kovačić, K., Ivanjko, E., & Jelušić, N. (2015). *Measurement of Road Traffic Parameters based on Multi-Vehicle Tracking*. 3–8. https://doi.org/10.20532/ccvw.2015.0002

[60.] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

[61.] Lan, J., Jiang, Y., Fan, G., Yu, D., & Zhang, Q. (2016). Real-Time Automatic Obstacle Detection method for Traffic Surveillance in Urban Traffic. *Journal of Signal Processing Systems*, *82*(3), 357–371. https://doi.org/10.1007/s11265-015-1006-4

[62.] Lee, G., Mallipeddi, R., Jang, G. J., & Lee, M. (2015). A Genetic Algorithm-Based Moving Object Detection for Real-time Traffic Surveillance. *IEEE Signal Processing Letters*, *22*(10), 1619–1622. https://doi.org/10.1109/LSP.2015.2417592

[63.] Li, F., Wang, Z., Nie, D., Zhang, S., Jiang, X., Zhao, X., & Hu, P. (n.d.). *Multi-Camera Vehicle Tracking System for AI City Challenge 2022*.

[64.] Li, S., & Chen, F. (2018). *3D-DETNet: a single stage video-based vehicle detector*. 57. https://doi.org/10.1117/12.2502012

[65.] Li, Y., He, K., Sun, J., & others. (2016). R-fcn: Object detection via region-based fully convolutional networks. *Advances in Neural Information Processing Systems*, *Nips*, 379–387. http://papers.nips.cc/paper/6465-r-fcn-object-detection-via-region-based-fully-convolutional-networks.pdf

[66.] Li, Z., Chen, Y., & Yin, Z. (2019). Vehicle tracking fusing the prior information of Kalman filter under occlusion conditions. *SN Applied Sciences*, *1*(8), 1–11. https://doi.org/10.1007/s42452-019-0852-2

[67.] Liang, S., Wu, H., Zhen, L., Hua, Q., Garg, S., Kaddoum, G., Hassan, M. M., & Yu, K. (2022). Edge YOLO: Real-Time Intelligent Object Detection System Based on Edge-Cloud Cooperation in Autonomous Vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *23*(12), 25345–25360. https://doi.org/10.1109/TITS.2022.3158253

[68.] Lim, K. H., Ang, L., Seng, K. P., & Chin, S. W. (2009). *Lane-Vehicle Detection and Tracking*. *II*, 5–10.

[69.] Liu, C., Zhang, Y., Hao Luo, †, Tang, J., Chen, W., Xu, X., Wang, F., Li, H., & Shen, Y.-D. (n.d.). *City-Scale Multi-Camera Vehicle Tracking Guided by Crossroad Zones*. https://www.aicitychallenge.org/

[70.] Liu, G., Shi, H., Kiani, A., Khreishah, A., Lee, J., Ansari, N., Liu, C., & Yousef, M. M. (2022). Smart Traffic Monitoring System Using Computer Vision and Edge Computing. *IEEE Transactions on Intelligent Transportation Systems*, *23*(8), 12027–12038. https://doi.org/10.1109/TITS.2021.3109481

[71.] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-319-46448-0_2

[72.] Liu, Y., Zhang, X., Zhang, B., Zhang, X., Wang, S., & Xu, J. (n.d.). *Multi-Camera Vehicle Tracking Based on Occlusion-aware and Inter-vehicle Information*. 3257–3264. https://openaccess.thecvf.com/content/CVPR2022W/AICity/papers/Liu_Multi-Camera_Vehicle_Tracking_Based_on_Occlusion-Aware_and_Inter-Vehicle_Information_CVPRW_2022_paper.pdf

[73.] Long, X., Deng, K., Wang, G., Zhang, Y., Dang, Q., Gao, Y., Shen, H., Ren, J., Han, S., Ding, E.,

& Wen, S. (2020). *PP-YOLO: An Effective and Efficient Implementation of Object Detector*. http://arxiv.org/abs/2007.12099

[74.] Lu, J., Xia, M., Gao, X., Yang, X., Tao, T., Meng, H., Zhang, W., Tan, X., Shi, Y., Li, G., & Ding, E. (n.d.). *Robust and Online Vehicle Counting at Crowded Intersections*.

[75.] M, M., & H. V., R. A. (2019). Design of Efficient Algorithms for Video Surveillance Applications using Artificial Intelligence. *Proceedings of ADCOM, October*. https://doi.org/10.34048/adcom.2019.phdforumpaper.5

[76.] Ma, X., & Grimson, W. E. L. (2005). Edge-based rich representation for vehicle classification. *Proceedings of the IEEE International Conference on Computer Vision*, *II*, 1185–1192. https://doi.org/10.1109/ICCV.2005.80

[77.] Mukhtar, A., & Tang, T. B. (2016). Vision based motorcycle detection using HOG features. *IEEE 2015 International Conference on Signal and Image Processing Applications, ICSIPA 2015 - Proceedings*, *March*, 452–456. https://doi.org/10.1109/ICSIPA.2015.7412234

[78.] Nam, Y., & Nam, Y. C. (2018). Vehicle classification based on images from visible light and thermal cameras. *Eurasip Journal on Image and Video Processing*, *2018*(1), 2–10. https://doi.org/10.1186/s13640-018-0245-2

[79.] Neumann, D., Langner, T., Ulbrich, F., Spitta, D., & Goehring, D. (2017). Online vehicle detection using Haar-like, LBP and HOG feature based image classifiers with stereo vision preselection. *IEEE Intelligent Vehicles Symposium, Proceedings*, *Iv*, 773–778. https://doi.org/10.1109/IVS.2017.7995810

[80.] Neupane, B., Horanont, T., & Aryal, J. (2022). Real-Time Vehicle Classification and Tracking Using a Transfer Learning-Improved Deep Learning Network. *Sensors*, *22*(10). https://doi.org/10.3390/s22103813

[81.] Nguyen-Ho, T. L., Pham, M. K., Nguyen, T. P., Nguyen, H. D., Do, M. N., Nguyen, T. V., & Tran, M. T. (2022). Text Query based Traffic Video Event Retrieval with Global-Local Fusion Embedding. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2022-June*, 3133–3140. https://doi.org/10.1109/CVPRW56347.2022.00353

[82.] Nugroho, H., Suwasono Adi, G., & Khoer Afandi, M. (n.d.). Detection of Empty/Occupied States of Parking Slots in Multicamera System Using Mask R-CNN Classifier. In *INTERNATIONAL JOURNAL OF APPLIED TECHNOLOGY RESEARCH* (Vol. 2023, Issue 1). https://ijatr.polban.ac.id/

[83.] Patrikar, D. R., & Parate, M. R. (2022). Anomaly detection using edge computing in video surveillance system: review. *International Journal of Multimedia Information Retrieval*, *11*(2), 85–110. https://doi.org/10.1007/s13735-022-00227-8

[84.] Perreault, H., Bilodeau, G.-A., Saunier, N., & Héritier, M. (2021). *FFAVOD: Feature Fusion Architecture for Video Object Detection*. http://arxiv.org/abs/2109.07298

[85.] Petrovskaya, A., & Thrun, S. (2009). Model based vehicle detection and tracking for autonomous urban driving. *Autonomous Robots*, *26*(2–3), 123–139. https://doi.org/10.1007/s10514-009-9115-1

[86.] Phan, H. N., Pham, L. H., Tran, D. N. N., & Ha, S. V. U. (2018). Occlusion vehicle segmentation algorithm in crowded scene for traffic surveillance system. *Advances in Intelligent Systems and Computing*, *672*(July), 584–595. https://doi.org/10.1007/978-981-10-7512-4_58

[87.] Pijpers, F. P. (2006). Performance metrics. *Astronomy and Geophysics*, *47*(6), 6.17-6.18. https://doi.org/10.1111/j.1468-4004.2006.47617.x

[88.] Prasad, D., Kapadni, K., Gadpal, A., Visave, M., & Sultanpure, K. (2019). HOG, LBP and SVM based Traffic Density Estimation at Intersection. *2019 IEEE Pune Section International Conference, PuneCon 2019*. https://doi.org/10.1109/PuneCon46936.2019.9105731

[89.] Premachandra, C., Ueda, S., & Suzuki, Y. (2019). Road intersection moving object detection by 360-degree view camera. *Proceedings of the 2019 IEEE 16th International Conference on Networking, Sensing and Control, ICNSC 2019*, *May*, 369–372. https://doi.org/10.1109/ICNSC.2019.8743319

[90.] Prevost, L., Negri, P., Clady, X., & Hanif, S. M. (2008). A cascade of boosted generative and discriminative classifiers for vehicle detection. *Eurasip Journal on Advances in Signal Processing*, *2008*.

https://doi.org/10.1155/2008/782432

[91.] Redmon, J., & Farhadi, A. (2018). *YOLOv3: An Incremental Improvement*. http://arxiv.org/abs/1804.02767

[92.] Revaud, J., & Humenberger, M. (n.d.). *Robust Automatic Monocular Vehicle Speed Estimation for Traffic Surveillance*. https://rebrand.ly/nle-cctv

[93.] Rezaei, M., Azarmi, M., & Mohammad Pour Mir, F. (2022). Traffic-Net: 3d Traffic Monitoring Using a Single Camera. *SSRN Electronic Journal*, 1–21. https://doi.org/10.2139/ssrn.4152596

[94.] Rybski, P. E., Huber, D., Morris, D. D., & Hoffman, R. (2010). Visual classification of coarse vehicle orientation using histogram of oriented gradients features. *IEEE Intelligent Vehicles Symposium, Proceedings*, 921–928. https://doi.org/10.1109/IVS.2010.5547996

[95.] Sentas, A., Kul, S., & Sayar, A. (2019). Real-Time Traffic Rules Infringing Determination over the Video Stream: Wrong Way and Clearway Violation Detection. *2019 International Conference on Artificial Intelligence and Data Processing Symposium, IDAP 2019*, 4–7. https://doi.org/10.1109/IDAP.2019.8875889

[96.] Shahbaz, A., Hariyono, J., & Jo, K. H. (2015). Evaluation of background subtraction algorithms for video surveillance. *2015 Frontiers of Computer Vision, FCV 2015*, *February 2019*, 1–4. https://doi.org/10.1109/FCV.2015.7103699

[97.] Shao, W., Xu, Y., Peng, L., Li, J., & Wang, H. (n.d.). *2301.04421*. https://arxiv.org/pdf/2301.04421

[98.] Shi, H., Ghahremannezhad, H., & Liu, C. (2021). Anomalous Driving Detection for Traffic Surveillance Video Analysis. *IST 2021 - IEEE International Conference on Imaging Systems and Techniques, Proceedings*, 0–5. https://doi.org/10.1109/IST50367.2021.9651372

[99.] Shladover, S. E., Vanderwerf, J., & Ragland, D. R. (2015). Robust Vehicle Detection and Distance Estimation Under Challenging Lighting Conditions. *IEEE Transactions on Intelligent Transportation Systems*, *16*(5), 8–15. https://eprints.whiterose.ac.uk/164835/1/T-ITS-Robust-Vehicle-Detection-2015.pdf

[100.] Sivaraman, S., & Trivedi, M. M. (2012). Real-time vehicle detection using parts at intersections. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 1519–1524. https://doi.org/10.1109/ITSC.2012.6338886

[101.] Snyder, C., & Do, M. N. (2019). STREETS: A novel camera network dataset for traffic flow. *Advances in Neural Information Processing Systems*, *32*(NeurIPS). https://proceedings.neurips.cc/paper/2019/file/ee389847678a3a9d1ce9e4ca69200d06-Paper.pdf

[102.] Specker, A., Florin, L., Cormier, M., & Beyerer, J. (2022). Improving Multi-Target Multi-Camera Tracking by Track Refinement and Completion. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2022-June*, 3198–3208. https://doi.org/10.1109/CVPRW56347.2022.00361

[103.] St-Charles, P. L., Bilodeau, G. A., & Bergevin, R. (2015). A self-adjusting approach to change detection based on background word consensus. *Proceedings - 2015 IEEE Winter Conference on Applications of Computer Vision, WACV 2015*, *June 2016*, 990–997. https://doi.org/10.1109/WACV.2015.137

[104.] Stadler, D., & Florin, L. (n.d.). *An Occlusion-aware Multi-target Multi-camera Tracking System*. https://openaccess.thecvf.com/content/CVPR2021W/AICity/papers/Specker_An_Occlusion-Aware_Multi-Target_Multi-Camera_Tracking_System_CVPRW_2021_paper.pdf

[105.] *stamp @ ieeexplore.ieee.org*. (n.d.). https://ieeexplore.ieee.org/iel7/6287639/9312710/09323062.pdf

[106.] Tang, X., Wang, W., Song, H., & Zhao, C. (2023). CenterLoc3D: monocular 3D vehicle localization network for roadside surveillance cameras. *Complex and Intelligent Systems*. https://doi.org/10.1007/s40747-022-00962-9

[107.] Tang, Zhen, Miao, Z., & Wan, Y. (2007). Background subtraction using running Gaussian average and frame difference. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *4740 LNCS*, 411–414. https://doi.org/10.1007/978-3-540-74873-1_50

[108.] Tang, Zheng, Naphade, M., Liu, M. Y., Yang, X., Birchfield, S., Wang, S., Kumar, R., Anastasiu, D., & Hwang, J. N. (2019). Cityflow: A city-scale benchmark for multi-target multi-camera vehicle tracking and re-identification. *Proceedings of the*

IEEE Computer Society Conference on Computer Vision and Pattern Recognition, *2019-June*, 8789–8798. https://doi.org/10.1109/CVPR.2019.00900

[109.] Thomas, S. S., Gupta, S., & Subramanian, V. K. (2018). Event Detection on Roads Using Perceptual Video Summarization. *IEEE Transactions on Intelligent Transportation Systems*, *19*(9), 2944–2954. https://doi.org/10.1109/TITS.2017.2769719

[110.] Tseng, D.-C., Hsu, C.-T., & Chen, W.-S. (2014). Blind-Spot Vehicle Detection Using Motion and Static Features. *International Journal of Machine Learning and Computing*, *4*(6), 516–521. https://doi.org/10.7763/ijmlc.2014.v6.465

[111.] Tuermer, S., Kurz, F., Reinartz, P., & Stilla, U. (2013). Airborne vehicle detection in dense urban areas using HoG features and disparity maps. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *6*(6), 2327–2337. https://doi.org/10.1109/JSTARS.2013.2242846

[112.] Vacavant, A., Chateau, T., & Wilhelm, A. (n.d.). *A Benchmark Dataset for Outdoor Foreground / Background Extraction – Supplementary Material*. *Section 2*, 2–7. http://chateaut.free.fr/publications/2012VacavantBMC.pdf

[113.] Vandeth, P., Tirtawangsa, J., & Nugroho, H. (2023). *Nighttime Motorcycle Detection for Sparse Traffic Images Using Machine Learning*. *17*(1), 81–92. https://journal.binus.ac.id/index.php/commit/article/download/8443/4727

[114.] Wang, Gaoang, Yuan, X., Zhang, A., Hsu, H. M., & Hwang, J. N. (2019). Anomaly candidate identification and starting time estimation of vehicles from traffic videos. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, *2019-June*, 382–390. https://openaccess.thecvf.com/content_CVPRW_2019/papers/AI City/Wang_Anomaly_Candidate_Identification_and_Starting_Time_Estimation_of_Vehicles_from_CVPRW_2019_paper.pdf

[115.] Wang, Guolin, Xiao, D., & Gu, J. (2008). Review on vehicle detection based on video for traffic surveillance. *Proceedings of the IEEE International Conference on Automation and Logistics, ICAL 2008*, *October*, 2961–2966. https://doi.org/10.1109/ICAL.2008.4636684

[116.] Wang, J., Dong, Y., Zhao, S., & Zhang, Z. (2023). A High-Precision Vehicle Detection and Tracking Method Based on the Attention Mechanism. *Sensors*, *23*(2). https://doi.org/10.3390/s23020724

[117.] Wang, K., Liu, Y., Gou, C., & Wang, F. Y. (2016). A Multi-view Learning Approach to Foreground Detection for Traffic Surveillance Applications. *IEEE Transactions on Vehicular Technology*, *65*(6), 4144–4158. https://doi.org/10.1109/TVT.2015.2509465

[118.] Wang, S., Liu, F., Gan, Z., & Cui, Z. (2016, November 21). Vehicle type classification via adaptive feature clustering for traffic surveillance video. *2016 8th International Conference on Wireless Communications and Signal Processing, WCSP 2016*. https://doi.org/10.1109/WCSP.2016.7752672

[119.] Wang, Y., Jodoin, P.-M., Porikli, F., Konrad, J., Benezeth, Y., & Ishwar, P. (n.d.). *CDnet 2014: An Expanded Change Detection Benchmark Dataset*.

[120.] Wen, J., Xu, Y., Tang, J., Zhan, Y., Lai, Z., & Guo, X. (2014). Joint video frame set division and low-rank decomposition for background subtraction. *IEEE Transactions on Circuits and Systems for Video Technology*, *24*(12), 2034–2048. https://doi.org/10.1109/TCSVT.2014.2333132

[121.] Wojke, N., & Häselich, M. (2012). Moving vehicle detection and tracking in unstructured environments. *Proceedings - IEEE International Conference on Robotics and Automation*, 3082–3087. https://doi.org/10.1109/ICRA.2012.6224636

[122.] Xu, Y., Yu, G., Wang, Y., Wu, X., & Ma, Y. (2016). A hybrid vehicle detection method based on viola-jones and HOG + SVM from UAV images. *Sensors (Switzerland)*, *16*(8). https://doi.org/10.3390/s16081325

[123.] Yao, H., Duan, Z., Xie, Z., Chen, J., Wu, X., Xu, D., & Gao, Y. (n.d.). *City-Scale Multi-Camera Vehicle Tracking based on Space-Time-Appearance Features*. https://github.com/ultralytics/yolov5

[124.] Ye, J., Yang, X., Kang, S., He, Y., Zhang, W., Huang, L., Jiang, M., Zhang, W., Shi, Y., Xia, M., & Tan, X. (n.d.). *A Robust MTMC Tracking*

System for AI-City Challenge 2021.

[125.] Yu, L., Chen, P., Liu, W., Kang, G., & Hauptmann, A. G. (2020). *Training-free Monocular 3D Event Detection System for Traffic Surveillance*. http://arxiv.org/abs/2002.00137

[126.] Yu, L., Zhang, D., Chen, X., & Hauptmann, A. (2018). *Traffic Danger Recognition With Surveillance Cameras Without Training Data*. http://arxiv.org/abs/1811.11969

[127.] Yu, Y., Xu, M., & Gu, J. (2019). Vision-based traffic accident detection using sparse spatio-temporal features and weighted extreme learning machine. *IET Intelligent Transport Systems*, *13*(9), 1417–1428. https://doi.org/10.1049/iet-its.2018.5409

[128.] Zaidi, S. S. A., Ansari, M. S., Aslam, A., Kanwal, N., Asghar, M., & Lee, B. (2022). A survey of modern deep learning based object detection models. *Digital Signal Processing: A Review Journal*, *126*, 1–18. https://doi.org/10.1016/j.dsp.2022.103514

[129.] Zapletal, D., & Herout, A. (2016). Vehicle Re-identification for Automatic Video Traffic Surveillance. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 1568–1574. https://doi.org/10.1109/CVPRW.2016.195

[130.] Zhang, C., Chen, S. C., Shyu, M. L., & Peeta, S. (2003). Adaptive background learning for vehicle detection and spatio temporal tracking. *ICICS-PCM 2003 - Proceedings of the 2003 Joint Conference of the 4th International Conference on Information, Communications and Signal Processing and 4th Pacific-Rim Conference on Multimedia*, *2*(December), 797–801. https://doi.org/10.1109/ICICS.2003.1292566

[131.] Zhang, Xingchen, Feng, Y., Angeloudis, P., & Demiris, Y. (2022). Monocular Visual Traffic Surveillance: A Review. *IEEE Transactions on Intelligent Transportation Systems*, 1–19. https://doi.org/10.1109/TITS.2022.3147770

[132.] Zhang, Xinxiang, Story, B., & Rajan, D. (2020). Night Time Vehicle Detection and Tracking by Fusing Sensor Cues from Autonomous Vehicles. *IEEE Vehicular Technology Conference*, *2020-May*(May). https://doi.org/10.1109/VTC2020-Spring48590.2020.9128989

[133.] Zhang, Xinxiang, Story, B., & Rajan, D. (2022). Night Time Vehicle Detection and Tracking by Fusing Vehicle Parts From Multiple Cameras. *IEEE Transactions on Intelligent Transportation Systems*, *23*(7), 8136–8156. https://doi.org/10.1109/TITS.2021.3076406

[134.] Zheng, Z., Jiang, M., Wang, Z., Wang, J., Bai, Z., Zhang, X., Yu, X., Tan, X., Yang, Y., Wen, S., & Ding, E. (n.d.). *Going Beyond Real Data: A Robust Visual Representation for Vehicle Re-identification*. https://github.com/NVlabs/DG-Net

[135.] Zhu, P., Wen, L., Du, D., Bian, X., Fan, H., Hu, Q., & Ling, H. (2022). Detection and Tracking Meet Drones Challenge. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(11), 7380–7399. https://doi.org/10.1109/TPAMI.2021.3119563

[136.] Zivkovic, Z. (2004). Improved adaptive Gaussian mixture model for background subtraction. *Proceedings - International Conference on Pattern Recognition*, *2*(May), 28–31. https://doi.org/10.1109/icpr.2004.1333992