Enhancing Road Safety: Leveraging CNN-LSTM and Bi-LSTM Models for Advanced Driver Behavior Detection

1st Ahmed Diaaeldin
Faculty of Computer Science
October University for
Modern Sciences and Arts (MSA)
Giza, Egypt
Email: ahmed.diaaeldin2@msa.edu.eg
ORCID: 0009-0007-9598-1217

2nd Moamen Zaher
Faculty of Computer Science
October University for
Modern Sciences and Arts (MSA)
Giza, Egypt
Email: moibrahem@msa.edu.eg
ORCID: 0009-0004-8560-4563

***Abstract*—Recognizing the importance of driver behavior is essential for enhancing road safety and optimizing traffic management systems. This study employs advanced deep learning techniques, specifically CNN-LSTM and Bi-LSTM models, to refine the prediction of driver behaviors using sensor data from the Honda Research Institute Driving Dataset (HDD). Our approach integrates a robust dataset encompassing a broad spectrum of sensor inputs, from vehicle dynamics to driver operational parameters, propelling advancements in driver behavior detection. The methodologies utilized enable the discernment of subtle and complex driving patterns, contributing to the reduction of road safety hazards. Our findings indicate that these models significantly improve the detection of hazardous driving behaviors, surpassing previous state-of-the-art methodologies with notable gains in mean average precision (mAP). These advancements affirm the potential of deep learning technologies in crafting sophisticated predictive safety systems, paving the way for future innovations.**

***Index Terms*—LSTM, CNN-LSTM, Bi-LSTM, Driver Behavior, Sensors**

I. INTRODUCTION

Dangerous driving behavior is a major factor contributing to road traffic accidents. Empirical research indicates that approximately 90% of road traffic accidents are due to human errors, with driver distraction and risky behaviors as primary contributors [1], [2]. Furthermore, it is estimated that around 95% of accidents involve driver-related dangerous behaviors, underscoring the significant impact of drivers' actions on road safety [3]. In 2015, 27% of traffic fatalities involved at least one driver who was speeding, alongside other aggressive behaviors such as failing to yield the right-of-way (7%) and unsafe lane changes (7.5%), highlighting the lethal consequences of aggressive driving [4], [5].

Machine learning (ML) and deep learning (DL) techniques have revolutionized the field of driver behavior detection, offering advanced capabilities to analyze complex datasets and predict driver actions with increased accuracy. These technologies enable a more nuanced understanding of driving behaviors than traditional rule-based systems, which struggle to interpret complex interactions and nuanced scenarios. General machine learning models such as Support Vector Machines (SVM), decision trees, and k-nearest neighbors (KNN) are extensively used to analyze various driving patterns by processing both vehicular data and contextual information from the driving environment [6]–[8]. Deep learning models, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and Long Short-Term Memory networks (LSTMs), further extend this capability, handling large volumes of unstructured data to detect subtle and complex patterns in driver behavior [9]–[11]. These models are trained on diverse datasets from real-world driving scenarios, enabling them to effectively predict driver reactions in real-time and improve overall road safety [12], [13]. The integration of these ML and DL models into driver assistance systems has led to significant improvements in the detection of hazardous driving behaviors, offering a robust framework for the development of predictive safety mechanisms.

Specific studies have focused on detecting critical driving behaviors like sudden acceleration, abrupt braking, rapid lane changes, and sharp turns, particularly using sensor data. Research has employed deep learning to detect reckless driving by analyzing video-extracted vehicle trajectory histograms, using data such as vehicle speed and trajectory derived from image sensors [9]. Similarly, lane-changing behavior detection systems that use sequence learning from dashcam footage determine lane deviations [10]. Additionally, smartphone sensors have been used to classify driving behaviors in real-world settings, demonstrating that accelerometers and gyroscopes can effectively detect rapid maneuvers and provide data to support the development of adaptive learning algorithms for improved accuracy [14]. These studies underscore the potential of using sensor-based data to improve the detection and prediction of hazardous driving behaviors, which could lead to more responsive and effective driver assistance systems.

The structure of this paper is as follows: The 'Related Work' section reviews current methodologies and deep learning models utilized for driver behavior detection. The 'Dataset' section provides an in-depth analysis of the dataset, highlighting its relevance to the study. The 'Methodology' section details the driver behavior classification models, including input processing and the evaluation metrics employed. In the 'Results and Discussion' section, the proposed model's results is compared with existing state-of-the-art approaches, followed by a discussion of the study's broader implications. Finally, the 'Future Work' section suggests potential avenues for further research.

II. Related Work

In systems that are specifically designed to efficiently supervise vehicle use, maintenance, and administrative tasks as fleet management systems the detection of driver behavior is crucial for improving operational efficiency and cutting costs. By collecting data on driver behaviors, these systems can offer valuable insights to improve driver safety and minimize the occurrence of accidents. Such systems utilize objective data to detect risky behaviors, provide feedback to drivers, and reduce crash risks [15].

However, monitoring systems frequently confront challenges such as privacy concerns, where drivers may perceive an invasion of their personal space due to the use of cameras, consequently engendering resistance to such monitoring practices [16]. Additionally, the management of the voluminous data generated necessitates extensive storage solutions and robust security measures to thwart potential breaches [17]. In response, sensors offer a less intrusive alternative, ameliorating privacy and data management issues by monitoring physiological and vehicular operational data to proficiently identify aggressive driving behaviors [18]. The privacy issue not only pertains to drivers but also extends to pedestrians, who may be subjected to unauthorized video surveillance, thereby complicating the ethical deployment of monitoring technologies.

In the quest for minimally invasive monitoring solutions, the integration of diverse sensor technologies such as GPS, gyroscopes, accelerometers, and Controller Area Network (CAN) sensors is imperative. These sensors adeptly track vehicle dynamics and driver behaviors with minimal intrusion into privacy. GPS systems provide precise real-time location data, while accelerometers and gyroscopes deliver valuable insights into vehicular motion and orientation, aiding in the detection of aggressive driving patterns without the need for visual data capture [19]. CAN sensors, by monitoring a range of vehicle metrics from speed to brake pressure, enrich the data set, facilitating a thorough analysis of driving habits without compromising the privacy of drivers or pedestrians [20]. Collectively, these technologies constitute a robust framework for detecting aggressive driving, emphasizing both effectiveness and ethical considerations.

A. Sensors

The integration of car and mobile sensors with machine learning (ML) and deep learning (DL) techniques has significantly transformed the landscape of driver behavior prediction. Utilizing a wide array of sensors embedded in vehicles and smartphones, researchers have developed sophisticated models that leverage vast amounts of data to evaluate and predict driver behavior with remarkable accuracy. These systems not only enhance road safety but also aid in managing fuel consumption and improving vehicle sustainability [21].

Deep learning models, particularly those employing convolutional neural networks (CNNs) and Long Short-Term Memory (LSTM) networks, have proven exceptionally adept at analyzing sensor data. This capability allows for the prediction of specific behaviors such as aggressive driving, drowsiness, and even the driver's focus of attention using real-time data streams [22].

The diversity of sensors in modern vehicles provides a rich dataset that enhances the functionality of driver assistance systems. By monitoring everything from vehicle speed to environmental conditions, these sensors offer insights into driver behavior and vehicle performance [23].

The synergy between car and mobile sensors with ML and DL technologies marks a significant advance in intelligent transportation systems. This integration not only boosts the accuracy of predicting driver behavior but also plays a crucial role in enhancing road safety, optimizing vehicle performance, and promoting sustainable driving practices. The continued development of these technologies promises even greater improvements in automotive safety and efficiency in the near future [24], [25].

B. Cameras

The integration of cameras with machine learning and deep learning significantly enhances the prediction and understanding of driver behavior. Modern studies have leveraged video data for crucial safety features such as detecting driver drowsiness using eye movement data and implementing in-cabin monitoring systems for real-time alerts [26], [27]. These methods, combined with deep learning techniques, improve road safety by providing real-time driver monitoring and behavior analysis [28].

C. Multi-model

Moreover, for further improvement in the prediction of driver behavior, integrating multi-modal inputs provides a comprehensive understanding of driver actions. This integration, using advanced deep learning algorithms like CNNs and LSTMs, enhances the prediction capabilities by analyzing data from diverse sources including cameras and sensors, thus improving the precision in detecting aggressive driving behaviors [29], [30].

III. Dataset

The Honda Research Institute Driving Dataset (HDD) is utilized in this research. HDD is a comprehensive dataset specifically created to facilitate research on driver behavior learning and causal reasoning in real-world environments [31]. The dataset comprises 104 hours of actual human driving in the San

Francisco Bay Area. The data was collected using a specially equipped vehicle fitted with various sensors, including three video cameras, a LiDAR sensor, an Automotive Dynamic Motion Analyzer, and a Vehicle Controller Area Network (CAN). This research focused on eight sensors, shown in figure 1, all sampled at 3 Hz. Of these, six sensors have been used in previous work: Acceleration Pedal, Steering Angle, Steering Speed, Speed, Brake Pedal, and Yaw Rate. Additionally, two turn signal sensors were included for analysis. The dataset has been divided into 137 driving sessions, with each session annotated using a novel methodology that incorporates a 4-layer annotation scheme. This scheme describes driver behaviors across four layers: Goal-oriented action, Stimulus-driven action, Cause, and Attention. The focus in this study was on the 11 goal-oriented actions, namely intersection passing, left turn, right turn, left lane change, right lane change, left lane branch, right lane branch, crosswalk passing, railroad passing, merging, and u-turn.

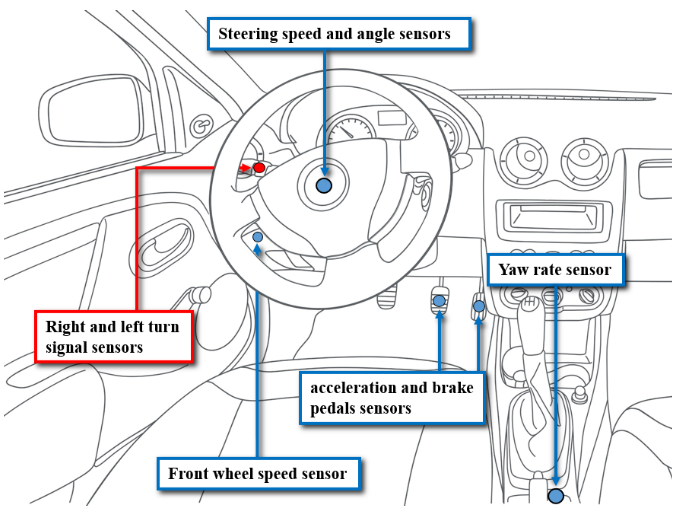

Fig. 1. The placements of the 8 sensors included in the dataset. The blue color represents the six sensors that were utilized in previous research studies, while the red color represents the two additional turn signal sensors.

IV. Methodology

The methodology section is organized into several subsections as follows: Firstly, 'Pre-processing,' which elucidates the approaches to data processing and addresses specific challenges encountered. Next, 'Action Classification' is dedicated to discussing the deep learning models selected for this study, including a justification for their use. Finally, 'Evaluation Metrics' describes the criteria and metrics employed to assess model performance.

A. Pre-processing

The train/test split outlined in the original dataset paper [31] and followed by all related work [32]–[36] was adhered to, dividing the 137 driving sessions into 100 sessions for training and 37 sessions for testing. This split was based on geolocation data to ensure minimal spatial overlap between training and test routes, preventing the model from being tested on the exact same locations used during training [31]. The training data was then divided into 90 training sessions and 10 validation sessions during training, an approach not implemented in previous research. The sequence length was set to 90, as used in earlier studies. Finally, past processing was conducted using six sensors to align with prior work, while two additional sensors were incorporated to analyze their impact on prediction accuracy.

B. Action Classification

Deep learning, particularly Long Short-Term Memory (LSTM) networks, has significantly advanced the detection of driver behaviors by leveraging complex data from sensors to identify nuanced patterns. Numerous studies have demonstrated the effectiveness of deep learning models in detecting various driver behaviors such as inattentiveness, aggression, distraction, and drowsiness [37]–[40]. These models have shown proficiency in processing time-series sensor data, enabling accurate recognition of driver states and behaviors, ultimately contributing to enhanced road safety and driver monitoring systems. The utilization of deep learning techniques has notably improved the accuracy and efficiency of driver behavior detection systems, showcasing the potential of these technologies in ensuring safer driving environments.

Building on the foundation laid by LSTM, Convolutional Neural Networks combined with Long Short-Term Memory networks (CNN-LSTM) have further enhanced the capability of deep learning models in this domain. Recent research discusses how sensor-based approaches, empowered by CNN-LSTM frameworks, adeptly identify aggressive driving patterns in two-wheeled vehicles [41]. Other researchers employ a CNN-LSTM car-following model that excels in various traffic scenarios, illustrating the generalization ability of these models across diverse driving conditions [42]. Additional studies validate the application of CNN-LSTM models for driver drowsiness prediction, showcasing their effectiveness in real-time environments [22]. These studies collectively affirm that CNN-LSTM frameworks are highly efficient in real-time, accurate behavioral assessment, leveraging both spatial and temporal data patterns captured from sensors.

In general time series data handling, Bi-directional Long Short-Term Memory (Bi-LSTM) networks have proven exceptionally effective as they analyze both past and future contexts [43]. This dual-direction approach is highly suitable for complex applications such as environmental modeling and stock price forecasting [44], [45]. These networks excel in diverse domains, leveraging their capacity to capture intricate temporal dependencies within sequential data. The versatility and superior performance of Bi-LSTM models make them indispensable tools in tasks requiring robust time series predictions [44], [46].

These models were selected due to their widespread application and the promising performance highlighted in the related work section.

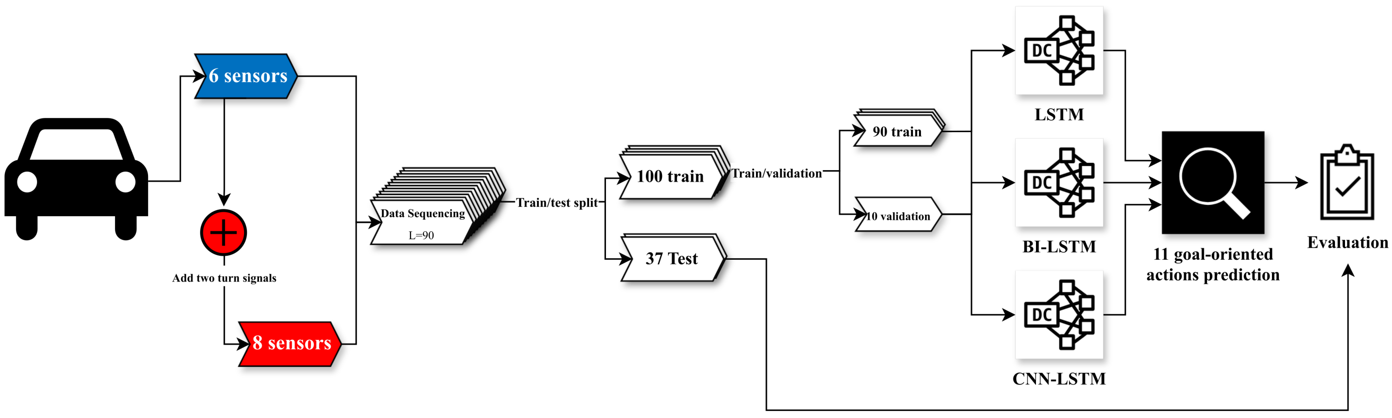

Fig. 2. System Architecture

C. Evaluation Metrics

Following previous works [32]–[36], per-frame mean average precision (mAP) was used to evaluate this work. This widely employed metric involves computing the average precision (AP) for each action class, followed by calculating their overall average [33]. In addition to mAP, precision, recall, and F1-score were also utilized to provide a comprehensive evaluation of model performance. However, mAP was listed in comparisons since it was consistently used in all state-of-the-art references, allowing for direct and standardized comparisons across different studies.

V. Results and Discussion

This section discusses the experiments conducted to evaluate the models. The first experiment utilized six sensors from the dataset, shown in Fig.1 in blue, to align with prior studies. In the second experiment, the analysis was extended by including two turn signal sensors, shown in Fig.1 in red, to assess their influence on model accuracy. During model training, various techniques were employed to enhance performance and promote generalization, such as EarlyStopping and ReduceLROnPlateau. Additional details regarding training constraints are provided in TABLE I.

TABLE I
Constant parameters across all experiments.

Parameter	Value
Batch Size	80
Learning Rate	Reduce Learning Rate on Plateau with an initial value of 0.001 with a patience of 4
Loss	Categorical Cross-Entropy
Number of Epochs	50 with Early Stop Callback with patience of 8
Loss Optimizer	Adam
Output layer Activation	Softmax

A. Experiment I

This section aims to conduct a comparative study between the selected models using the employed preprocessing and training methodologies. The results are then compared against the state-of-the-art. Six sensors were used to ensure comparability with past research.

TABLE II
Comparison of proposed models overall mean average precision with state-of-the-art results

Method	mAP
CNN [32]	22.7
LSTM [31]	23.8
ED [32]	27.4
TRN [32]	29.2
OadTR [33]	29.8
Colar [34]	30.6
GateHUB [35]	32.1
MAT [36]	32.7
LSTM (Ours)	28.8
BI-LSTM (Ours)	**33.1**
CNN-LSTM (Ours)	**36.7**

The results presented in TABLE II indicate that the CNN-LSTM model significantly outperforms previous state-of-the-art models in overall mean average precision, exceeding the most recent MAT model [36] by 4%. The BI-LSTM model achieved a modest improvement over MAT by 0.4%, while the LSTM model surpassed its state-of-the-art counterpart by 6.1%, demonstrating the impact of minor adjustments during training. Ultimately, the CNN-LSTM model proved advantageous, underscoring its superior suitability for this specific problem.

The results presented in TABLE III reveal advantages in some classes regarding per-class average precision, such as turns and merging, while highlighting predictive failures across other classes, such as crosswalk and railroad passing. Although state-of-the-art models demonstrated better results in these classes, the overall performance remains inadequate for practical application. This trend is consistent across all state-of-the-art results, suggesting that accurate prediction may

TABLE III
COMPARISON OF PROPOSED MODELS PER CLASS AVERAGE PRECISION WITH STATE-OF-THE-ART RESULTS

Method	Inter. Pass.	Left Turn	Right Turn	Left Lane Change	Right Lane Change	Left Lane Branch	Right Lane Branch	Cross. Pass.	Rail. Pass.	Merge	U-turn
CNN [32]	34.2	72.0	74.9	16.0	8.5	7.6	1.2	0.4	0.1	2.5	32.5
LSTM [31]	36.4	66.2	74.2	26.1	13.3	8.0	0.2	0.3	0.0	3.5	33.5
ED [32]	43.9	73.9	75.7	31.8	15.2	15.1	2.1	0.5	0.1	4.1	39.1
TRN [32]	46.5	**75.2**	77.7	35.9	19.7	18.5	**3.8**	**0.7**	**0.1**	2.5	40.3
OadTR [33]	-	-	-	-	-	-	-	-	-	-	-
Colar [34]	-	-	-	-	-	-	-	-	-	-	-
GateHUB [35]	-	-	-	-	-	-	-	-	-	-	-
MAT [36]	-	-	-	-	-	-	-	-	-	-	-
LSTM	**67.6**	74.9	75.0	37.6	0.0	0.0	0.0	0.0	0.0	7.4	54.4
BI-LSTM	66.3	70.9	**82.3**	**50.0**	**33.3**	0.0	0.0	0.0	0.0	**25.0**	36.6
CNN-LSTM	65.4	71.8	77.1	31.1	22.2	**55.6**	0.0	0.0	0.0	21.1	**59.2**

Models represented with dashes (-) on specific class names did not report per-class precision metrics in their publications, providing only the overall mean Average Precision (mAP).

require additional features, particularly visual ones, rather than relying solely on the six sensors provided.

B. Experiment II

TABLE IV
COMPARISON BETWEEN PROPOSED MODELS OVERALL MEAN AVERAGE PRECISION WHEN USING 6 AND 8 SENSORS

Features	Method	mAP
6	LSTM	28.8
	BI-LSTM	33.1
	CNN-LSTM	**36.7**
8	LSTM	38.4
	BI-LSTM	36.0
	CNN-LSTM	**40.1**

TABLE IV shows that incorporating two turn signal sensors improved overall mean average precision across all models, with LSTM demonstrating the highest increase of 9.6%. Additionally, the CNN-LSTM model still shows the highest overall mAP with a 3.4% improvement over its six-sensor version. These enhancements underscore the benefit of adding the extra sensors for improved prediction accuracy.

Additionally, The results in TABLE VI demonstrate improved precision across most classes for each model after incorporating two additional sensors, indicating the benefit of adding more sensors to enhance overall class mean average precision.

C. Model Complexity

The results in TABLE V demonstrate that the CNN-LSTM model continuously exhibited the shortest testing time among all models evaluated, highlighting its efficiency. Additionally, The LSTM model demonstrated a significant size advantage, being almost half the size of other models while achieving a comparable testing time to CNN-LSTM. This combination of reduced size, rapid testing time, and increased accuracy with eight sensors makes it a strong contender for applications that

TABLE V
COMPARISON BETWEEN EVERY PROPOSED MODEL'S COMPLEXITY WHEN USING 6 AND 8 SENSORS

Features	Method	Model Size (MB)	Parameters	Testing Time (s)
6	**LSTM**	**3.5**	**303,756**	28.8
	BI-LSTM	6.98	605,836	44.31
	CNN-LSTM	6.52	565,644	**27.09**
8	**LSTM**	**3.53**	**305,804**	28.92
	BI-LSTM	7.02	609,932	43.75
	CNN-LSTM	6.53	567,180	**25.49**

prioritize these benefits. Moreover, CNN-LSTM also showed a small advantage over BI-LSTM in size in addition to testing time, which further demonstrates how well it fits this prediction problem.

VI. CONCLUSION

This research have demonstrated the capabilities of deep learning models, specifically CNN-LSTM, LSTM, and Bi-LSTM, in predicting driver behavior using sensor data from the Honda Research Institute Driving Dataset (HDD). Experiments results show substantial improvements in model performance, thereby contributing to enhanced road safety and driver monitoring systems.

The first experiment focused on the use of six sensors, where the CNN-LSTM model emerged as a standout, outperforming all previous models by achieving a mean average precision (mAP) increase of 4% over the most recent state-of-the-art models. This significant improvement not only highlights the efficacy of the CNN-LSTM integration but also underscores its robustness in handling complex, multi-dimensional data.

The Bi-LSTM model, with its unique ability to process data from both past and future contexts, demonstrated a modest performance advantage over state-of-the-art models, specifically a 0.4% increase in mAP. However, the CNN-LSTM model showcased further advantages over the Bi-LSTM in terms

TABLE VI

COMPARISON OF PROPOSED MODELS PER CLASS AVERAGE PRECISION WHEN USING 6 AND 8 SENSORS

Features	Method	Inter Pass.	Left Turn	Right Turn	Left Lane Change	Right Lane Change	Left Lane Branch	Right Lane Branch	Cross. Pass.	Rail. Pass.	Merge	U-turn
6	**LSTM**	67.6	74.9	75.0	37.6	0.0	0.0	0.0	0.0	0.0	7.4	54.4
	BI-LSTM	66.3	70.9	82.3	50.0	33.3	0.0	0.0	0.0	0.0	25.0	36.6
	CNN-LSTM	65.4	71.8	77.1	31.1	22.2	**55.6**	0.0	0.0	0.0	21.1	**59.2**
8	**LSTM**	63.90	**76.60**	**82.58**	55.25	38.70	33.33	0.0	0.0	0.0	16.67	56.36
	BI-LSTM	64.76	74.03	78.21	55.11	40.08	19.05	0.0	0.0	0.0	20	44.89
	CNN-LSTM	**65.42**	76.28	80.86	**56.13**	**52.24**	33.33	0.0	0.0	0.0	**50**	27.2

of performance, model size, and testing time, reinforcing its superior suitability for this specific problem.

In the second experiment, the inclusion of two additional turning signal sensors provided a clearer understanding of driver intent, particularly beneficial for LSTM's architecture, which showed an impressive accuracy improvement of 9.6% in overall mean average precision. This highlights the LSTM's enhanced performance when analyzing extended sensor data, making it particularly useful for applications where compact model size and faster testing times are crucial. Moreover, the CNN-LSTM model again demonstrated superior performance in this setup, achieving a 3.4% improvement in mAP over its six-sensor configuration, further confirming its effectiveness and reliability in more complex sensor integrations.

In summary, the refined accuracy of these models, especially under different sensor configurations, provides a promising direction for future enhancements in driver assistance technologies. The demonstrated advancements in CNN-LSTM and LSTM models not only enhance decision-making but also ensure a higher level of safety in dynamic driving environments.

VII. FUTURE WORK

For future advancements in this research, exploring the utility of GPS data, which is available in the Honda Research Institute Driving Dataset (HDD) but not currently mapped to annotated session videos, presents a promising avenue. By integrating GPS data with the existing eight-sensor setup, models could be more enrich with spatial contextual information that may enhance the prediction accuracy of driver behaviors.

Furthermore, investigating the effectiveness of GPS data only for prediction could serve as a low-cost solution within the application market if it demonstrates sufficient accuracy. This approach would reduce reliance on more complex sensor systems, offering a simpler and potentially more cost-effective method for monitoring and predicting driver behaviors.

REFERENCES

[1] M. Siam, M. Isa, N. Borhan, A. Sukardi, and W. Voon, "Measurement of driver distraction in malaysia's traffic environment: a driving simulator study," *Journal of Mechanical Engineering and Sciences*, vol. 8, pp. 1472–1480, 2015.

[2] A. Hassen, A. Godesso, L. Abebe, and E. Girma, "Risky driving behaviors for road traffic accident among drivers in mekele city, northern ethiopia," *BMC Research Notes*, vol. 4, no. 1, 2011.

[3] M. Bazzaz, A. Zarifian, M. Emadzadeh, and V. Vakili, "Driving behaviors in iran: A descriptive study among drivers of mashhad city in 2014," *Global Journal of Health Science*, vol. 7, no. 7, 2015.

[4] L. Tasca, "A review of the role of aggressive driving in traffic accidents," *Aggressive Behavior*, 2000.

[5] National Center for Statistics and Analysis, "Traffic safety facts: 2017 data," *U.S. Department of Transportation*, 2017.

[6] L. Zhang, L. Yan, Y. Fang, X. Fang, and X. Huang, "A machine learning-based defensive alerting system against reckless driving in vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 68, pp. 12227–12238, 2019.

[7] H. Woo, Y. Ji, H. Kono, Y. Tamura, Y. Kuroda, T. Sugano, Y. Yamamoto, A. Yamashita, and H. Asama, "Lane-change detection based on vehicle-trajectory prediction," *IEEE Robot. Autom. Lett.*, vol. 2, pp. 1109–1116, 2017.

[8] Y. Zhang, X. Shi, S. Zhang, and A. Abraham, "A xgboost-based lane change prediction on time series data using feature engineering for autopilot vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, pp. 19187–19200, 2022.

[9] J.-M. Jang, "Detection of reckless driving using deep learning," pp. 853–858, 2020.

[10] J. Gao, Y. Murphey, J. Yi, and H. Zhu, "A data-driven lane-changing behavior detection system based on sequence learning," *Transp. B Transp. Dyn.*, vol. 10, pp. 831–848, 2020.

[11] Z. Chen, C. Wu, Z. Huang, N. Lyu, Z. Hu, M. Zhong, Y. Cheng, and B. Ran, "Dangerous driving behavior detection using video-extracted vehicle trajectory histograms," *J. Intell. Transp. Syst.*, vol. 21, pp. 409–421, 2017.

[12] Q. Shangguan, T. Fu, J. Wang, T. Luo, and S. Fang, "An integrated methodology for real-time driving risk status prediction using naturalistic driving data," *Accid. Anal. Prev.*, vol. 156, p. 106122, 2021.

[13] A. Dosovitskiy, G. Ros, F. Codevilla, A. López, and V. Koltun, "Carla: An open urban driving simulator," in *Conference on Robot Learning*, pp. 1–16, 2017.

[14] H. Eftekhari and M. Ghatee, "A similarity-based neuro-fuzzy modeling for driving behavior recognition applying fusion of smartphone sensors," *J. Intell. Transp. Syst. Technol. Plan. Oper.*, vol. 23, pp. 72–83, 2019.

[15] A. K. Pradhan, J. T. Lin, C. Wege, and F. Babel, "Effects of behavior-based driver feedback systems on commercial long haul operator safety," in *Driving Assessment Conference*, vol. 9, University of Iowa, 2017.

[16] O. Kumtepe, G. Akar, and E. Yuncu, "Driver aggressiveness detection using visual information from forward camera," in *2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1–6, 2015.

[17] J. Lee and K. Jang, "A framework for evaluating aggressive driving behaviors based on in-vehicle driving records," *Transportation Research Part F: Traffic Psychology and Behaviour*, 2017.

[18] I. Lashkov and A. Kashevnik, "Aggressive behavior detection based on driver heart rate and hand movement data," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pp. 1490–1495, 2021.

[19] R. Chhabra, S. Verma, and C. Krishna, "Detecting aggressive driving behavior using mobile smartphone," in *Proceedings of 2nd International Conference on Communication, Computing and Networking*, 2018.

[20] D. A. Johnson and M. Trivedi, "Driving style recognition using a smartphone as a sensor platform," in *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 1609–1615, 2011.

[21] K. Vellenga, H. Steinhauer, A. Karlsson, G. Falkman, A. Rhodin, and A. Koppisetty, "Driver intention recognition: state-of-the-art review," *Ieee Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 602–616, 2022.
[22] M. Gomaa, R. Mahmoud, and A. Sarhan, "A cnn-lstm-based deep learning approach for driver drowsiness prediction," *Journal of Engineering Research*, vol. 0, pp. 0–0, 2022.
[23] A. Khandakar, M. Chowdhury, R. Ahmed, A. Dhib, M. Mohammed, N. Al-Emadi, and D. Michelson, "Portable system for monitoring and controlling driver behavior and the use of a mobile phone while driving," *Sensors*, vol. 19, p. 1563, 2019.
[24] E. Khosravi, A. Hemmatyar, M. Siavoshani, and B. Moshiri, "Safe deep driving behavior detection (s3d)," *Ieee Access*, vol. 10, pp. 113827–113838, 2022.
[25] P. Khosravinia, "Enhancing road safety through accurate detection of hazardous driving behaviors with graph convolutional recurrent networks," *Ieee Access*, pp. 1–1, 2023.
[26] J. Waleed, T. Abbas, and T. Hasan, "Implementation of driver's drowsiness assistance model based on eye movements detection," *Eastern-European Journal of Enterprise Technologies*, vol. 5, pp. 6–13, 2020.
[27] S. Suryavanshi, "In cabin driver monitoring and alerting system for passenger cars using machine learning.," *Journal of Physics Conference Series*, vol. 2601, p. 012040, 2023.
[28] A. Ziryawulawo, M. Kirabo, C. Mwikirize, J. Serugunda, E. Mugume, and S. Miyingo, "Machine learning based driver monitoring system: a case study for the kayoola evs," *Saiee Africa Research Journal*, vol. 114, pp. 40–48, 2023.
[29] P. Khosravinia, "Enhancing road safety through accurate detection of hazardous driving behaviors with graph convolutional recurrent networks," *Ieee Access*, pp. 1–1, 2023.
[30] J. Bajaj, N. Kumar, R. Kaushal, G. L, F. Flammini, and R. Natarajan, "System and method for driver drowsiness detection using behavioral and sensor-based physiological measures," *Sensors*, vol. 23, p. 1292, 2023.
[31] V. Ramanishka, Y.-T. Chen, T. Misu, and K. Saenko, "Toward driving scene understanding: A dataset for learning driver behavior and causal reasoning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7699–7707, 2018.
[32] M. Xu, M. Gao, Y.-T. Chen, L. S. Davis, and D. J. Crandall, "Temporal recurrent networks for online action detection," in *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 5532–5541, 2019.
[33] X. Wang, S. Zhang, Z. Qing, Y. Shao, Z. Zuo, C. Gao, and N. Sang, "Oadtr: Online action detection with transformers," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 7565–7575, 2021.
[34] L. Yang, J. Han, and D. Zhang, "Colar: Effective and efficient online action detection by consulting exemplars," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3160–3169, June 2022.
[35] J. Chen, G. Mittal, Y. Yu, Y. Kong, and M. Chen, "Gatehub: Gated history unit with background suppression for online action detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 19925–19934, 2022.
[36] J. Wang, G. Chen, Y. Huang, L. Wang, and T. Lu, "Memory-and-anticipation transformer for online action understanding," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 13824–13835, 2023.
[37] M. Alkinani, W. Khan, and Q. Arshad, "Detecting human driver inattentive and aggressive driving behavior using deep learning: recent advances, requirements and open challenges," *Ieee Access*, vol. 8, pp. 105008–105030, 2020.
[38] M. Dua, R. Singla, S. Raj, and A. Jangra, "Deep cnn models-based ensemble approach to driver drowsiness detection," *Neural Computing and Applications*, vol. 33, pp. 3155–3168, 2020.
[39] P. Panwar, P. Roshan, R. Singh, M. Rai, A. Mishra, and S. Chauhan, "Ddnet- a deep learning approach to detect driver distraction and drowsiness," *Evergreen*, vol. 9, pp. 881–892, 2022.
[40] V. Chirra, U. Reddy, and V. Kolli, "Deep cnn: a machine learning approach for driver drowsiness detection based on eye state," *Revue D Intelligence Artificielle*, vol. 33, pp. 461–466, 2019.
[41] D. Kim, "The detection of aggressive driving patterns in two-wheeled vehicles using sensor-based approaches," *Applied Sciences*, vol. 13, p. 12475, 2023.
[42] P. Qin, H. Li, Z. Li, W. Guan, and Y. He, "A cnn-lstm car-following model considering generalization ability," *Sensors*, vol. 23, p. 660, 2023.
[43] Z. Huang, W. Xu, and K. Yu, "Bidirectional lstm-crf models for sequence tagging," *arXiv preprint arXiv:1508.01991*, 2015.
[44] C. Han and X. Fu, "Challenge and opportunity: deep learning-based stock price prediction by using bi-directional lstm model," *Frontiers in Business Economics and Management*, vol. 8, pp. 51–54, 2023.
[45] Z. Li, "Modelling nitrogen oxide emission trends from the municipal solid waste incineration process using an adaptive bi-directional long and short-term memory network," *The Canadian Journal of Chemical Engineering*, vol. 102, pp. 1225–1237, 2023.
[46] J. Liu, "Prediction of existing tunnel deformation induced by shield undercrossing based on bo-bi-lstm," 2023.