Advances in Computer, Signals and Systems (2024)
Clausius Scientific Press, Canada

DOI: 10.23977/acss.2024.080707
ISSN 2371-8838 Vol. 8 Num. 7

Research and design of illegal driving behavior detection model based on deep learning

Bingqing Chen, Yan Zhou, Hailing Zhang, Shuhan Zuo

Anhui University of Finance and Economics, Bengbu, 233030, China

Keywords: Illegal driving behavior, deep learning, traffic monitoring, intelligent transportation, road safety

Abstract: The rapid development of transportation systems and the growing number of vehicles on roads have significantly increased traffic-related risks, especially due to illegal driving behaviors such as speeding, distracted driving, and unauthorized lane changes. These behaviors not only disrupt traffic flow but also contribute to severe accidents, property damage, and fatalities. Traditional traffic monitoring techniques, such as radar-based systems and manual surveillance, are inadequate to address these complex challenges due to their dependency on predefined rules and limited scalability. This research introduces a robust illegal driving behavior detection model built on the principles of deep learning. By combining convolutional neural networks (CNNs) for spatial feature extraction and long short-term memory (LSTM) networks for temporal analysis, the proposed model captures complex driving patterns from traffic video data. A large-scale dataset featuring diverse driving scenarios and behaviors was used to train and validate the model, achieving a remarkable accuracy of 95%. The study not only demonstrates the potential of deep learning in traffic law enforcement but also highlights its advantages in scalability, automation, and real-time decision-making. This paper provides valuable insights for researchers and policymakers aiming to implement intelligent traffic management systems.

1. Introduction

Illegal driving behaviors represent a pressing issue for traffic safety and management worldwide. These behaviors, including speeding, distracted driving, illegal U-turns, and reckless lane changes, are major contributors to road accidents[1]. According to the World Health Organization (WHO), approximately 1.3 million fatalities and 50 million injuries occur annually due to road accidents, with many incidents linked to unsafe driving practices. Such behaviors often result from human errors, negligence, or violations of traffic laws. Efforts to monitor and mitigate these violations have traditionally relied on technologies such as radar for speed detection, manual video surveillance for lane violations, and rule-based algorithms for automated systems. However, these methods are often constrained by environmental conditions, such as poor visibility or heavy traffic, and fail to adapt to dynamic road scenarios. Moreover, manual observation is labor-intensive and prone to errors, further limiting its scalability.

Deep learning, a branch of artificial intelligence (AI), offers significant potential to address these limitations. Unlike traditional methods, deep learning models are capable of automatically extracting meaningful patterns from raw data, eliminating the need for manual feature engineering. These models excel in analyzing complex, unstructured data such as traffic video streams, making them ideal for detecting illegal driving behaviors in real time[2].This paper aims to present a novel deep learning-based illegal driving detection model that integrates CNN and LSTM architectures. The model is designed to analyze spatial and temporal features of traffic video data, ensuring accurate detection of unsafe driving patterns. The contributions of this research are twofold: Development of a hybrid CNN-LSTM model tailored for traffic video analysis.

2. Related Work

2.1. Traditional Methods for Traffic Monitoring

Conventional approaches for monitoring traffic violations often rely on tools like radar systems for speed measurement and closed-circuit television (CCTV) cameras for surveillance. These systems typically use predefined rules to identify violations, such as exceeding speed limits or crossing solid lane markers[3]. While effective in controlled environments, these methods are highly susceptible to errors in complex scenarios, such as overlapping vehicles, sudden lighting changes, or obstructed views caused by weather conditions. Rule-based algorithms, another commonly used technique, employ hard-coded instructions to detect illegal behaviors[4]. For instance, algorithms may calculate a vehicle’s speed by analyzing consecutive video frames. However, these systems lack adaptability to nuanced behaviors, such as distracted driving, and struggle in detecting violations that involve subjective judgments[5].

2.2. Machine Learning in Traffic Analysis

Machine learning (ML) methods have been employed to improve the efficiency of traffic monitoring systems. Algorithms such as support vector machines (SVMs), decision trees, and random forests have shown moderate success in classifying traffic behaviors. These methods rely on structured datasets with hand-crafted features, such as vehicle speed, position, and trajectory[6]. Although ML models outperform traditional methods in some cases, their dependency on manual feature extraction limits their ability to handle unstructured and dynamic data like video streams.

2.3. Deep Learning Approaches

Recent advancements in deep learning have revolutionized traffic behavior analysis. Convolutional neural networks (CNNs) are particularly effective in extracting spatial features from images and video frames, such as vehicle shapes, lane markings, and driver postures. Recurrent neural networks (RNNs), especially long short-term memory (LSTM) networks, are adept at processing sequential data, enabling the analysis of temporal patterns in driving behaviors[7]. Studies combining CNN and LSTM architectures have shown promising results in capturing both spatial and temporal dynamics. For instance, Zhou et al. (2023) utilized a CNN-LSTM model to detect reckless driving behaviors in urban environments, achieving an accuracy of 92%. These models demonstrate a significant improvement over traditional methods, particularly in scenarios involving complex traffic interactions.

3. Methodology

3.1. Data Collection and Preprocessing

Data collection involved the compilation of a diverse dataset comprising video footage of simulated and real-world driving scenarios. Footage was sourced from both open datasets and specially designed experiments that included a variety of illegal driving behaviors, such as distracted driving, illegal lane changes, and speeding. To ensure the model's adaptability, data were recorded under varying lighting, weather, and traffic conditions. Preprocessing included frame extraction from videos at 30 frames per second, resizing frames to 224x224 pixels, and normalization to enhance model convergence. Moreover, data augmentation techniques such as random cropping, flipping, and brightness adjustments were applied to improve generalizability[8]. Annotations were created using manual labeling tools, ensuring high accuracy in identifying illegal behaviors across the dataset.

3.2. Model Architecture

The CNN-LSTM hybrid architecture was designed to leverage the strengths of both convolutional and sequential networks. The CNN component utilized a ResNet-50 backbone, pre-trained on the ImageNet dataset, to extract spatial features from video frames. Features such as driver posture, vehicle movement, and environmental context were encoded as high-dimensional representations[9]. The LSTM component, consisting of two layers with 128 hidden units each, processed the sequential dependencies in the extracted features, enabling the detection of temporal patterns in driving behavior. This combination of spatial and temporal modeling was critical for accurately identifying behaviors such as brief distractions or gradual lane drifts, which require both spatial context and time-series analysis.

3.3. Training Process

The training process was conducted using a supervised learning approach, with illegal driving behaviors as categorical labels. The dataset was divided into training (70%), validation (15%), and testing (15%) subsets to evaluate model performance. The Adam optimizer was employed with an initial learning rate of 0.001, gradually reduced using a step decay scheduler to prevent overfitting. Cross-entropy loss was used as the objective function due to its suitability for multi-class classification tasks. The model was trained for 50 epochs with a batch size of 32, leveraging early stopping based on validation accuracy to prevent overtraining. GPU acceleration was utilized to expedite the process, enabling efficient handling of the large dataset.

3.4. Evaluation Metrics

Model performance was assessed using metrics including accuracy, precision, recall, F1-score, and confusion matrices. Accuracy provided a general performance overview, while precision and recall offered insights into the model's ability to minimize false positives and negatives, respectively. The F1-score, a harmonic mean of precision and recall, served as the primary metric for evaluating classification balance. Confusion matrices provided detailed analyses of misclassification trends, aiding in the identification of behaviors prone to errors. Additionally, real-time inference speed was measured to evaluate the model's practicality for deployment in traffic monitoring systems.

3.5. Hardware and Software Environment

The experimentation was conducted on a high-performance computing environment equipped with NVIDIA Tesla V100 GPUs and 32GB of memory. TensorFlow and Keras frameworks were employed for model implementation, offering flexibility and ease of integration with additional tools. Video preprocessing and augmentation were performed using OpenCV and Albumentations libraries, ensuring high-quality input for the neural network. The combination of state-of-the-art hardware and optimized software tools enabled efficient training, testing, and evaluation of the detection model, facilitating real-time performance capabilities.

4. Results and Analysis

4.1. Performance Metrics

The CNN-LSTM model demonstrated an ability to balance high precision and recall across different illegal driving behaviors, maintaining a consistent F1-score above 90% for all categories. This level of performance indicates a high degree of generalization and robustness. In comparison, previous studies using traditional ML methods often achieved performance metrics in the range of 70–80%, highlighting the advantage of deep learning in processing complex, non-linear patterns inherent in driving behavior. Furthermore, the higher precision in distracted driving detection showcases the model's sensitivity to subtle variations in driver posture and actions, critical for tackling one of the leading causes of road accidents globally.

4.2. Confusion Matrix Analysis

The confusion matrix revealed nuanced challenges in detecting behaviors with overlapping characteristics. For example, false positives for illegal lane changes often arose when vehicles momentarily strayed due to road conditions, without constituting actual violations. Similarly, false negatives in distracted driving cases occurred in situations with poor cabin lighting or occluded views of the driver's face. These findings suggest that refining pre-processing techniques, such as adaptive brightness adjustments or supplementary infrared imaging, could further enhance detection accuracy. The ability of the model to minimize misclassification in the majority of scenarios reinforces its applicability to diverse traffic environments.

4.3. Comparative Performance

The model was benchmarked against single-layer CNNs and LSTMs, demonstrating superior multi-modal learning capability. For instance, CNN alone excelled in spatial feature recognition, but lacked temporal context, often misclassifying short-term actions. LSTM alone, while proficient in analyzing temporal sequences, struggled with the rich spatial data necessary to differentiate nuanced actions like illegal lane changes. By combining the strengths of both, the hybrid architecture achieved balanced performance, offering insights into how spatial-temporal data fusion can enhance behavior detection in complex systems like traffic monitoring.

4.4. Real-Time Performance

Latency analysis highlighted the model's suitability for real-time application, with inference times of 35 ms per frame. This allows processing of high-definition video streams in dynamic traffic conditions, a critical factor for real-world deployment. Unlike rule-based systems that exhibit significant delays during computationally intensive scenarios, the proposed model benefits from optimized GPU acceleration. Such low-latency performance ensures that illegal behaviors can be flagged and recorded almost instantaneously, paving the way for practical integration into live surveillance infrastructure.

4.5. Robustness Testing

Testing the model across different environmental conditions revealed its adaptability. In low-light scenarios, the CNN component effectively leveraged contrast-enhancement techniques, while in adverse weather, edge-detection filters played a crucial role. The 2% reduction in accuracy under heavy rain, though minor, underscores the need for future integration of auxiliary data sources like LiDAR. Meanwhile, the model's stability in high-traffic scenarios, maintaining a 93% detection accuracy, indicates its potential for scaling across urban and suburban monitoring systems where vehicular density fluctuates widely.

5. Discussion

5.1. Interpretation of Results

The model's ability to achieve high detection accuracy across diverse illegal behaviors validates its architectural design, particularly the synergy between CNN for spatial feature extraction and LSTM for temporal pattern recognition. For instance, distracted driving detection was most successful due to the model's capacity to identify even brief deviations from normal postures, underscoring its potential to address critical road safety issues. However, the slight underperformance in lane-change detection highlights the challenges posed by overlapping vehicles and road complexities. Integrating advanced segmentation methods could address these limitations.

5.2. Practical Implications

The proposed model offers significant practical value by enabling automated detection of illegal behaviors, potentially revolutionizing traffic law enforcement. Real-time detection allows law enforcement agencies to intervene promptly, reducing accidents caused by distracted or reckless driving. Moreover, the scalability of the system for use in intelligent transportation infrastructure aligns with smart city initiatives. Beyond enforcement, the model could serve as a foundation for driver education programs by providing detailed feedback on unsafe practices detected during simulations or training.

5.3. Ethical and Social Considerations

The deployment of AI-based behavior detection systems raises important ethical concerns. Continuous monitoring might lead to privacy infringement, particularly if camera feeds capture sensitive personal data. Addressing these concerns requires implementing robust anonymization techniques, ensuring that only relevant behavioral information is processed and stored. Additionally, there is a risk of algorithmic bias stemming from imbalanced datasets, which could disproportionately flag certain demographic groups. A commitment to regular auditing and inclusive dataset collection is essential to foster public trust and system fairness.

5.4. Limitations

While the model's high performance demonstrates its potential, certain limitations remain. Its dependence on video data makes it susceptible to environmental challenges, such as fog or glare, where visual clarity is reduced. Similarly, the reliance on high-performance GPUs for real-time processing may limit its applicability in resource-constrained settings. Finally, the dataset's limited geographical and cultural diversity may affect the model's generalizability. Expanding the dataset to include varied road conditions and driving cultures will be crucial for global deployment.

5.5. Future Directions

Future research should explore the integration of multimodal sensor data to overcome visual challenges. For example, combining video input with radar, LiDAR, or audio data could enhance the system's robustness under adverse conditions. Additionally, attention-based neural architectures could refine the model's focus on key features, improving accuracy for behaviors like illegal lane changes. Another promising avenue is the development of lightweight variants for deployment on edge devices, enabling real-time processing in cost-sensitive environments. Finally, exploring the model's integration with autonomous vehicle systems could extend its application beyond monitoring, contributing to proactive safety interventions in self-driving cars.

6. Conclusion

This study developed a deep learning-based illegal driving behavior detection model using a hybrid CNN-LSTM architecture. The model effectively identifies behaviors such as distracted driving, illegal lane changes, and speeding by combining spatial feature extraction with temporal pattern analysis. Trained on a diverse dataset, it showed strong performance in terms of accuracy, precision, recall, and F1-score, making it suitable for real-time traffic monitoring and smart city applications. While the results are promising, further improvements can be made by expanding the dataset and incorporating additional sensor data, such as LiDAR or GPS, for better detection accuracy. Overall, the model contributes to road safety by offering an effective solution for identifying and addressing illegal driving behaviors. As smart infrastructure evolves, such models will play a crucial role in enhancing traffic management and reducing accidents.

Acknowledgements

This work is supported by Innovation and Entrepreneurship Training Project for College Students of Anhui University of Finance and Economics in 2024, Project number: 202410378128.

References

[1] Zhang X, Wang Y, Liu Z, et al. Predictive analysis of illegal driving behaviors using deep convolutional neural networks. Journal of Traffic and Transportation Engineering, 2023, 9(4): 1-15.

[2] Lee S, Park H, Choi J, et al. Real-time traffic monitoring using hybrid deep learning models. Journal of Intelligent Transportation Systems, 2022, 14(3): 45-58.

[3] Huang Y, Yang T, Xu L, et al. A hybrid model for illegal driving detection in smart cities. IEEE Transactions on Vehicular Technology, 2021, 70(5): 2340-2351.

[4] Li J, Zhou W, Tang X, et al. A deep learning-based approach to identify unsafe driving behaviors using video data. Transportation Research Part C: Emerging Technologies, 2021, 123: 314-325.

[5] Xu J, Zhao X, Li K, et al. Detection of driver distractions using convolutional and recurrent neural networks. IEEE Access, 2020, 8: 125-136.

[6] Sun W, Liu Y, Ma C, et al. Smart traffic monitoring system using deep learning for violation detection. International Journal of Computer Vision and Image Processing, 2022, 12(1): 22-36.

[7] Wang L, Chen Z, Zhang Q, et al. A deep learning model for vehicle behavior analysis in urban traffic. Journal of Artificial Intelligence and Traffic Safety, 2023, 5(2): 89-101.

[8] He X, Song Y, Zhang T, et al. Multimodal illegal driving behavior detection using deep neural networks. Computer Vision and Image Understanding, 2023, 214: 103-114.

[9] Li Q, Jiang P, Sun L, et al. Spatio-temporal pattern recognition of traffic violations using CNN-LSTM networks. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(7): 2025-2037.