Journal of Al-Qadisiyah for Computer Science and Mathematics Vol.17.(2) 2025,pp.Comp 162–173

University Of AL-Qadisiyah

Available online at www.qu.edu.iq/journalcm

JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS

ISSN:2521-3504(online) ISSN:2074-0204(print)

A Review of Deep Learning Approaches for Vehicle Classification and Tracking in Diverse Conditions

Zahraa Hussain Jumah[a], Adala M. Chyaid[b]

[a] *College of Computer Science and Information Technology , University of Basrah, Iraq. Email:zahraa.h.jumah@uobasrah.edu.iq*

[b] *College of Computer Science and Information Technology , University of Basrah, Iraq. Email:adala.gyad@uobasrah.edu.iq*

ARTICLE INFO

Article history:
Received: 05 /05/2025
Rrevised form: 01 /06/2025
Accepted : 03 /06/2025
Available online: 30 /06/2025

Keywords:

Vehicle Classification

Real-time Tracking

Deep Learning

CNN

LSTM

Computer Vision

Traffic Monitoring

Intelligent Transportation Systems (ITS)

ABSTRACT

Vehicle classification and tracking are fundamental components of intelligent transportation systems (ITS), enabling real-time traffic monitoring, surveillance, and autonomous driving. This paper provides a comprehensive review of modern deep learning and artificial intelligence methods used in vehicle classification and tracking. It highlights key methods such as CNNs, YOLO, Faster R-CNN, Kalman Filters, DeepSORT, and LSTM networks, with a particular focus on hybrid CNN-LSTM models that combine spatial and temporal features for robust performance in dynamic scenes. The review identifies major challenges reported in the literature, including occlusion, adverse weather conditions, real-time processing requirements, and data privacy concerns. It also outlines the most common application scenarios, such as smart surveillance, urban traffic control, and autonomous navigation. Based on the current trends, the paper recommends future directions involving vision transformers, reinforcement learning, edge computing, and multimodal sensor fusion. The goal is to offer researchers and practitioners a structured overview of the state-of-the-art, while highlighting opportunities for improving adaptability, scalability, and efficiency in smart transportation systems.

MSC..

https://doi.org/10.29304/jqcsm.2025.17.22196

1. Introduction

In recent years, the development of intelligent transportation systems (ITS) has significantly advanced the fields of vehicle detection, classification, and tracking. These components are vital for enabling real-time traffic monitoring, efficient road infrastructure utilization, enhanced safety, and long-term urban planning. As traffic authorities increasingly rely on accurate and automated data collection, the demand for intelligent, real-time solutions has grown accordingly [1]. The proliferation of surveillance cameras in urban and highway environments has transformed traffic monitoring practices. These systems are now widely used for real-time vehicle tracking and event detection. However, manual inspection of video footage remains inefficient—it often

∗Corresponding author: zahraa Hussain Jumah

Email addresses: *zahraa.h.jumah@uobasrah.edu.iq*

requires analysis of specific vehicle details such as color, type, and direction of movement, which is time-consuming and prone to human error due to fatigue and decreased attention span [2]. Recent breakthroughs in high-performance computing and deep learning have enabled the emergence of intelligent video analysis systems capable of automating vehicle classification and multi-object tracking tasks with high accuracy. These systems, particularly those based on convolutional neural networks (CNNs), YOLO, Faster R-CNN, Kalman Filters, DeepSORT, and LSTM networks, represent the current state-of-the-art (SOTA) in ITS applications [3]. Processing image sequences quickly and accurately is essential for many real-time computer vision tasks. During this process, intermediate representations are generated, enabling systems to capture conceptual information about scene dynamics and enhancing situational awareness [4].

This review categorizes the recent literature into three primary areas:

1.1 Vehicle Classification (VC):

Vehicle classification plays a fundamental role in ITS applications such as toll collection, traffic safety, autonomous driving, and parking management. Various machine learning methods have been applied, including Fuzzy Logic, Decision Trees, Adaboost, Random Forests, and Neural Networks. Among these, CNNs have received special attention for their ability to perform robust visual pattern recognition. Notably, Faster R-CNN and YOLO are widely recognized for delivering high-accuracy, real-time results in complex traffic environments [2][5][6].

1.2 Vehicle Tracking (VT):

Vehicle tracking involves re-identifying vehicle instances across consecutive video frames using information such as pixel values, object shapes, colors, and bounding boxes [7]. Some methods have relied on Kalman Filters for prediction, though they are limited by memory and computational constraints [8]. Other approaches have explored Kernelized Correlation Filters (KCF) [9] and LSTM-based architectures designed to support long-term, real-time tracking of multiple vehicles [10].

1.3 Integrated Vehicle Classification and Tracking (VC-VT):

The integration of vehicle classification and tracking provides a more complete understanding of traffic dynamics over time. For example, one study utilized OC-SVM for 3D feature-based classification and adaptive Kalman Filters for temporal motion estimation [11]. Other research adopted deep learning models—such as YOLO—for real-time classification and tracking in challenging environments [12]. Hybrid approaches combining SVMs, Adaboost with cascade classifiers, and Kalman Filters have also shown promising results in enhancing accuracy and minimizing tracking loss [13].

Research Objectives

This study aims to provide a comprehensive review of modern techniques used for vehicle classification and tracking through AI-driven approaches. The main objectives are:

1- Analyzing state-of-the-art techniques with a focus on integrating CNNs and LSTMs to enhance adaptability in diverse conditions.
2- Evaluating the performance of different models in terms of accuracy, computational efficiency, and robustness to environmental variations.
3- Exploring current challenges, such as the impact of environmental factors and the ability to improve performance with limited or imbalanced data.

By conducting this review, we seek to highlight recent advancements in the field and provide insights into future directions that could further improve the accuracy and efficiency of vehicle classification and tracking systems.

2. Fundamentals of Vehicle Classification and Tracking

The available methods for retrieving moving vehicle features can be grouped under the term vehicle classification, a term related to traffic engineering. VC is a basic unit for classifying vehicles into different groups and plays an important role in intelligent transportation systems (ITS). It has been widely used in many fields, including traffic flow surveillance, security management, etc. Fixed devices, attack zones, global coverage, or a combination of these components can be used to monitor vehicles as they move on roads. Through these devices, a variety of information can be collected, such as the number of vehicles, their size, speed, weight, distance between them, type, model, and stop plate number [14].

Object tracking is the task of locating and tracking a moving object (or multiple objects) in real time. The objective is to track these targets using their visual appearance between successive video frames by detection or association over time [15].

In recent times, sustainable intelligent transportation research has focused on real-time traffic management on highways with vehicle detection and tracking. Nevertheless, the performance of existing deep learning technology based approaches remain difficult since the vehicle is of different sizes, occlusions among other complex traffic scenes [16].

3. Traditional vs. AI-Based Techniques

Before artificial intelligence, deep learning and related techniques had developed, methods for vehicle classification were based mainly on traditional methods which used fixed sensors and video cameras including image analysis tools of low complexity. These methods were based on simple physical properties of vehicles, such as size, weight, and wheel numbers, to classify heterogeneous vehicles. While they worked well in various environments, these solutions were not adaptable to dynamic settings and conditions, e.g. changes in traffic or weather. They were human and therefore capable of errors and could be slow to process.

Owing to the recent developments in the field of computer vision and deep learning, algorithms based on deep learning have slowly taken place of traditional methods for real-time object detection. Specifically, Convolutional Neural Networks (CNN) have shown to achieve high accuracy in challenging environments and became an essential tool for vehicle recognition and tracking. These AI based systems are widely used in recent years to predict the traffic flow by using machine learning and deep learning algorithms which provides superior solution than traditional solutions [16].

Table 1: Comparison Between Vehicle Sensor Technologies and Traditional Classification Algorithms [17][18].

Category	Type / Model	Underlying Principle	Advantages	Limitations
Sensor-Based Methods	**Radar**	Measures physical width and length of vehicles using electromagnetic waves	More robust to environmental changes; structurally accurate	Less effective in high traffic density due to vehicle overlap
	Infrared Sensors	Analyze reflected light from vehicle surfaces and match patterns stored in a database	Effective in good visibility conditions; useful for identifying vehicle-specific patterns	Sensitive to weather changes (e.g., rain or fog), which may reduce accuracy
	Acoustic Sensors	Utilize vehicle-generated sound features that are independent of speed	Not affected by vehicle speed; suitable in limited-visibility environments	May be influenced by ambient noise or nearby vehicle sounds

Traditional Classification Algorithms	**Naive Bayes (NB)**	Probabilistic classification based on feature independence	Simple and effective for problems with independent variables	Assumes feature independence, which may not always hold true
	K-Nearest Neighbors (KNN)	Classifies vehicles based on distance from closest data points in feature space	No training phase needed; effective for small datasets	Slow with large datasets; sensitive to noise
	Random Forest (RF)	Ensemble of decision trees for improved generalization	Resistant to overfitting; handles complex data well	Slower prediction time; requires careful parameter tuning
	Support Vector Machine (SVM)	Maximizes the margin between classes using nonlinear decision boundaries	High accuracy, especially with high-dimensional data	Computationally intensive; performance degrades in noisy environments

4. Deep Learning Techniques in Vehicle Classification and Tracking

Recent breakthrough in artificial intelligence (in particular deep learning) has been speeding up its application in diversified fields! Vehicle classification, serving an important domain for intelligent transportation systems, traffic management, security alert, and autonomous driving, is one of the major applications of deep learning among them [19]. Deep learning (DL), the short form of Deep machine learning in AI, is used to extract more intricate representations from raw data for more powerful and precise analyses [20]. The depth in deep learning refers to the computational layers in the model which is more than that of the shallow learning methods [21]. DL can bring about major gains in domains where we have largescale, high-dimensional data, which is why deep neural networks generally work much better than shallow machine learning algorithms on tasks involving textual data, images, video, audio, and speech [22]. Also, detection and tracking of multiple objects in video streams is a basic task with applications from event detection, autonomous driving, and robot navigation [23].

4.1 Convolutional Neural Networks (CNNs): CNNs are the most commonly used deep learning algorithm and they have shown to be instrumental in working with images and assigning them into the classes that they were trained with. [24] The development of CNN architectures have improved vehicle classification systems by advancing models such as LetNet, AlexNet, VGG and GoogLeNet, which reflects progress in providing better classification accuracy and coping with real-world visual complexity [19].

4.2 Object Detection Algorithms: YOLO and Faster R-CNN are some of the popular networks. YOLO has very fast detection, so it is appropriate for real-time applications and it is also better in comparison with other deep learning methods [25]. while Faster R-CNN is its high accuracy and has big advantages in the object detection. It is superior to the classical ML in a large margin. These methods are widely exploited in traffic monitoring systems [26].

Table 2: comprehensive comparison between Faster R-CNN and YOLO in terms of speed, accuracy, and their suitability for real-time applications [78].

Criterion	Faster R-CNN	YOLO
Speed	Based on a two-stage approach that combines region proposal networks with region classification. While this improves accuracy, it	Formulates detection as a regression problem and performs detection in a single pass through the network, resulting in fast

	significantly reduces speed, making it less suitable for real-time applications.	processing suitable for real-time scenarios.
Accuracy	Achieves high detection accuracy due to its detailed region proposal and classification mechanism. Ideal for scenarios where precision is critical, such as offline video analysis.	Offers fast detection but with generally lower accuracy compared to Faster R-CNN, especially in complex scenes involving occlusions or poor lighting conditions.
Suitability for Real-Time Applications	Due to its multi-stage pipeline and slower inference time, it is less appropriate for real-time traffic surveillance systems.	Its high processing speed makes it well-suited for real-time traffic monitoring and control systems.

4.3 Tracking Algorithms: Kalman filters offer predictive tracking even in noisy data and are often used in object tracking systems [27]. They are among the most popular computational-based approach and has shown to perform well in the presence of noise in the input video scenes [28].

SORT integrates the spatio-temporal and appearance information to provide more robust tracking, which can track the video frames, fetch the target movement and track it further when the target reverse the direction of movement [15]. However, DeepSORT introduces an image feature extraction layer with deep method [29]. It is considered as one of the fastest tracking algorithm [30] and can cope with increasing number of vehicles as well as their visual similarity [31].

LSTM is applied to offer long-term temporal coherences, which helps accurate tracking in the cases of long-time sequences. It has been employed for the analysis of vehicle motion [32]. and it can track a moving source to high accuracy [33].

4.4 CNN-LSTM Hybrid Models: CNN-LSTM Hybrid Model has achieved some success in different fields like traffic prediction, in-vehicle security and driver safety. For example, a Hybrid CNN-LSTM Model (HCLM) was proposed for traffic volume prediction and it was more accurate compared with traditional traffic models such as ARIMA in terms of MAE and RMSE. The model was also highly efficient, building 70% faster than the nearest competitor [34]. In vehicle security, another hybrid model, HybridSecNet, employs LSTM and CNN for intrusion detection in vehicle CAN networks. The classification accuracy of the model was 99.5% in the detection of the attack, thereby enhancing the security of current vehicles from cyber attacks [35]. Regarding road safety: CNN-LSTM hybrid models to detect driver distractions with AI for driving safety by achieving a 93% accuracy in detecting distractions. This improves autonomous driving and is congenial to the Advanced Driver Assistance System (ADAS) [36].

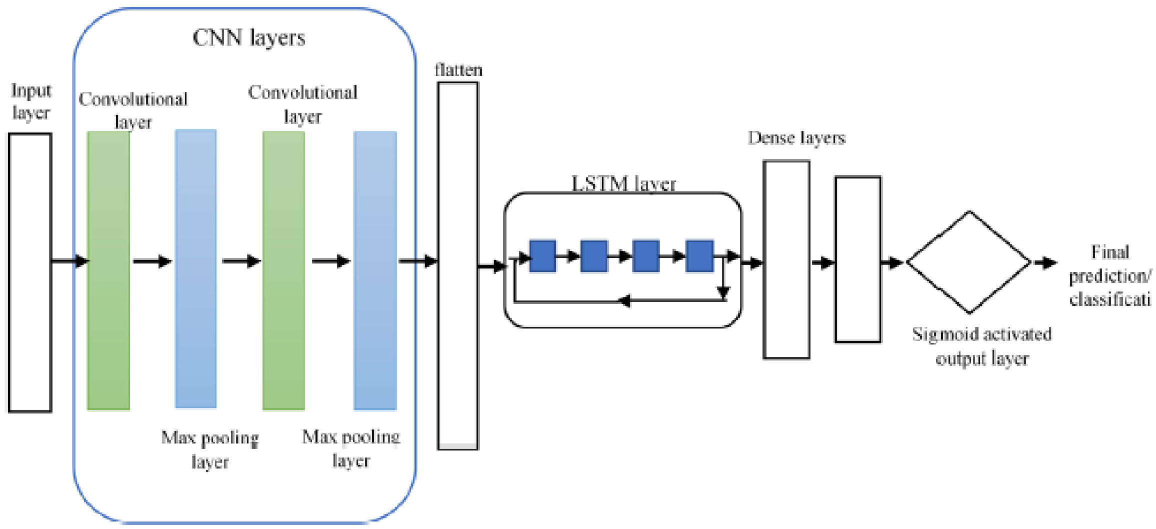

Fig. 1 Structure of Proposed Hybrid CNN-LSTM model [78].

5. Datasets for Model Training and Evaluation

The accuracy and efficiency of vehicle classification systems largely depend on the quality of the data used for their training and operation. Below is a comparison between the types of data used:

Table 3: Comparison of Dataset Types Used in Vehicle Classification and Tracking

Dataset Type	Advantages	Limitations	Typical Use Cases	References
Image/ Video	-Rich in visual content -Useful for distinguishing vehicle types -Enables dynamic behavior analysis over time -Can extract license plate, speed, type, and location -Lower cost and easier to install/maintain than loop sensors	-Image-based recognition more explored in faces than vehicles -May struggle in poor lighting or occlusion	-Vehicle detection and classification -Traffic pattern analysis -Surveillance-based tracking	[37], [38]
LiDAR	-Highly precise 3D spatial data -Provides range, intensity, and angle information -Excellent for object detection and classification in autonomous driving	-Accuracy affected by object distance from sensor -Higher cost and complexity	-Autonomous vehicle perception -Roadside classification systems	[18], [39]
Radar	-Effective in low visibility (fog, snow, etc.) -Robust under adverse weather -Bistatic/Forward-Scattering	-Requires objects to pass through specific baseline geometry Limited visual detail	-Vehicle classification in poor weather -Advanced target	[40], [41]

	-Radars enable accurate 180° tracking and target classification	tracking systems

6. Practical Applications

- **Urban Traffic Management:** Traffic flow in today's cities is one of the most significant issues that plagues modern cities, and manifests as road traffic traveling at slower or stopped speeds than would be anticipated without the presence of poor traffic flow management. Urban traffic authorities try to improve traffic management by using historical and present vehicle and urban infrastructure data. The evolution towards intelligent systems based on the artificial intelligence has contributed to increase the precision of the traffic data analysis and classification in relation to traditional mechanisms of detection [42]. These approaches use intelligent traffic prediction systems, real time data processing, dynamic multi-hop routing, and user friendly monitoring systems to minimize the traffic congestion and maximize the transportation efficiency [43]. As increasingly large high-quality datasets are becoming available, deep learning methods are increasingly playing a key role in detecting, predicting, and mitigating congestion [44].

- **Smart Surveillance Robots:** The next generation of security technology has been released with these smart security patrol robots that are revolutionizing the way businesses protect their property. These robots, which are embedded with sensors and real-time cameras, are aimed at identifying threats, tracking movement, and realizing persistent monitoring under less human intervention [45]. When these video frames collected from field environment are fed into deep learning techniques such as CNNs, they can detect and classify the event as normal or anomalous by comparing the captured frames pretrained datasets [46]. They have been put into public and private use applying them with integrated fire alarm, visitor automation and parking assistance [47]. Today's systems also connect to the web and APIs as well – e.g., the Flask API of the Telegram system is used for notifications and for user interaction, making it a connected and adaptive response scenario. Together, these robots provide a low cost, highly reactive security solution to address 24/7 monitoring needs in a variety of environments.

- **Autonomous Navigation: Autonomous navigation** Autonomous navigation is a central topic in robotics, which include localization, path planning and obstacle avoidance. Anew optimal control technique for light driverless vehicles application has been proposed, which allows an iterative process to optimize the path and reduce the travelling time, improve energy efficiency and comply with environment requirements. It has been implemented with ROS and middleware, and demonstrates an efficient response in restrictive indoor environments [48]. As for drones, autonomous navigation is crucial for dynamic tasks, such as, rescue and inspection. These systems are based on integrated perception, planning, and control algorithms, albeit utilizing different sensors (e.g. cameras or lidar) for generating paths in real-time and avoiding obstacles [49]. Furthermore, a new Deep Reinforcement Learning (DRL) based method has been proposed without any manual parameter tuning. The robot combines sensor data, local mapping, and goal setting to automatically derive navigation commands and in simulations it outperforms standard techniques [50].

- **Toll and Parking Automation:** Toll and parking automation are considered to play a crucial role in contemporary urban planning, trying to improve the efficiency and alleviate the congestion in the urban environment. Most of the traditional systems have manual payments and physical tickets which is more tiring and inconvenient. The current trends expand on these themes with developments such as ETC (electronic toll collection) devices, intelligent garage systems, and free-flow wireless systems to ease congestion. For example, ETC tags and automatic identification collectors have been used to process vehicle entry and exit, and the fees are directly deducted without stopping, which is conducive to it being used for parking fees, and the parking is high efficiency and user convenience [51]. New intelligent systems integrate vehicle data acquisition, wireless communication, automated gate control and alarm processing components and are all controlled by a central processor. Such systems assure accurate calculation of fee and save a lot of parking space offering seamless and automated solutions to the users [52, 53].

- **Accident Detection:** Accident Detection: Road accidents are considered as one of the greatest source of death and disability of the young in developed countries, so a timely and appropriate accident detection is very important. Contemporary study for this type of systems concentrates on the shortening of reporting period and increasing the detection precision [54]. In most developed countries about 40% of daily accidents result in a fatal injury, while in developing countries this is about 70% [55]. In nations such as India, the growing demise per year of vehicle drivers is due to the poor monitoring power and not limiting the drivers count [56].

7. Challenges in Vehicle Classification and Tracking

- **Environmental Factors:** Harsh weather, such as heavy snow, fog, rain, and sand or dust storms, bring a great challenge to the camera function, as it decreases the visibility and causes the unsafe driving. Indeed, these constraints influence the efficiency of traffic monitoring systems and autonomous driving applications in the context of detection and tracking algorithms [57]. The light waves emitted by cameras and LiDAR cannot penetrate through fog, heavy rain, or snow and cannot effectively detect the objects that are to be recognized [58]. Recent research has proposed the integration of radar sensors, which are less affected by atmospheric conditions, with vision-based systems to enhance detection robustness. Additionally, deep learning models trained on weather-augmented datasets have shown better generalization under adverse conditions. Future research may focus on multi-modal sensor fusion and domain adaptation techniques to improve detection reliability in extreme weather scenarios.

- **Occlusions and Overlaps:** Full or partial occlusion and vibration are situations of complete occlusion, which make vehicle tracking challenging and may even cause the tracking to fail [59]. These are all important challenges that lead to loss of tracking accuracy or to the inability of the system to perform its task optimally, in some cases making tracking impracticable [7]. To mitigate this, some tracking-by-detection methods leverage temporal consistency and motion prediction models such as Kalman filters or LSTM-based predictors to estimate the trajectory of occluded vehicles. Advanced re-identification (Re-ID) networks are also being used to recover lost tracks when the vehicle reappears. Further research is needed to improve object permanence reasoning and occlusion-aware tracking models.

- **Real-Time Requirements:** Low-latency inference with acceptable accuracy is important. The major difficulty lies on the realization of this vision system in real time conditions to get a precise localization and classification of vehicles in on-road traffic [59]. Recent studies have addressed this by employing lightweight CNN architectures such as MobileNet and YOLO variants, which allow for efficient inference on edge devices. Furthermore, pruning and quantization techniques help reduce model complexity. A promising direction involves dynamic inference models that adjust computational load based on scene complexity, thus maintaining real-time performance without sacrificing accuracy.

- **Data Privacy and Ethics:** Worries about privacy are sparked by surveillance footage. Efficient and intelligent data processing is a prerequisite to reach the performance desired and secure a scalable trustfulness of reliability, safety, and quality [60]. To address this, anonymization techniques such as blurring or obfuscation of personal data (e.g., license plates or faces) are increasingly incorporated. Federated learning also offers a privacy-preserving approach by training models locally on edge devices without transferring raw data to central servers. Future developments in ethical AI guidelines and legislation will further influence how data is collected, processed, and shared in intelligent transportation systems.

8. Emerging Trends and Advanced Technologies

- **Vision Transformers (ViT):** The vision transformer or simply Transformer is a type of AI model was originally proposed by Google Research in 2017 in the famous paper called *Attention is All You Need* [61]. Such a model uses self-attention mechanism, which allows it to compute the significance of each piece of the input without relying for anybody else. But its use in image classification jobs were initially restricted. Specialized networks even force us to resize and reshape input images unnecessarily! In order to overcome this, the authors introduced an architecture named as Vision Transformer (ViT). This works by dividing the input image into 2D patches and then flattening them to a one-dimensional sequence, which is then fed to the model [62]. Recently, transformer-based architectures, such as ViT, have achieved comparable performance, and in some cases outperformed, in image classification to the standard convolutional architectures such as ResNet [63]. Moreover, recent studies have explored the integration of ViTs with CNN-based backbones and LSTM-based temporal models to enhance spatio-temporal learning. Such hybrid architectures allow the system to benefit from ViT's global attention capabilities while leveraging the CNN-LSTM pipeline's strength in capturing local spatial features and sequential dependencies, making them promising candidates for vehicle classification and tracking tasks in complex environments.
- **Reinforcement Learning (RL):** is a computer vision subtopic of artificial intelligence, aims to detect and tracking objects in dynamic environments [64]. It is used for many modern applications including but not limited to surveillance, autonomous driving, robotics, human-computer interaction, vehicle navigation, drones, security systems, augmented reality, and intelligent transportation. One of the main strength of RL is the possibility to learn from the interaction with environment, since it can be successfully adapted to complex conditions. These would be cases of occlusion, visual artifacts, rapid movement, changes in lighting, and motion blur [64]. Furthermore, RL is employed to train effective tracking policies, which are learned for preserving temporal consistency of frames and optimizing long-term tracking performance to improve the robustness and accuracy of object tracking in dynamic environments [65]. In recent advancements, RL techniques have been successfully integrated with CNN-LSTM frameworks to optimize tracking policies based on continuous feedback from the environment. This fusion enables the model not only to learn spatial and temporal representations but also to dynamically adapt its behavior in response to environmental changes, thereby enhancing tracking accuracy and robustness in real-world transportation systems.
- **Predictive analytics:** It is one of the fundamental use cases of artificial intelligence in the domain of intelligent transportation systems, which is used to predict vehicle trajectories, predict future traffic trend for better traffic management and to avoid congestion. The technique involves processing large data sets from sensors, cameras and monitoring systems using machine learning algorithms and deep neural networks. In autonomous driving, the prediction part predicts future states of neighboring cars given the current and past observations of the surrounding. This predictive functionality gives the vehicle an earlier sense of potential risks, contributing to safer and more intelligent decisions [66].

- **Edge AI and IoT Integration:** Edge AI integration with Internet of Things (IoT) devices, such as sensors, marks a significant architectural trend for smart systems as it enables data processing on the device. This translates into lower latency, bandwidth savings, stronger privacy, and faster decision-making [67]. This disruptive paradigm is necessary to create effective, secure and responsive technology ecosystems in domains like healthcare and intelligent cities. Moreover, Edge AI assists in the creation of self-sufficient systems through data processing, resource allocation, and decision making on the edge by optimizing the same and provides better resource utilization, allowing autonomous systems to operate independently without the need for constantly communicating with centralized servers. This incorporation is important when constructing adaptive and robust intelligent systems [68].

- **Multimodal Sensor Fusion:** includes data fusion from different sensor types, which can exponentially increase the reliability for target detection and classification by utilizing the strong points of each and compensating for the drawbacks [69].

  Additionally, this approach facilitates the information verification and the systems security as well as its reliability by adding redundancy and complementarity between the different sensing means [70]. Due to the development of machine-learning and deep learning, sensor fusion is becoming intelligent and automatic, and it is widely used for a range of purposes in, for example, robotics, medical and environmental monitoring. As an example, a suggested approach combines information from a microphone and a visible-spectrum CCD camera in an effort to enhance defect detection during L-DED. The raw acoustic signals and the melt images are directly integrated in a hybrid CNN, and the end-to-end model achieves an 98.5% accuracy in defect prediction, the robot tool-center-point synchronization enables accurate localization [71].

9. Future Research Directions

- **End-to-End Learning Architectures: End-to-end learning architectures** to address detection, classification, and tracking processes have attracted substantial interests, particularly in multi-object tracking (MOT) and real-time applications. Unfortunately, these architectures suffer from difficulty of integration of two separate system of detection and tracking into one framework which is also inefficient and inaccurate. Recent approaches have adopted global response networks to directly encode image sequences and thus performance and efficiency are both improved on tracking [72]. An end-to-end real-time network for surveillance systems has been constructed bringing together detection, tracking and re-identification, getting high accuracy figures (i.e. 97% mAP and 98.3% MOTA) [73].

- **Explainable AI:** Explainable AI (XAI) is an important field in AI research, with the goal of increasing the transparency and interpretability of AI systems. With new, more advanced AI technologies, it is more important now than ever to understand why and how a machine is making decisions. This is important to gain trust and maintain accountability, in particular in high-stakes fields such as healthcare and finance. XAI uses a variety of methods to make the decision-making process of AI models more transparent and understandable to actual humans even when dealing with complex, opaque algorithms. Explainability in AI is how than can be useful in a few different ways. The Importance of trust and accountability First of all, trust and accountability are crucial, since XAI tackles the trustability of AI systems, specifically in domains that are safety critical and decisions have high impact [74]. Regulatory compliance also makes this an important issue; GDPR requires that users receive the "right to explanation" for automat ed decisions, and the adoption of AI systems in regulated industries also requires both explanation and demonstrating the absence of bias. Lastly, user engagement can be improved by making AI "explain" its decisions, which encourages cooperation between AI and humans and increases the quality of decisions [74].

- **Sustainable Computing:** Sustainable computing is a fundamental answer to the environmental concerns caused by the accelerating development of computing technologies. The issue of energy consumption and electronic waste has led to a requirement for green alternatives that meet technological progress with environmental responsibility. Recent studies identify barriers, such as technical, economic, and policy challenges to this transition [75]. The theme goes back to the more general concept of sustainability, which attempts to integrate economic progress with environmental protection by encouraging energy-efficient solutions and by linking innovations to the United Nations' Sustainable Development Goals (SDGs) [76]. Rising awareness of green computing within the industry has been driven by the public and economic benefits of the approach, as well as energy consumption regulations such as Energy Star program [77]. The sustainable computing spectrum now spans efficient algorithms, using renewable energy, recycling of materials, and smart cooling solutions—each necessitating cooperation between academia, business, and government [75].

Conclusion

Vehicle classification and tracking have become cornerstones of modern intelligent transportation systems, driven by the rapid advancements in deep learning and artificial intelligence. This review highlighted how these technologies—particularly convolutional neural networks, object detection frameworks like YOLO and Faster R-CNN, and sequence modeling tools such as LSTM and DeepSORT—have redefined the precision, efficiency, and scalability of vehicle monitoring systems. Through the exploration of traditional versus AI-based methods, it is evident that deep learning offers superior adaptability to complex real-world environments, including occlusion, adverse weather, and high traffic density. The integration of hybrid models like CNN-LSTM has shown promise in handling spatial-temporal dynamics, enhancing both classification accuracy and tracking persistence. Furthermore, the paper emphasized the importance of diverse datasets—ranging from video imagery to LiDAR and radar data—for training and benchmarking models in varied traffic conditions. Despite these advances, several practical and ethical challenges remain, including real-time processing constraints, environmental variability, and data privacy concerns. To address these, the field is moving toward more advanced solutions such as vision transformers, reinforcement learning, and multimodal sensor fusion. Moreover, the deployment of AI on edge devices and integration with IoT frameworks is enabling low-latency, real-time analytics crucial for smart city applications. Looking ahead, the development of end-to-end learning architectures, improved domain adaptation techniques, and explainable AI will be key to building transparent and generalizable models. Sustainability also emerges as a critical factor, with a growing need for energy-efficient computing in large-scale deployments.

References

1. Sharma, D. and Z.A. Jaffery, *Categorical vehicle classification and tracking using deep neural networks.* International Journal of Advanced Computer Science and Applications, 2021. **12**(9).
2. Tan, Z., *Vehicle classification with deep learning.* Master's thesis, Fırat University, Institute of Science, Elazığ, 63p, 2019.
3. Neupane, B., T. Horanont, and J. Aryal, *Real-time vehicle classification and tracking using a transfer learning-improved deep learning network.* Sensors, 2022. **22**(10): p. 3813.
4. Rad, R. and M. Jamzad, *Real time classification and tracking of multiple vehicles in highways.* Pattern Recognition Letters, 2005. **26**(10): p. 1597-1607.
5. Mostafa, M., et al. *Automatic Vehicle Classification and Speed Tracking*. in *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*. 2023. IEEE.
6. Faruque, M.O., H. Ghahremannezhad, and C. Liu, *Vehicle classification in video using deep learning.* Machine Learning and Data Mining in Pattern Recognition, MLDM, 2019: p. 117-131.
7. Azimjonov, J. and A. Özmen, *A real-time vehicle detection and a novel vehicle tracking systems for estimating and monitoring traffic flow on highways.* Advanced Engineering Informatics, 2021. **50**: p. 101393.
8. Liu, S., et al., *Vehicle tracking by detection in UAV aerial video.* Science China. Information Sciences, 2019. **62**(2): p. 24101.
9. Xu, Y. and J. Wang, *A unified neural network for object detection, multiple object tracking and vehicle re-identification.* arXiv preprint arXiv:1907.03465, 2019.
10. Farhodov, X., et al., *LSTM network with tracking association for multi-object tracking.* Journal of Korea Multimedia Society, 2020. **23**(10): p. 1236-1249.
11. Velazquez-Pupo, R., et al., *Vehicle detection with occlusion handling, tracking, and OC-SVM classification: A high performance vision-based system.* Sensors, 2018. **18**(2): p. 374.
12. Ahn, H. and Y.-H. Lee, *Vehicle classification and tracking based on deep learning.* Journal of web engineering, 2022. **21**(4): p. 1283-1294.
13. Saha, R., T. Debi, and M.S. Arefin. *Developing a framework for vehicle detection, tracking and classification in traffic video surveillance*. in *Intelligent Computing and Optimization: Proceedings of the 3rd International Conference on Intelligent Computing and Optimization 2020 (ICO 2020)*. 2021. Springer.
14. Shokravi, H., et al., *A review on vehicle classification and potential use of smart vehicle-assisted techniques.* Sensors, 2020. **20**(11): p. 3274.
15. Bathija, A. and G. Sharma, *Visual object detection and tracking using yolo and sort.* International Journal of Engineering Research Technology, 2019. **8**(11): p. 345-355.
16. Kumar, S., et al., *Fusion of deep sort and Yolov5 for effective vehicle detection and tracking scheme in real-time traffic management sustainable system.* Sustainability, 2023. **15**(24): p. 16869.
17. Gholamhosseinian, A. and J. Seitz, *Vehicle classification in intelligent transport systems: An overview, methods and software perspective.* IEEE Open Journal of Intelligent Transportation Systems, 2021. **2**: p. 173-194.
18. Wu, J., et al., *Automatic vehicle classification using roadside LiDAR data.* Transportation Research Record, 2019. **2673**(6): p. 153-164.
19. Li, T. and R. Stern, *Car-following-response-based vehicle classification via deep learning.* Journal on Autonomous Transportation Systems, 2024. **1**(1): p. 1-23.
20. Wang, Y. and H.A. Karimi, *Advanced deep learning models and algorithms for spatial-temporal data*, in *Big Data*. 2024, CRC Press. p. 227-253.
21. Endo, T., *Analysis of Conventional Feature Learning Algorithms and Advanced Deep Learning Models.* Journal of Robotics Spectrum, 2023. **1**: p. 001-012.
22. Janiesch, C., P. Zschech, and K. Heinrich, *Machine learning and deep learning.* Electronic markets, 2021. **31**(3): p. 685-695.

23. Pramanik, A., et al., *Granulated RCNN and multi-class deep sort for multi-object detection and tracking.* IEEE Transactions on Emerging Topics in Computational Intelligence, 2021. **6**(1): p. 171-181.
24. Jahan, N., S. Islam, and M.F.A. Foysal. *Real-time vehicle classification using CNN*. in *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*. 2020. IEEE.
25. Zuraimi, M.A.B. and F.H.K. Zaman. *Vehicle detection and tracking using YOLO and DeepSORT*. in *2021 IEEE 11th IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)*. 2021. IEEE.
26. Wang, X., et al., *Real-time vehicle type classification with deep convolutional neural networks.* Journal of Real-Time Image Processing, 2019. **16**: p. 5-14.
27. Khalkhali, M.B., A. Vahedian, and H.S. Yazdi, *Vehicle tracking with Kalman filter using online situation assessment.* Robotics and Autonomous Systems, 2020. **131**: p. 103596.
28. Farahi, F. and H.S. Yazdi, *Probabilistic Kalman filter for moving object tracking.* Signal Processing: Image Communication, 2020. **82**: p. 115751.
29. Sukkar, M., et al., *Enhancing pedestrian tracking in autonomous vehicles by using advanced deep learning techniques.* Information, 2024. **15**(2): p. 104.
30. Azhar, M.I.H., et al. *People tracking system using DeepSORT*. in *2020 10th IEEE international conference on control system, computing and engineering (ICCSCE)*. 2020. IEEE.
31. Lin, L., et al., *Realtime vehicle tracking method based on YOLOv5+ DeepSORT.* Computational Intelligence and Neuroscience, 2023. **2023**(1): p. 7974201.
32. Xiao, H., et al., *An improved LSTM model for behavior recognition of intelligent vehicles.* IEEE Access, 2020. **8**: p. 101514-101527.
33. Hussain, L.A., et al., *A predictive target tracking framework for IoT using CNN–LSTM.* Internet of Things, 2023. **22**: p. 100744.
34. Mead, M.A., *Hybrid CNN and LSTM Model (HCLM) for Short-Term Traffic Volume Prediction.* International Journal of Intelligent Computing and Information Sciences, 2022. **22**(4): p. 51-61.
35. Chougule, A., et al., *HybridSecNet: in-vehicle security on controller area networks through a hybrid two-step LSTM-CNN model.* IEEE Transactions on Vehicular Technology, 2024.
36. Kothuri, S.R., et al. *Hybrid CNN-LSTM machine learning algorithm for driver distraction detection*. in *2024 International Conference on Integrated Circuits and Communication Systems (ICICACS)*. 2024. IEEE.
37. Ni, X. and H. Huttunen, *Vehicle attribute recognition by appearance: computer vision methods for vehicle type, make and model classification.* Journal of Signal Processing Systems, 2021. **93**: p. 357-368.
38. Kim, H., *Multiple vehicle tracking and classification system with a convolutional neural network.* Journal of Ambient Intelligence and Humanized Computing, 2022. **13**(3): p. 1603-1614.
39. Liu, B., D. Zhao, and H. Zhang, *Road classification using 3D LiDAR sensor on vehicle.* Measurement Science and Technology, 2023. **34**(6): p. 065201.
40. Zhao, Z., et al., *Point cloud features-based kernel SVM for human-vehicle classification in millimeter wave radar.* IEEE Access, 2020. **8**: p. 26012-26021.
41. Kanona, M.E., et al., *A machine learning based vehicle classification in forward scattering radar.* IEEE Access, 2022. **10**: p. 64688-64700.
42. Nampally, R., *Leveraging AI in Urban Traffic Management: Addressing Congestion and Traffic Flow with Intelligent Systems.* Journal of Artificial Intelligence and Big Data, 2021. **1**(1): p. 86-99.
43. Kumar, A., et al. *Navigating Urban Mobility: A Review of AI-Driven Traffic Flow Management in Smart Cities*. in *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*. 2024. IEEE.
44. Kumar, N. and M. Raubal, *Applications of deep learning in congestion detection, prediction and alleviation: A survey.* Transportation Research Part C: Emerging Technologies, 2021. **133**: p. 103432.
45. Niharika, R., et al. *Voice Controlled Moving Robot for Smart Surveillance*. in *2024 5th International Conference on Circuits, Control, Communication and Computing (I4C)*. 2024. IEEE.
46. Avinashiappan, A., et al., *Smart surveillance using deep learning.* International Journal of Reconfigurable and Embedded Systems, 2023. **12**(3): p. 423.
47. Mahajan, Y., et al., *Smart CCTV Surveillance System with Visitor Management System, Face Detection, Fire Alarm, Parking Assistant for Residential Areas.* International Journal for Research in Applied Science and Engineering Technology 2023. **11**(11): p. 1340-1345.
48. Valera, Á., et al., *Navigation of autonomous light vehicles using an optimal trajectory planning algorithm.* Sustainability, 2021. **13**(3): p. 1233.
49. Lodi, M., S. Sarkar, and A. Dadhich. *Autonomous Drone Navigation in Dynamic and Uncertain Environments*. in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. 2024. IEEE.
50. Yang, R., Y. Tang, and G. Xiong. *Autonomous Navigation Method Based on Deep Reinforcement Learning with Dual-Layer Perception Mechanism*. in *2024 IEEE International Conference on Unmanned Systems (ICUS)*. 2024. IEEE.
51. Almeqbaali, N.N.A., et al. *Parking and Travel Ticket Automation–Recent Research Solutions*. in *2024 15th Annual Undergraduate Research Conference on Applied Computing (URC)*. 2024. IEEE.
52. Long, J., Qu, G., & Shi, S. , *New parking toll collector management system.* 2019.
53. Yang, D., Lin, G., Li, G., Song, Y., & Fang, Z., *Automatic toll collection system and ETC-based intelligent mechanical garage system.* 2020.
54. Kumar, N., D. Acharya, and D. Lohani, *An IoT-based vehicle accident detection and classification system using sensor fusion.* IEEE Internet of Things Journal, 2020. **8**(2): p. 869-880.
55. Pemila, M., et al., *Real time classification of vehicles using machine learning algorithm on the extensive dataset.* IEEE Access, 2024.
56. Sekhar, G.B., et al. *Vehicle Tracking and Speed Estimation Using Deep Sort*. in *2022 Second International Conference on Interdisciplinary Cyber Physical Systems (ICPS)*. 2022. IEEE.
57. Hassaballah, M., et al., *Vehicle detection and tracking in adverse weather using a deep learning framework.* IEEE transactions on intelligent transportation systems, 2020. **22**(7): p. 4230-4242.
58. Akita, T. and S. Mita. *Object tracking and classification using millimeter-wave radar based on LSTM*. in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. 2019. IEEE.
59. Lin, C.-J., S.-Y. Jeng, and H.-W. Lioa, *A Real-Time Vehicle Counting, Speed Estimation, and Classification System Based on Virtual Detection Zone and YOLO.* Mathematical Problems in Engineering, 2021. **2021**(1): p. 1577614.
60. Komorkiewicz, M., et al., *Intelligent Data Handling in Current and Next-Generation Automated Vehicle Development—A Review.* IEEE Access, 2023. **11**: p. 32061-32072.
61. Vaswani, A., et al., *Attention is all you need.* Advances in neural information processing systems, 2017. **30**.

62. Maurício, J., I. Domingues, and J. Bernardino, *Comparing vision transformers and convolutional neural networks for image classification: A literature review.* Applied Sciences, 2023. **13**(9): p. 5521.
63. Bhojanapalli, S., et al. *Understanding robustness of transformers for image classification*. in *Proceedings of the IEEE/CVF international conference on computer vision*. 2021.
64. Medina, M.C.C., B.J. Fernandes, and P.V. Barros, *The use of reinforcement learning algorithms in object tracking: A systematic literature review.* Neurocomputing, 2024: p. 127954.
65. Zhang, D., et al., *Deep reinforcement learning for visual object tracking in videos.* arXiv preprint arXiv:1701.08936, 2017.
66. Mozaffari, S., et al., *Deep learning-based vehicle behavior prediction for autonomous driving applications: A review.* IEEE Transactions on Intelligent Transportation Systems, 2020. **23**(1): p. 33-47.
67. Ali, K.S., *Edge AI and IoT: Direct integration for on-the-device data processing.* Advances in Engineering Innovation, 2023. **5**: p. 20-23.
68. Chenna, D., *EdgeAI in Self-Sustaining Systems With AI and IoT.* 2024.
69. Roheda, S., H. Krim, and B.S. Riggan, *Robust multi-modal sensor fusion: An adversarial approach.* IEEE Sensors journal, 2020. **21**(2): p. 1885-1896.
70. Wang, H., et al., *Recent progress on artificial intelligence-enhanced multimodal sensors integrated devices and systems.* Journal of Semiconductors, 2025. **46**(1): p. 011610.
71. Chen, L., et al. *Multimodal sensor fusion for real-time location-dependent defect detection in laser-directed energy deposition*. in *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. 2023. American Society of Mechanical Engineers.
72. Wan, X., et al., *Tracking beyond detection: learning a global response map for end-to-end multi-object tracking.* IEEE Transactions on Image Processing, 2021. **30**: p. 8222-8235.
73. Lei, M., et al., *End-to-end network for pedestrian detection, tracking and re-identification in real-time surveillance system.* Sensors, 2022. **22**(22): p. 8693.
74. Zodage, P., et al., *Explainable AI (XAI): History, basic ideas and methods.* International Journal of Advanced Research in Science, Communication and Technology, 2024. **3**.
75. Mathrani, R., *Challenges and Innovations in Achieving Sustainable Computing: A Comprehensive Analysis.* energy. **20305**: p. 6.
76. *Sustainable Computing—An Overview.* 2022.
77. *Sustainable computing.* IEEE Access, 2024. **46**(3).
78. *Pandey, A., Mannepalli, P. K., Gupta, M., Dangi, R., & Choudhary, G. (2024). A deep learning-based hybrid CNN-LSTM model for location-aware web service recommendation. Neural Processing Letters, 56(5), 234.*
79. Liu, X., & Rodelas, N. C. (2024). *Vehicle Detection and Tracking Techniques Based on Deep Learning in Road Traffic Surveillance*. [Journal/Conference Name if available], [Volume(Issue)], pages.