ISSN 2959-6157

A Comprehensive Examination of Machine Learning and Deep Learning Techniques for Driver Distraction Recognition

Ruijie Fan

Telecommunication Engineering,
Xi'an Jiaotong-Liverpool University,
Suzhou, China
Ruijie.Fan22@student.xjtlu.edu.cn

Abstract:

As the number of vehicles owned worldwide has rapidly increased, traffic accidents caused by distracted driving have become a serious public safety concern. Despite advancements in driver monitoring systems, it is still challenging to accurately identify distracted behavior. This paper thoroughly analyzes recent developments in driver distraction detection, with a focus on Deep Learning (DL) and conventional Machine Learning (ML) methods. It begins by analyzing the benefits of conventional ML methods, such as Support Vector Machines (SVM), Decision Trees (DT), and Logistic Regression (LR), with regard to interpretability, computational effectiveness, and practical use. The DL methods, such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Video Transformer Networks (VTN), are then thoroughly examined. Because of their strong feature extraction capabilities and capacity to represent intricate temporal and spatial dependencies present in driver behaviors; these DL techniques provide better performance. This review also discusses important issues that DL-based systems must deal with, like interpretability, applicability, and privacy. The paper also concludes by discussing potential avenues for future research, highlighting the significance of cause-aware intervention mechanisms, multimodal data fusion, and human-in-the-loop frameworks. These strategies aim to improve overall safety, user trust, and detection accuracy. In addition to summarizing contemporary approaches, this thorough review offers practitioners and researchers insightful information that will help them reduce the frequency of traffic accidents and enhance technologies for detecting driving behavior.

Keywords: Driver distraction detection; machine learning; deep learning.

ISSN 2959-6157

1. Introduction

In recent years, vehicle ownership has increased significantly. With countless drivers on the road every day, ensuring safe driving has become more important than ever. One of the trickiest problems in the world is road traffic accidents (RTAs). Millions of people have been harmed or even lost their lives as a result of RTA every year all over the world [1]. Three key elements have been found in studies to have a major impact on road traffic safety: the vehicle, the road environment, and the human component. The human element is one of the most crucial of these [2]. Over the past decade, driver safety has improved more than ever before. Even with these improvements, there are still a lot of terrible incidents that happen all around the world. Human errors, such as texting, chatting on the phone, eating, and drinking while driving, are mostly to blame for these collisions. Furthermore, exhaustion, drowsiness, and distractions all play a part in serious and sometimes fatal accidents [3]. Therefore, it is crucial to find ways to detect whether a driver is distracted. To better detect distracted driving, Numerous approaches have been explored in several studies within the last many years. Yu et al., for instance, created a smartphone-based monitoring system that employed neural networks and SVM to examine sensor data gathered while really operating a vehicle. Additionally, Tango and Botta used neural networks, the Adaptive Neuro Fuzzy Inference System (ANFIS), and SVM in a simulated setting with predetermined secondary tasks to build a nonintrusive distraction detection system based on vehicle signals [4]. While traditional methods rely on external sensors and devices to detect driver distraction, recent advancements have introduced more techniques based on ML. Nowadays, sensor-equipped driver monitoring systems can be used to collect driver behavior data, which can then be analyzed using intelligent algorithms. With the increasing prevalance of ML and Artificial Intelligence (AI), integrating them into Driver Monitoring Systems (DMSs) enables the motorist to be warned when they are distracted [5]. Although there has been an increasing number of relevant studies in recent years that contribute to the application of ML in detecting drivers' driving behaviors, Comprehensive evaluations of the most recent developments are still lacking.

This presentation will present a thorough analysis of current DL-based methods for classifying driver behavior, with a focus on distraction and drowsiness detection. This work provides a systematic summary of representative methods, analyzes their technical routes and design choices, and identifies key innovations and remaining challenges.

2. Method

2.1 Traditional ML-Based Detection Methods

Traditional methods have been widely used in early research for detecting driver distraction nowadays. Traditional ML models have two key advantages. One is high interpretability, which allows transparent and explainable decision-making. And the other is low computational cost, which enabling fast training and real-time inference. Several classical methods, such as SVM, DT, and LR, that have been used in driver distraction detection will be introduced in the section on traditional M.

2.1.1 SVM

SVM method, which was first developed by Vapnik [6], are powerful statistical learning tools used to model complex, nonlinear relationships in classification tasks. SVM has been used in fields like handwriting analysis, text classification jobs, and facial identification. They all reach successful achievement. SVM has also been adopted in driver distraction detection due to their robustness and computational efficiency [7, 8]. For example, to identify drivers by detecting their behavior, Qian et al. developed an SVM-based driver identification system by collecting driving behavior data from a simulated environment and extracting features using FFT [9]. Their results showed that FFT outperformed PCA and ICA, and the SVM classifier achieved over 85% accuracy in distinguishing individual drivers. These results confirm that SVM is suitable and accurate for identifying drivers' behavior based on behavioral data.

2.1.2 DT

One of the most popular and well-known techniques in ML for categorization and decision-making problems is DT. To analyze driver behavior, Abdullah and Sipos adopted a DT approach [10]. They collected via structured questionnaires from 1172 drivers in Duhok, Iraq, covering variables such as age, gender, speeding behavior, road type, and accident time to build a dataset. Responses were numerically encoded and stored in CSV format for processing. The decision tree was successful in identifying the high-risk behavior patterns in the dataset. Driver age, speeding habits, and support for speed control measures were all significant predictors. Notably, drivers who frequently exceeded the speed limit were greater likelihood of participating in serious collisions, particularly on multilane roads. Interpretable behavioral rules were also identified by the model, such as the association between older drivers and more serious but fewer accidents. This work demonstrates how interpretable ML models, like DTs, can be used to detect risky driving behavior based on structured behavioral data. The reported classification accuracy of 79.34% further supports its practical applicability in traffic safety research.

2.1.3 LR

LR is one of the most widely used classification algorithms in industry [11]. It is a supervised ML algorithm developed for learning classification problems [12]. Compared to other algorithms, LR has more accuracy, faster training speed, simplicity and efficiency. It is still one of the most favourite algorithms among many researchers in big data competitions for its efficiency and comprehensiveness [11]. Therefore, LR is also suitable for use in scenarios detecting whether the driver is distracted. Almadi et al. employed LR to assess driver behavior and forecast crash severity [13]. The method used binary logistic regression, and the dependent variable was crash severity (severe vs. non-severe). For model input, independent features such as driver age, speeding frequency, and mobile phone usage were normalized and numerically encoded. The model determines the probability of severe crashes as a function of the input variables, with coefficients showing the impact of each feature. For instance, getting older was linked to less serious collisions, but speeding was often linked to more serious ones. Despite having a slightly lower accuracy than DT (76% vs. 79%), LR produced results that were easy to understand and highlighted important behavioral factors that increase the risk.

2.2 DL-based Detection Methods

DL has become increasingly prominent in distraction detection tasks due to its powerful feature learning capabilities. DL enables the learning of data representations at different levels of abstraction using computational models with multiple processing layers. These techniques have significantly improved the state-of-the-art in a number of fields, including voice recognition, visual object identification, and object detection [14]. Compared to traditional ML, DL can automatically extract high-level features from raw input, making it especially suitable for complex scenarios and DL algorithm has been deployed in many scenarios. This section about DL will introduce, CNN, RNN and VTN. The three kinds of network are all part of Deep Neural Network (DNN).

2.2.1 CNN

CNN is a unique type of DNN. According to Kapoor et al. [15], a real-time driver distraction monitoring system based on CNN was proposed. In order to optimize pretrained CNN models like MobileNetV1, MobileNetV2, InceptionV3, and VGG-16, the system uses transfer learning on the State Farm Distracted Driver dataset. In order to identify distracting behaviors like texting or talking on the phone, the system analyzes real-time driver photos from an Android device installed on the dashboard of the car. According to their analysis, MobileNetV1 was the most effective distraction detection system in terms of top-1 accuracy, which made it appropriate for real-time driver safety applications.

2.2.2 RNN

RNN uses the concept of supervised learning, which specializes for sequential data and are capable of retaining information over time through their hidden state. RNNS are ideal for tasks involving temporal dependencies due to their capacity to process sequential data and store information in a hidden state over time. R. K. R and A. M. Mathew used a ReLU-BiLSTM model to identify driver distraction [16]. The BiLSTM, an extension of RNNs, captures both forward and backward temporal dependencies, whereas the ReLU activation does not have to deal with the vanishing gradient problem. By examining image sequences of driver behaviors (like eating or texting) and employing an attention mechanism to focus on crucial information, this model increases the accuracy of distraction detection. For real-time applications, the RNN approach outperformed other methods such as CNNs and SVMs, proving its effectiveness.

2.2.3 VTN

VTN is a video recognition framework based on transformers. Using a VTN model that employs a contrastive learning technique to categorize normal and anomalous driving conditions, A new methodology for identifying driver distraction was introduced by Koay et al [17]. To identify distractions, this model uses multimodal and Multiview video input from depth and infrared sensors. The Driver Anomaly Detection (DAD) dataset performs better than convolutional-based models, enabling the system to handle unobserved anomalous activities in an open-set recognition scenario. Swin-Tiny achieved an accuracy of 97.02% and an AUC of 0.9892. The findings demonstrate how transformer-based architectures may improve driver distraction detection.

3. Discussion

3.1 Challenges

Interpretability: In DL-based driver behavior detection systems, interpretability is still a major problem [18]. Despite the high accuracy of these models, their "black-box" design makes it challenging to comprehend the decision-making process, which is crucial in safety-critical applications such as driver monitoring [18]. Drivers and regulatory agencies may be less willing to rely on systems they do not fully understand due to a lack of transparency. Moreover, it is difficult to communicate model predictions in a way that is understandable to humans. For example,

ISSN 2959-6157

it is crucial to give explicit justifications for the alert, such as "driver's gaze deviated from the road for 3 seconds," if a system detects distraction [19]. Users might doubt the system's dependability in the absence of such transparency, which could result in abuse.

Applicability: There are several obstacles to overcome when putting DL-based driver behavior detection systems into practice. The cameras and sensors needed for these kinds of systems are often absent from cars, and updating them can be costly and technically challenging [20]. Moreover, models developed in controlled environments frequently do not generalize to a range of real-world situations, including different kinds of roads, traffic, and lighting [20]. Older or less expensive cars might not have the processing capability needed for real-time data processing, even with the right hardware [20]. Scalable hardware solutions, flexible algorithms, and effective models designed for embedded deployment are needed to address these applicability issues. The widespread integration of these systems into various vehicles and environments would be facilitated by these improvements.

Privacy: There are serious privacy issues with the use of DL in driver behavior detection systems. To assess a driver's level of engagement and alertness, these systems frequently gather physiological signals, eye movements, and facial expressions [21]. A significant amount of personal data may be gathered as a result of the ongoing monitoring necessary for precise detection, increasing the risk of unauthorized access and data leaks [22]. Researchers have suggested privacy-preserving techniques to allay these worries. Federated learning, for example, improves data privacy by enabling model training across decentralized devices without sharing raw data [22]. Differential privacy techniques can facilitate efficient model training, but they can also add noise to the data, making it challenging to identify specific drivers [23]. Even with these advancements, finding a balance between the need for accurate driver behavior detection and the protection of individual privacy remains challenging [21]. Methods must be created and refined via continuous research to guarantee the effectiveness of driver monitoring systems and the preservation of driver privacy [23].

3.2 Future Prospects

Human-in-the-loop: A promising future direction is the use of human-in-the-loop (HITL) frameworks, where driver feedback and contextual cues guide model adaptation. Unlike fully autonomous systems, HITL allows personalized, real-time adjustment, reducing false alarms and improving trust. This is particularly crucial for cars that are semi-autonomous, where human oversight remains essential for safety and decision-making. HITL enhances both interpretability and system reliability in real-world applications [24].

Multimodal Fusion: To increase accuracy and resilience, future driver monitoring systems are anticipated to integrate a variety of data sources, including physiological signals, vehicle telemetry, and facial images. Models can now account for elements like lighting variations, occlusions, or sensor noise that frequently impede unimodal approaches thanks to this multimodal fusion [25]. Furthermore, the effectiveness of unsupervised deep multimodal fusion frameworks for distraction detection has been demonstrated in recent studies. By combining visual inputs, vehicle telemetry, and behavioral signals, these systems can recognize behaviors such as eating, texting, and phone use without the need for large annotated datasets. The proposed deep network architecture combines multi-scale features from multiple modalities and employs reconstruction error for unsupervised learning, resulting in real-time performance and improved accuracy under various driving conditions [26].

Cause-Aware Intervention Mechanisms: Future driver state monitoring systems should be able to recognize driver states and provide appropriate interventions according to the underlying cause of the state. For example, to respond appropriately, the system must be able to differentiate between exhaustion and drowsiness brought on by illnesses. As a result, more targeted and effective interventions will increase driver safety. Additionally, understanding the underlying causes of specific states can help develop preventative measures to avert dangerous situations before they materialize [27].

4. Conclusion

In conclusion, the field of road safety has greatly advanced with the incorporation of DL and ML techniques into driver behavior detection systems. DL models like CNN, RNN, and VTN offer strong capabilities for complex and real-time behavior analysis, while traditional ML techniques like SVM, decision trees, and logistic regression offer intelligible and useful responses. Interpretability, applicability in a range of driving scenarios, and privacy concerns are just a few of the challenges that remain despite the advancements. Future developments should focus on protecting data privacy, increasing generalizability in real-world scenarios, and improving model transparency. In terms of improving road safety and reducing accident rates, driver behavior monitoring systems have a bright future due to the development of multimodal fusion, human-in-the-loop frameworks, and cause-aware.

References

[1] Le K G, Liu P, Lin L T. Using GIS-based spatio-temporal statistical analysis techniques to identify road traffic accident

hotspots in Hanoi, Vietnam. Geospatial Information Science, 2019, 23(2): 153–164
[2] Goniewicz K, Goniewicz M, Pawłowski W, et al. Road accident rates: Strategies and programmes for improving road traffic safety. European Journal of Trauma and Emergency Surgery, 2016, 42: 433–438.
[3] Kaplan S, Guvensan M A, Yavuz A G, Karalurt Y. Driver behavior analysis for safe driving: A survey. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(6): 3017–3032.
[4] Shahverdy M, Fathy M, Berangi R, Sabokrou M. Driver behavior detection and classification using deep convolutional neural networks. Expert Systems with Applications, 2020, 149: 113240
[5] Santos M, Coelho P J, Pires I M, Gonçalves P, Dias G P. An overview of machine learning algorithms to reduce driver fatigue and distraction-related traffic accidents. Procedia Computer Science, 2024, 238: 97–102.
[6] Shimizu Y, Yoshimoto J, Toki S, Takamura M, Yoshimura S, et al. Toward probabilistic diagnosis and understanding of depression based on functional MRI data analysis with logistic group LASSO. PLOS ONE, 2015, 10(5): e0123524.
[7] Liang Y, Reyes M L, Lee J D. Real-time detection of driver cognitive distraction using support vector machines. IEEE Transactions on Intelligent Transportation Systems, 2007, 8(2): 340–350.
[8] Byun H, Lee S W. Applications of support vector machines for pattern recognition: A survey. In: Lee S W, Verri A, eds. Pattern Recognition with Support Vector Machines (Lecture Notes in Computer Science, vol. 2388). Springer, 2002: 341–350.
[9] Qian H, Ou Y, Wu X, Meng X, Xu Y. Support vector machine for behavior-based driver identification system. Journal of Robotics, 2010, 2010: 397865.
[10] Abdullah P, Sipos T. Drivers' behavior and traffic accident analysis using decision tree method. Sustainability, 2022, 14(18): 11339.
[11] Jijo B T, Abdulazeez A M. Classification based on decision tree algorithm for machine learning. Journal of Applied Science and Technology Trends, 2021, 2(1): 20–28.
[12] Kang K, Gao F, Feng J. A new multi-layer classification method based on logistic regression. In: 2018 13th International Conference on Computer Science & Education (ICCSE). Colombo, Sri Lanka: IEEE, 2018: 1–4.
[13] Almadi I M, Al Mamlook R E, Ullah I, Alshboul O, Bandara N, Shehadeh A. Vehicle collisions analysis on highways based on multi-user driving simulator and multinomial logistic regression model on US highways in Michigan. International Journal of Crashworthiness, 2022, 28(6): 770–785.
[14] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature, 2015, 521: 436–444.
[15] Kapoor K, Pamula R, Murthy S V. Real-time driver distraction detection system using convolutional neural networks. In: Singh P, Panigrahi B, Suryadevara N, Sharma S, Singh A, eds. Proceedings of ICETIT 2019. Lecture Notes in Electrical Engineering, vol. 605. Cham: Springer, 2020: 263–270.
[16] R K R, Mathew A M. An extremely lightweight driver distraction detection system using recurrent neural network for enhanced road safety. In: Proceedings of the 2024 International Conference on Computational Intelligence and Network Systems (CINS). Dubai, UAE: IEEE, 2024: 1–7.
[17] Koay H V, Chuah J H, Chow C O. Contrastive learning with video transformer for driver distraction detection through multiview and multimodal video. In: Proceedings of the 2023 IEEE Region 10 Symposium (TENSYMP). Canberra, Australia: IEEE, 2023: 1–6.
[18] Li X, Xiong H, Li X, et al. Interpretable deep learning: interpretation, interpretability, trustworthiness, and beyond. Knowledge and Information Systems, 2022, 64: 3197–3234.
[19] Carvalho D V, Pereira E M, Cardoso J S. Machine learning interpretability: A survey on methods and metrics. Electronics, 2019, 8(8): 832.
[20] Elamrani Abou Elassad Z, Mousannif H, Al Moatassime H, Karkouch A. The application of machine learning techniques for driving behavior analysis: A conceptual framework and a systematic literature review. Engineering Applications of Artificial Intelligence, 2020, 87: 103312.
[21] Zhou W, Jia Z, Feng C, et al. Towards driver distraction detection: a privacy-preserving federated learning approach. Peer-to-Peer Networking and Applications, 2024, 17: 896–910
[22] Chhabra R, Singh S, Khullar V. Privacy enabled driver behavior analysis in heterogeneous IoV using federated learning. Engineering Applications of Artificial Intelligence, 2023, 120: 105881.
[23] Zhang L, Saito H, Yang L, Wu J. Privacy-preserving federated transfer learning for driver drowsiness detection. IEEE Access, 2022, 10: 80565–80574.
[24] Alkinani M H, Khan W Z, Arshad Q. Detecting human driver inattentive and aggressive driving behavior using deep learning: Recent advances, requirements and open challenges. IEEE Access, 2020, 8: 105008–105030.
[25] El-Nabi S A, El-Shafai W, El-Rabaie E S M, et al. Machine learning and deep learning techniques for driver fatigue and drowsiness detection: A review. Multimedia Tools and Applications, 2024, 83: 9441–9477.
[26] Zhang Y, Chen Y, Gao C. Deep unsupervised multi-modal fusion network for detecting driver distraction. Neurocomputing, 2021, 421: 26–38.
[27] Qu Y, Hu H, Liu J, Zhang Z, Li Y, Ge X. Driver state monitoring technology for conditionally automated vehicles: Review and future prospects. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1–20. Art. no. 3000920.