AI-Driven Framework for Real-Time Prediction of Microscopic and Macroscopic Driving Risk Using Holistic Data

Dimitrios I. Tselentis
Research Associate
Department of Transportation Planning and Engineering
National Technical University of Athens, Athens, Greece, GR15773
Email: dtsel@central.ntua.gr

Thodoris Garefalakis
Research Associate
Department of Transportation Planning and Engineering
National Technical University of Athens, Athens, Greece, GR15773
Email: tgarefalakis@mail.ntua.gr

Dimitrios Nikolaou
Research Associate
Department of Transportation Planning and Engineering
National Technical University of Athens, Athens, Greece, GR15773
Email: dnikolaou@mail.ntua.gr

Eva Michelaraki
Research Associate
Department of Transportation Planning and Engineering
National Technical University of Athens, Athens, Greece, GR15773
Email: evamich@mail.ntua.gr

George Yannis
Professor
Department of Transportation Planning and Engineering
National Technical University of Athens, Athens, Greece, GR15773
Email: geyannis@central.ntua.gr

Word Count: 6624 words + 2 tables (250 words per table) = 7,124 words

Submitted: July 29, 2024

ABSTRACT

Accurate estimation of driving risk is essential for improving urban road safety and reducing crashes. Despite advancements, a significant gap exists in the real-time estimation of microscopic driving risk using advanced data technologies. This study presents an AI-based framework for assessing microscopic driving risk using the pNEUMA dataset, which comprises drone-collected traffic data from Panepistimiou Street, a five-lane urban arterial in Athens, Greece. Over five days, ten drones captured three hours of traffic data daily. The study aims to predict the probability of the ego vehicle encountering risk-related events such as speeding, lane changing, harsh braking, harsh acceleration, and short time-to-collision scenarios. The proposed methodology employs Long Short-Term Memory (LSTM) neural networks, both uni-directional and bi-directional, to analyze and predict driving risks based on the behavior of surrounding traffic. LSTM models showed high precision in identifying risk-related events, demonstrating the effectiveness of deep learning techniques in this context. Additionally, the results were aggregated at the road section level, allowing real-time risk estimation across different arterial subsections. This aggregation supports the development of real-time monitoring systems that provide actionable feedback to drivers, thereby enhancing road safety. The findings highlight the potential of integrating comprehensive data sources like drones to monitor and assess real-time driving behavior using advanced AI techniques. This approach offers granular insights into both micro and macroscopic driving risks. The study paves the way for future research into the application of similar methodologies in various urban settings, ultimately aiming to develop comprehensive, real-time road safety systems.

Keywords: Traffic Safety; Driving Risk Prediction; Microscopic and Macroscopic Risk; Real-Time Monitoring; Drone Data; LSTM Models

INTRODUCTION

Microscopic driving behavior analysis

Road safety is a global concern with traffic crashes causing significant human, economic, and social costs. Approximately 1.19 million people lose their lives annually in road traffic crashes, and 20 to 50 million suffer non-fatal injuries (*1*). Human error has proven a major cause of these crashes (*2*). Understanding and mitigating these errors can reduce the frequency and severity of road incidents.

Traffic models are categorized into three major types according to the scope of analysis and the desired degree of granularity: macroscopic, mesoscopic, and microscopic. Macroscopic analysis examines overall traffic flow characteristics like vehicle density and speed. Microscopic analysis focuses on individual driver behaviors and vehicle interactions, using sub-models like car-following and lane-changing to simulate traffic at the vehicle level. Mesoscopic analysis offers an intermediate perspective, capturing group behaviors without the full complexity of microscopic models.

Microscopic and macroscopic driving risk estimation

A comprehensive assessment of driving risks requires integrating both microscopic and macroscopic approaches. This holistic perspective enhances the precision of risk estimations. Microscopic analysis focuses on the actions and interactions of individual drivers and vehicles, crucial for understanding behaviors that affect overall traffic flow and safety. For example, frequent lane changes can disrupt traffic and increase collision risks (*3*). By simulating these behaviors, researchers can identify high-risk situations and develop targeted interventions (*4*). Advanced models using microscopic data can enhance the predictive capabilities of risk perception and warning systems, providing real-time alerts to high-risk drivers (*4*).

In contrast, macroscopic risk estimation models analyze aggregated traffic data to identify broader trends and potential risks. These models are valuable for urban planning and large-scale traffic management, predicting how changes in traffic patterns or infrastructure impact safety. For instance, speed dispersion has been identified as a reliable macroscopic indicator for assessing road safety, as it correlates with microscopic risks such as potential collisions (*5*). Advances in data analysis and machine learning have led to sophisticated models that improve the prediction and mitigation of driving risks, thus enhancing overall traffic safety.

Integrating microscopic and macroscopic approaches provides a comprehensive framework for understanding and mitigating road safety risks. Despite their importance, this integration has not been extensively explored, especially with groundbreaking data collection technologies like drones.

Objectives

To address the gaps in existing research, this study leverages drone data to monitor and predict driving risk on an urban arterial, contributing to the sparse literature in this domain. Through the implementation of advanced aerial data collection techniques, this research aims to obtain real-time, high-resolution information on driving behavior and traffic patterns that are typically challenging to holistically observe via traditional ground-based methods. The primary objective of this research is the development of AI-based models that make use of microscopic data to provide estimations of traffic risk probability, both on a microscopic and macroscopic level.

The paper is structured as follows: At the beginning, a thorough literature review is presented to highlight the gaps that this research aims to address. Subsequently, the research methodology is outlined, followed by a detailed description of the data and preprocessing steps. Finally, of the LSTM model on both microscopic and macroscopic driving risk probabilities are presented and discussed leading to significant conclusions.

BACKGROUND

Microscopic Driving Behavior and Risk

A comprehensive understanding of individual vehicle movements and driver interactions on the road is essential for assessing risks and developing interventions to improve road safety. Various scientific approaches and models have been employed to study microscopic driving behavior and risk.

Car-following behavior, a fundamental aspect of microscopic driving analysis, has been extensively studied. Zatmeh-Kanj and Toledo (2021) focused on the impact of driver distractions, such as phone usage, on car-following behavior using a driving simulator. Their findings revealed that drivers showed reduced awareness of the preceding vehicle while distracted, leading to increased speed volatility and reduced safety margins (*6*). Additionally, Tan et al. (2022) developed a risk field model that integrates behavior theories like risk homeostasis theory and preview-follower theory to uniformly model driving behavior in different scenarios. This model was validated using naturalistic data in car-following scenarios, demonstrating its effectiveness in risk quantification and motion planning (*7*).

In the realm of lane-changing behavior, Chen et al. (2021) conducted a comprehensive study on lane-changing risks by analyzing a vehicle trajectory dataset (HighD dataset). They proposed a lane-changing risk index (LCRI) and revealed that risk levels were associated with variables such as gap distance, vehicle speed, and acceleration. This study provided an in-depth analysis of the factors influencing the risks associated with lane-changing, thus offering a tool for the development of effective traffic safety measures (*8*). Furthermore, Ali et al. (2020) examined the effects of connected environments on discretionary lane changing (DLC) decision-making. Their study revealed that drivers in connected environments exhibited larger spacing, longer DLC durations, and lower acceleration noise, indicating improved safety and situational awareness (*9*).

Gap acceptance behavior at intersections, another critical component of microscopic driving behavior analysis, was examined by Li et al. (2020). They investigated the impact of mobile phone use on gap acceptance behavior at intersections using a driving simulator. While the distraction did not significantly affect gap acceptance decisions, it increased the time taken to complete intersection crossings, highlighting the complex interplay between driver distraction and gap acceptance behavior (*10*).

Estimation of Real-time and/or Microscopic Crash Probability

In the context of microscopic driving behavior analysis, numerous studies emphasize the estimation of real-time and microscopic crash probability, highlighting its critical role in enhancing traffic safety and enabling proactive traffic management. Haque et al. (2020) developed a model to estimate head-on crash probability using real-time motion trajectory data on two-lane undivided highways. Their model considered drivers' overtaking decisions and time-to-collision, incorporating factors like vehicle speed and opposing vehicle characteristics, with buses having the highest crash probability (*11*).

Elassad et al. (2020) developed real-time crash prediction models using driving simulator data that captured driver input responses, vehicle kinematics, and weather conditions. They employed machine learning techniques, demonstrating high prediction accuracy, especially under varying weather conditions (*12*). Moreover, Guo et al. (2021) developed a traffic crash risk prediction model based on risky driving behavior and traffic flow data. Combining detailed driver behavior data with machine learning techniques, they accurately predicted 84.48% of crashes, highlighting the potential of real-time behavior monitoring to improve road safety (*13*).

Integrative LSTM Techniques for Mobility Prediction

The integration of Long Short-Term Memory (LSTM) networks in mobility prediction has significantly advanced predictive modeling and data analysis capabilities. Mou et al. (2019) developed the Temporal Information Enhancing LSTM (T-LSTM), which integrates recurrent time labels to capture temporal dynamics more effectively, enhancing the accuracy of short-term traffic flow predictions (*14*).

Altché and de La Fortelle (2017) explored the application of LSTM networks for highway trajectory prediction, focusing on predicting driving behaviors in a macroscopic context (*15*).

In public transportation, LSTM networks have been employed to predict passenger flow and enhance service planning. Pasini et al. (2020) used an LSTM encoder-predictor model for short-term train load forecasting, demonstrating its effectiveness in improving efficiency (*16*). Additionally, LSTM networks have been utilized for predicting road traffic crashes by analyzing historical accident data and various influencing factors. Li et al. (2019) developed an LSTM-CNN-based model to predict traffic crashes, incorporating features such as traffic flow characteristics, signal timing, and weather conditions, which proved effective in identifying patterns leading to crashes and enabling proactive safety measures (*17*).

Application of LSTM in Microscopic Driving Behavior Analysis

LSTM networks have also proven to be highly effective in analyzing time-series data, particularly in microscopic driving behavior analysis. Saleh et al. (2017) used smartphone sensor data with LSTM to classify driving behaviors, achieving high accuracy in categorizing normal, aggressive, and drowsy driving behaviors (*18*). Deo and Trivedi (2018) employed LSTMs to model and predict lane change intentions, showing that these models can accurately predict unsafe lane changes by analyzing sequences of vehicle movements (*19*). Jia et al. (2021) further used LSTM models to predict lane-changing behavior, effectively identifying patterns leading up to crash events (*20*).

Moreover, LSTM networks have been used to model interactions between vehicles and pedestrians. Ridel et al. (2019) implemented an LSTM model to predict pedestrian trajectories, which is crucial for the development of advanced driver-assistance systems (ADAS) and autonomous driving technologies (*21*). Tselentis and Papadimitriou (2023) analyzed driving behavior through detailed time-series data, focusing on speed and heart rate measurements to recognize patterns preceding risky driving scenarios (*22*).

In conclusion, LSTM neural networks have emerged as powerful tools in the microscopic analysis of driving behavior. Their ability to model and predict sequential data makes them invaluable for improving the accuracy of driving behavior models, enhancing risk prediction, and ultimately contributing to safer and more efficient transportation systems. The diverse applications of LSTM in the field of road safety are strongly represented in the existing literature, from trajectory prediction to modeling complex interactions in driving environments.

METHODS

Theoretical Background of LSTM

A long short-term memory (LSTM) network, a variant of recurrent neural networks (RNNs), has become essential for capturing and modeling intricate sequential patterns. LSTM networks address the vanishing gradient problem that often impairs traditional RNNs, enhancing performance in tasks involving long-term dependencies. Their ability to retain information over extended sequences makes them well-suited for modeling sequential data.

LSTM networks are structured like a chain of repeating modules, each containing multiple gates that regulate information flow, enabling the system to effectively learn and retain dependencies in sequential data over long periods. Due to their advanced architecture, capable of capturing and retaining long-term dependencies, LSTMs have been successfully applied to a wide array of tasks, including activity recognition and language translation. Their precision in processing time-series data makes them particularly suitable for applications where understanding temporal dynamics is crucial.

A typical LSTM architecture includes three primary gates:

1. **Input Gate:** Regulates information included in the memory cell. It has two components: an input gate layer using a sigmoid function to select values to be updated and a hyperbolic tangent (tanh) layer creating a vector of candidate values for potential inclusion in the cell state.
2. **Forget Gate:** Determines which information from the cell state should be discarded using a sigmoid function where 1 means retention and 0 means forgetting.
3. **Output Gate:** Controls information output from the cell state, employing a sigmoid activation function to filter relevant parts of the cell state at the current time step. This filtered state is then scaled using a tanh function, optimizing values for the next layer or output.

Methodological approach for the estimation of microscopic event probability

LSTM models were chosen for their proficiency in handling sequential data from driving metrics, capturing temporal dependencies critical for predicting driving risk events. Both uni-directional and bi-directional LSTM configurations were tested. The uni-directional LSTM processes data in a single temporal direction, capturing sequential dependencies and trends in road user behavior. In contrast, the bi-directional LSTM processes data in both forward and backward temporal directions, allowing it to consider the full context of traffic behavior both before and after each moment in time. This dual perspective can potentially enhance the model's understanding of complex driving patterns, thereby improving the accuracy of risk prediction.

The LSTM was trained to estimate the microscopic event probability of the ego vehicle. Driving metrics considered as features of the LSTM model include metrics for the ego vehicle and surrounding vehicles: vehicle type, speed, longitudinal acceleration, distance traveled, relative distance, speed, and time to collision from the leading vehicle. These metrics were also considered for the four vehicles close to the ego vehicle: the vehicle in front, the vehicle following, and the two vehicles on the right and left. The target variables used to train the model were lane changing, speeding, harsh acceleration, and harsh braking.

All LSTM models in this study were structured with at least two LSTM layers using the 'tanh' activation function and incorporating L2 regularization to mitigate overfitting. The models included dropout layers with a dropout rate of over 40% to further prevent overfitting. The output layer used a dense layer with a 'sigmoid' activation function, suitable for binary classification tasks. The models were trained with a batch size of 32 for up to 400 epochs, with early stopping when validation loss did not improve. Different model configurations were tested by varying the number of LSTM layers, their direction (unidirectional or bidirectional), the number of neurons, dropout rate, and optimizer type, such as Adam and SGD with different learning rates and momentum settings.

Figure 1 illustrates the driving metrics of the five vehicles considered in the LSTM model training. The ego vehicle is shown in green, while surrounding vehicles are in black. Figure 1 also shows how spatiotemporal risk is estimated on a road segment level through the aggregation of the microscopic risk estimated using the LSTM model. This approach considers both space and time dimensions, providing a time evolution of road risk in a specific cross-section and a risk snapshot for a specific road segment.

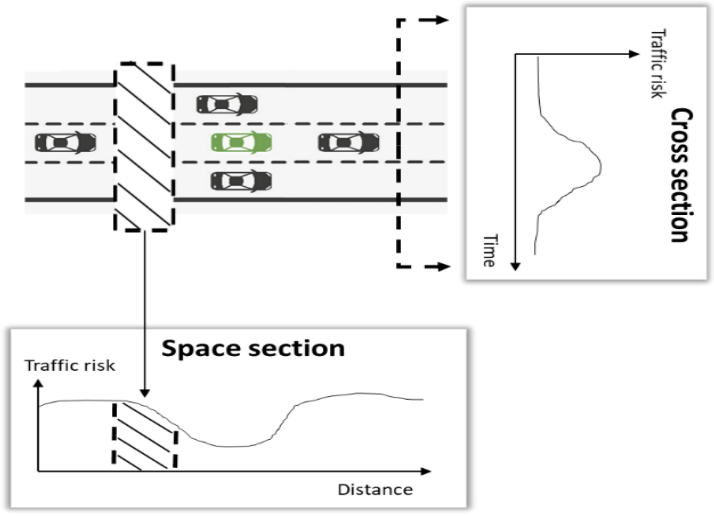

Figure 1 The ego vehicle (green color) and the surrounding vehicles whose metrics are considered in the LSTM model trained and tested herein. (b) Right figure illustrates how risk can be aggregated in space and time to assess spatiotemporal risk of a road

The length of the time-series before the event used to predict the risky event is 10 seconds. The event flag marks the occurrence of an event within a rolling window of 2 seconds. There is a 1-second gap between the end of the time-series features and the event flag, considered as the driver's reaction time. As shown in **Figure 2**, the LSTM model is then trained to predict the probability of the ego vehicle participating in a risky event within the future window of 1 to 3 seconds ahead, skipping the future window of 0 to 1 second ahead of the prediction.

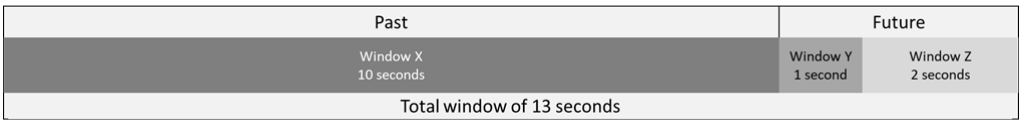

Figure 2 Past time-series window X where driving metrics are recorded and future time-series window Z where an event may occur. Window Y is not considered

Where in **Figure 2**:
X: 10-second window of driving metrics.
Y: 1-second window before the event occurrence, not considered in prediction.
Z: 2-second window where event occurrence is flagged and used as a target variable.

The LSTM model is trained to predict the probability of the ego vehicle participating in an event within window Z, using driving data from the five vehicles within window X. The lengths of the time-series windows were chosen after trial and error but require further exploration to maximize model predictability.

The input LSTM array has a shape of (250, 35, 69505), with X being 250 past observations, Y being 35 features, and Z being 69505 time-series snapshots. The target time-series shape is (1, 69505), representing whether an event occurred within the 2-second window. The train and test sets were split 80% and 20%, respectively. The study aims to provide short-term predictions of the probability of a vehicle being involved in risky driving events based on the characteristics of the vehicle and surrounding vehicles. The rolling basis calculation of time-series windows creates input and target arrays for LSTM model training.

DATA

Data Description

The pNEUMA dataset (*23, 24*) was developed from a data collection experiment conducted in Athens, Greece, in 2018. Over five days, a swarm of ten drones captured real-time traffic data for three hours during peak morning hours. As shown in **Figure 3**, the drones covered a 1.3 km² area with over 100 kilometers of roadways and nearly 100 intersections, capturing nearly half a million vehicle trajectories. This data included vehicle speed, acceleration, and position with a data collection frequency of 25Hz, (e.g., captured every 0.04 seconds).

Figure 3 Blocks covered by each drone of the swarm

This study focused on Panepistimiou Street, a five-lane urban arterial in Athens **(Figure 4)**, due to its volume of traffic and frequency of traffic events. The studied blocks were IDs: 2, 3, and 5.

Figure 4 Study area: Panepistimiou Street

The pNEUMA dataset organized each row to detail a single vehicle's data. Initial columns provide the vehicle's trajectory, including trackID, vehicle type, distance traveled, and average speed. Following columns repeat data for different timeframes, including latitude, longitude, speed, longitudinal acceleration, lateral acceleration, and time.

Data Pre-Processing

The dataset was restructured to ensure that each column of time data corresponded to 0.04 seconds intervals. Furthermore, Panepistimiou lanes were delineated as polygons using Google Earth coordinates for accurate vehicle position tracking. Based on the pNEUMA dataset, three data frames were developed: (a) individual vehicle metrics, (b) lane-specific traffic metrics, and (c) traffic events.

The first dataframe aggregated data for vehicle pairs (following and leading vehicles) and risk metrics, including track ID, position, type, speed, longitudinal acceleration, and lane-polygon. The Time-to-Collision (TTC) metric (*25*) was also calculated **(Equation 1)**.

$$TTCi = \frac{X_{i-1}(t) - X_i(t) - l_i}{\dot{X}_i(t) - \dot{X}_{i-1}(t)} \quad \forall \dot{X}_i(t) > \dot{X}_{(i-1)}(t), \qquad (1)$$

The second dataframe captured traffic metrics for each polygon at every 0.04-second timeframe, including (i) average vehicle speed **(Equation 2)**, (ii) total number of vehicles, (iii) density (vehicles/km) **(Equation 3)**, and (iv) traffic flow (vehicles per hour) **(Equation 4)**.

$$v_{avg} = \frac{\sum_{i=1}^{n} v_i}{n}, \qquad (2)$$

where vavg represents the average speed (km/h), v_i speed of the i-th vehicle (km/h) and n the total number of vehicles observed.

$$k = \frac{n}{L}, \qquad (3)$$

where n represents the total number of vehicles observed and L the length of the road segment (km).

$$q = k \cdot v_{avg}, \qquad (4)$$

where k represents the density (vehicles/km) and vavg the average speed (km/h).

The third dataframe identified risky traffic events, setting thresholds for safety indicators such as TTC, harsh acceleration, harsh braking, and speeding. The TTC threshold was set at 1.5 seconds (*26*), harsh braking (*27*) and acceleration at 0.5g (4.9 m/s²) (*28*), and speeding incidents at 10km/h above the speed limit. Lane change detection flagged vehicles in different polygons between two consecutive time frames.

In the final phase, the three dataframes were integrated into a unified dataset, aligning each entry with 0.04-second intervals, creating a cohesive time-series format essential for advanced predictive model analysis.

RESULTS

Microscopic driving risk probability

Table 1 outlines various LSTM model configurations tested for microscopic risk prediction expressed through speeding. Each configuration is detailed with respect to the direction of the LSTM (uni-directional or bi-directional), the number of LSTM layers, dropout percentage, the number of neurons in each LSTM layer, type of optimizer, learning rate, batch size, and the number of epochs used for training. Specifically, two uni-directional models were tested with different layer and neuron configurations, all optimized using the Adam optimizer with a learning rate of 0.0001 and trained for 50 epochs. The majority of the configurations are bi-directional LSTMs, varying in the number of layers (2 or 3), dropout rates (40% or 50%), and neuron counts (ranging from 16 to 128 in various layers). These bi-directional models employ either the Adam or SGD optimizer, with the SGD variants using momentum, decay, and in some cases, Nesterov acceleration. Learning rates for SGD models vary between 0.001 and 0.0001, with training epochs extending up to 400 for certain configurations. This diverse set of configurations aims to identify the optimal architecture and training parameters for accurately predicting driving risk at a microscopic level. Other configurations were also tested, for

instance uni-directional models and models with lower dropout percentages, which demonstrated a very poor performance. Models trained for the prediction of lane changing, harsh acceleration, and harsh braking showed poor performance as well, which is mainly attributed to the low number of events recorded during the drone videos processed and this is the reason why the results of these models are not presented in this section.

TABLE 1 Configuration of the LSTM tested for the prediction of speeding events

Model ID	Direction	# LSTM layers	Dropout %	# Neurons in 1st LSTM layer	# Neurons in 2nd LSTM layer	# Neurons in 3rd LSTM layer	Optimizer	Learning rate	Batch size	# epochs
1	Uni	3	40%	64	32	32	Adam	1E-4	64	50
2	Uni	2	50%	32	16	-	Adam	1E-4	64	50
3	Bi	2	50%	32	16	-	Adam	1E-4	64	50
4	Bi	2	50%	64	32	-	Adam	1E-4	64	100
5	Bi	3	50%	32	16	16	Adam	1E-4	64	50
6	Bi	3	50%	64	32	32	Adam	1E-4	64	50
7	Bi	3	40%	128	64	64	Adam	5E-4	64	50
8	Bi	3	40%	64	32	32	SGD (momentum=0.4, decay=0.0001, nesterov=False)	1E-3	64	60
9	Bi	3	40%	32	16	16	SGD (momentum=0.6, decay=0.00001, nesterov=False)	1E-4	64	60
10	Bi	2	40%	64	32	-	SGD (momentum=0.6, decay=0.00001, nesterov=True)	1E-4	64	60
11	Bi	2	40%	64	32	-	SGD (momentum=0.5, decay=0.0005, nesterov=True)	1E-3	64	60
12	Bi	3	40%	128	64	64	SGD (momentum=0.5, decay=0.0005, nesterov=True)	1E-4	256	200
13	Bi	3	40%	32	16	16	SGD (momentum=0.5, decay=0.0005, nesterov=True)	1E-4	32	400

The performance of the LSTM models was evaluated using the key metrics of accuracy, precision, recall, and the Area Under the Curve (AUC). These performance metrics provide a comprehensive assessment of the model's ability to accurately predict driving risk events, ensuring a robust evaluation of the model's effectiveness in various scenarios. This systematic evaluation allows for the identification of the most suitable LSTM configuration and temporal parameters for real-time driving risk assessment. **Table 2** presents the performance metrics of various uni-directional and bi-directional

LSTM models trained to predict speeding events. The bi-directional models generally outperformed the uni-directional models in terms of precision, likely due to their ability to consider future states.

For example, Model 7, a bi-directional LSTM with three layers and 40% dropout, achieved the highest precision of 94%. However, this model, like others, exhibited a relatively low recall rate of 18%, indicating that while the model is effective at correctly identifying speeding events when they occur (high precision), it struggles to detect all occurrences of such events (moderate recall). In contrast, Model 1, a uni-directional LSTM with three layers and 40% dropout, showed a slightly lower precision of 84% and a slightly higher recall of 22%. The performance metrics of the bi-directional LSTM models for predicting speeding events reveal several important insights.

Models using the Adam optimizer consistently demonstrated higher precision and recall compared to those using the SGD optimizer. For instance, Model 6, a bi-directional LSTM with three layers and an Adam optimizer, achieved a recall rate of 27% and a precision of 86%, reflecting a relatively better balance between identifying and correctly predicting a reasonable number of speeding events. This model also demonstrated the highest binary accuracy of 86% among all LSTM models. This suggests that increasing the number of neurons in each layer, up to a certain point, can enhance the model's ability to learn from the data without overfitting. On the other hand, models with higher complexity, such as Model 12 with three layers and 128 neurons in the first LSTM layer, showed significant increases in training time without corresponding improvements in prediction accuracy or recall. This indicates a point of diminishing returns in model complexity, where additional layers and neurons do not necessarily enhance performance but rather contribute to overfitting, as seen in models with high precision but low recall.

TABLE 2 Performance metrics for the LSTM models predicting speeding events

Model ID	Binary accuracy	Precision	Recall	Area Under the Curve (AUC)
1	84%	76%	22%	75%
2	83%	87%	10%	71%
3	84%	89%	17%	75%
4	85%	85%	25%	77%
5	84%	79%	18%	74%
6	86%	86%	27%	77%
7	85%	94%	18%	76%
8	82%	52%	27%	69%
9	81%	37%	4%	64%
10	81%	100%	0%	66%
11	82%	76%	4%	68%
12	82%	72%	2%	65%
13	82%	68%	3%	66%

Models using the SGD optimizer showed varied performance. Model 10, which uses a bi-directional LSTM with two layers and 64 neurons in the first layer, achieved a perfect precision of 100% but an extremely low recall of 0.1%. This indicates a severe case of overfitting where the model correctly identifies very few instances, but those it identifies are almost always correct. The lower AUC values in SGD-optimized models, such as 64% for Model 9 and 65% for Model 12, further suggest that these models are less reliable in distinguishing between positive and negative classes across different threshold settings compared to their Adam-optimized counterparts. This highlights the need for more sophisticated regularization techniques or the exploration of other optimizers to improve the recall rates and overall

robustness of these models. Model 7 is picked for the macroscopic risk estimation in the following section, as the model showing the highest accuracy.

Macroscopic driving risk probability

In order to provide a comprehensive assessment of driving risk at a road section level, the microscopic driving risk probabilities predicted by the LSTM models were aggregated. This aggregation process involved collating the risk probabilities of individual vehicles and their interactions within specific road segments over time. This aggregation was made for a short time period of 102 seconds. By doing so, the detailed microscopic predictions were transformed into a broader macroscopic perspective, offering valuable insights into the overall risk profile of different road sections. As explained in the previous section, Model 7 was used to initially estimate microscopic risk expressed as speeding events occurring at the vehicle level and subsequently aggregated this risk at the road section level. This estimation should be extended in the future to account for various events and situations that express traffic risk, such as lane changing, harsh acceleration, and harsh braking, and should also represent the traffic risk of all individual road users within the section.

To achieve this aggregation, the road was divided into discrete segments, and the predicted risk probabilities for each vehicle within these segments were collected over the specified time intervals. These probabilities were then summed to generate a risk score for each segment, reflecting the overall likelihood of risky events occurring within that area. The resulting aggregated risk data, normalized based on the maximum risk probability observed, are visualized in **Figure 5**, highlighting the variation in driving risk for a specific road section across time. This visualization shows the dynamic nature of traffic risk when focusing on a road section of the network. Future research should enrich this analysis by studying multiple road sections of the road network at the same time to assess relative risk within the network. This would not only underscore the areas with higher risk but would also provide a dynamic view of how risk evolves over time and space, enabling more targeted, prioritized and effective traffic safety interventions.

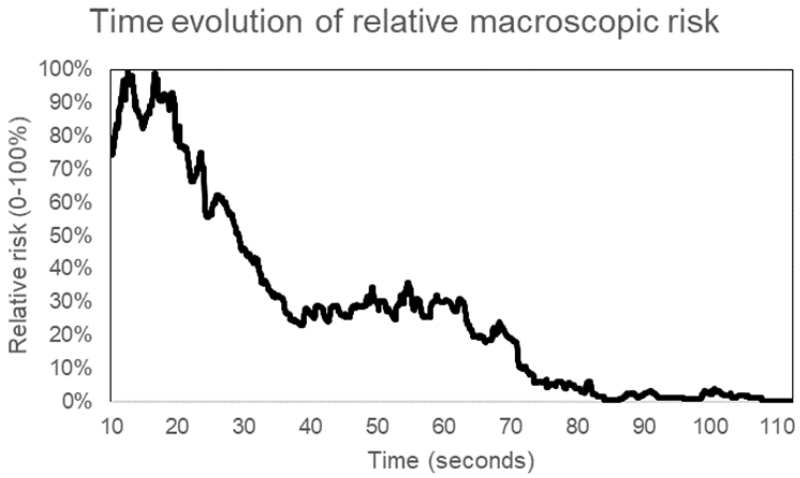

Figure 5 Evolution of aggregated macroscopic risk in time for a specific section of Panepistimiou street for 100 seconds

DISCUSSION

The performance of the uni-directional and bi-directional LSTM models was evaluated using several key metrics, revealing notable differences between the two configurations. The bi-directional LSTM models demonstrated slightly better performance compared to their uni-directional counterparts, likely due to their capacity to consider future states. This ability allows bi-directional models to capture

more context around the event, leading to a more accurate prediction of risky driving behaviors. Both model types exhibited tendencies to overfit, as evidenced by high precision and recall during testing phases. However, the implementation of neuron dropout effectively mitigated rapid overfitting, ensuring that the models maintained generalizability to new data. The high validation precision of the LSTM models indicates their capability to correctly distinguish between risky and non-risky event situations, demonstrating their potential utility in real-world driving risk assessments.

Despite the high precision, the models showed a relatively low recall on the validation dataset, highlighting a significant limitation and a need for further refinement to improve the identification of actual risky events. Overall, while the bi-directional LSTM models show promise due to their slightly better precision, both types of LSTM models require further refinement to improve recall rates. The low recall suggests that while the models are adept at identifying true non-risky events, they struggle to detect a substantial portion of actual risky events. This inefficiency could reduce the model's practical effectiveness in real-time applications where the identification of most risky events is critical for preventing incidents.

The findings suggest that while LSTM models are a valuable tool for driving risk prediction, careful attention must be paid to model design and evaluation metrics to ensure they can be effectively deployed in real-world scenarios. Expanding the methodology to different urban settings will help generalize the findings and validate the model's robustness. Integrating drone data with other sources, such as vehicle, road, and environmental data, could provide a more comprehensive risk assessment and enhance the model's predictive capabilities.

Increasing the complexity of the models, by adding more neurons and layers, did not yield a corresponding improvement in predictability. Instead, these more complex models resulted in significantly longer training times without substantial gains in performance. This indicates that there is a limit to the benefits of increasing model complexity for this specific application and emphasizes the importance of balancing model complexity with practical considerations such as training time and computational resources. Overall, while the bi-directional LSTM models show promise due to their slightly better precision, both types of LSTM models require further refinement to improve recall rates. Future work could explore alternative architectures or additional features that might enhance the model’s ability to identify risky events more accurately. Additionally, strategies to address overfitting, such as more sophisticated regularization techniques or data augmentation, could be beneficial. The findings suggest that while LSTM models are a valuable tool for driving risk prediction, careful attention must be paid to model design and evaluation metrics to ensure they can be effectively deployed in real-world scenarios.

In the future, these models should be extended to account for various road users, including pedestrians, cyclists, e-scooters, and other vulnerable road users, and incorporate their specific metrics such as trajectories, distraction, interactions, and maneuvers. Additionally, factors related to the road itself (e.g., road geometry, infrastructure, road defects, traffic lights phasing), vehicles (e.g., age, type, engine capacity, fleet characteristics), traffic conditions (e.g., volume, speed, density, travel time), and the environment (e.g., weather, time of day, daylight) should be considered. Integrating crash data, including real-time crashes, high-risk collisions, and historical crash records, will provide a more comprehensive risk assessment. By studying these elements both individually and in combination, future research can better understand how the interaction of all road users affects traffic risk and identify which factors contribute most significantly to increased traffic risk. This holistic approach will be essential in developing more effective and inclusive traffic safety measures.

Comparison with past literature is not possible because, to the best of our knowledge, no previous studies have employed a similar methodology for driving risk prediction using drone-collected data and LSTM models. This novel approach sets a precedent in the field, highlighting the innovative nature of our research. Consequently, our findings represent a pioneering step in driving risk assessment. While traditional studies rely on static sensors or vehicular data, our drone-based approach offers a broader and more dynamic perspective, capturing intricate traffic patterns and behaviors. This method allows for comprehensive monitoring of multiple vehicles and their interactions in real-time, providing a richer

dataset for analysis. The versatility and mobility of drones enable the collection of high-resolution traffic data over extensive urban areas, enhancing the accuracy and applicability of the driving risk predictions.

CONCLUSIONS

This research underscores the potential of using AI modeling to predict driving risk in real-time and ultimately enhance urban traffic safety. The LSTM-based approach proved effective in terms of precisely predicting risky driving events in both microscopic and macroscopic level, demonstrating its utility in real-time risk monitoring systems. By leveraging drone-collected traffic data, this study provides a comprehensive methodological framework for estimating traffic risk, contributing to safer driving environments. One of the main contributions of this paper is the introduction of a methodological approach for the estimation of microscopic and macroscopic traffic risk. The comparison of uni-directional and bi-directional LSTM models revealed that bi-directional models offer slightly better precision, likely due to their ability to consider future states. Despite some overfitting tendencies, the models demonstrated high validation precision, indicating their capacity to accurately distinguish between risky and non-risky events.

Future work could explore alternative architectures or additional features that might enhance the model's ability to identify risky events more accurately. Additionally, strategies to address overfitting, such as more sophisticated regularization techniques or data augmentation, could be beneficial. Findings suggest that LSTM models can be potentially valuable in driving risk prediction but need further refinement to ensure effective deployment in real-world scenarios. Moreover, future research should focus on improving the accuracy of driving event predictions by exploring different models and optimizing the lengths of the time-series windows used in model training. Training the models into additional urban settings will improve robustness and assist generalization of results. It would also be highly recommended to integrate the drone data with other sources related to traffic, road, infrastructure, and environment, as well as from other road users such as AVs, pedestrians, cyclists and scooters. Collision data would also be of significant value to improve model performance. This could provide a more comprehensive risk assessment and enhance the model's predictive capabilities.

The methodology and results of this study have significant potential for real-time applications aimed at enhancing traffic safety. The aggregated risk data can be integrated into real-time monitoring systems to provide immediate feedback to drivers, thereby potentially reducing traffic incidents. By aggregating results on a road section level, this approach enhances the practical application of the data by offering real-time detection of risk profiles for different road segments as well as real-time detection of risk hotspots in the road network. This real-time risk estimation can be used to develop and test driver feedback systems that alert drivers to potential dangers, enabling them to take preventive actions promptly. Additionally, a network-wide system could be developed to monitor and assess the real-time evolution of driving risk across an entire urban area, providing valuable insights for traffic management and control strategies. Such applications could lead to more proactive traffic safety measures, ultimately contributing to safer road environments.

In summary, this study represents a pioneering step in driving risk assessment using AI and drone data. The findings highlight the potential of this approach to contribute significantly to real-time traffic safety measures, paving the way for more proactive and informed traffic management strategies in urban environments.

ACKNOWLEDGMENTS

The authors would like to acknowledge the use of the pNEUMA dataset in this study– open-traffic.epfl.ch

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design D. I. Tselentis, G. Yannis; data collection: D. I. Tselentis, T. Garefalakis, D. Nikolaou; analysis and interpretation of results:

D. I. Tselentis, T. Garefalakis; draft manuscript preparation: D. I. Tselentis, T. Garefalakis, D. Nikolaou, G. Yannis. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. World Health Organization. *Global Status Report on Road Safety 2023*. Geneva, 2023. https://www.who.int/publications/i/item/9789240086517. Accessed Mar. 6, 2024.

2. Staubach, M. Factors Correlated with Traffic Accidents as a Basis for Evaluating Advanced Driver Assistance Systems. *Accident Analysis & Prevention*, Vol. 41, No. 5, 2009, pp. 1025–1033. https://doi.org/10.1016/j.aap.2009.06.014.

3. Treiber, M., and A. Kesting. *Traffic Flow Dynamics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

4. Cascan, E. T., J. Ivanchev, D. Eckhoff, A. Sangiovanni-Vincentelli, and A. Knoll. Multi-Objective Calibration of Microscopic Traffic Simulation for Highway Traffic Safety. 2019.

5. Zhao, X., Q. Li, D. Xie, J. Bi, R. Lu, and C. Li. Risk Perception and the Warning Strategy Based on Microscopic Driving State. *Accident Analysis & Prevention*, Vol. 118, 2018, pp. 154–165. https://doi.org/10.1016/j.aap.2018.02.012.

6. Zatmeh-Kanj, S., and T. Toledo. Car Following and Microscopic Traffic Simulation Under Distracted Driving. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2675, No. 8, 2021, pp. 643–656. https://doi.org/10.1177/03611981211000357.

7. Tan, H., G. Lu, and M. Liu. Risk Field Model of Driving and Its Application in Modeling Car-Following Behavior. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 23, No. 8, 2022, pp. 11605–11620. https://doi.org/10.1109/TITS.2021.3105518.

8. Chen, Q., H. Huang, Y. Li, J. Lee, K. Long, R. Gu, and X. Zhai. Modeling Accident Risks in Different Lane-Changing Behavioral Patterns. *Analytic Methods in Accident Research*, Vol. 30, 2021, p. 100159. https://doi.org/10.1016/j.amar.2021.100159.

9. Ali, Y., Z. Zheng, Md. Mazharul Haque, M. Yildirimoglu, and S. Washington. Understanding the Discretionary Lane-Changing Behaviour in the Connected Environment. *Accident Analysis & Prevention*, Vol. 137, 2020, p. 105463. https://doi.org/10.1016/j.aap.2020.105463.

10. Li, X., O. Oviedo-Trespalacios, and A. Rakotonirainy. Drivers' Gap Acceptance Behaviours at Intersections: A Driving Simulator Study to Understand the Impact of Mobile Phone Visual-Manual Interactions. *Accident Analysis & Prevention*, Vol. 138, 2020, p. 105486. https://doi.org/10.1016/j.aap.2020.105486.

11. Haque, N., Md. Hadiuzzaman, F. Rahman, and M. R. K. Siam. Real-Time Motion Trajectory Based Head-on Crash Probability Estimation on Two-Lane Undivided Highway. *Journal of Transportation Safety & Security*, Vol. 12, No. 10, 2020, pp. 1312–1337. https://doi.org/10.1080/19439962.2019.1597001.

12. Elamrani Abou Elassad, Z., H. Mousannif, and H. Al Moatassime. Class-Imbalanced Crash Prediction Based on Real-Time Traffic and Weather Data: A Driving Simulator Study. *Traffic Injury Prevention*, Vol. 21, No. 3, 2020, pp. 201–208. https://doi.org/10.1080/15389588.2020.1723794.

13. Guo, M., X. Zhao, Y. Yao, P. Yan, Y. Su, C. Bi, and D. Wu. A Study of Freeway Crash Risk Prediction and Interpretation Based on Risky Driving Behavior and Traffic Flow Data. *Accident Analysis & Prevention*, Vol. 160, 2021, p. 106328. https://doi.org/10.1016/j.aap.2021.106328.

14. Mou, L., P. Zhao, H. Xie, and Y. Chen. T-LSTM: A Long Short-Term Memory Neural Network Enhanced by Temporal Information for Traffic Flow Prediction. *IEEE Access*, Vol. 7, 2019, pp. 98053–98060. https://doi.org/10.1109/ACCESS.2019.2929692.

15. Altche, F., and A. de La Fortelle. An LSTM Network for Highway Trajectory Prediction. 2017.

16. Pasini, K., M. Khouadjia, A. Samé, F. Ganansia, and L. Oukhellou. LSTM Encoder-Predictor for Short-Term Train Load Forecasting. 2020.

17. Li, P., M. Abdel-Aty, and J. Yuan. Real-Time Crash Risk Prediction on Arterials Based on LSTM-CNN. *Accident Analysis & Prevention*, Vol. 135, 2020, p. 105371. https://doi.org/10.1016/j.aap.2019.105371.

18. Saleh, K., M. Hossny, and S. Nahavandi. Driving Behavior Classification Based on Sensor Data Fusion Using LSTM Recurrent Neural Networks. 2017.

19. Deo, N., and M. M. Trivedi. Multi-Modal Trajectory Prediction of Surrounding Vehicles with Maneuver Based LSTMs. 2018.

20. Jia, S., F. Hui, C. Wei, X. Zhao, and J. Liu. Lane-Changing Behavior Prediction Based on Game Theory and Deep Learning. *Journal of Advanced Transportation*, Vol. 2021, 2021, pp. 1–12. https://doi.org/10.1155/2021/6634960.

21. Ridel, D. A., N. Deo, D. Wolf, and M. Trivedi. Understanding Pedestrian-Vehicle Interactions with Vehicle Mounted Vision: An LSTM Model and Empirical Analysis. 2019.

22. Tselentis, D. I., and E. Papadimitriou. Time-Series Clustering for Pattern Recognition of Speed and Heart Rate While Driving: A Magnifying Lens on the Seconds around Harsh Events. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 98, 2023, pp. 254–268. https://doi.org/10.1016/j.trf.2023.09.010.

23. Barmpounakis, E., and N. Geroliminis. On the New Era of Urban Traffic Monitoring with Massive Drone Data: The PNEUMA Large-Scale Field Experiment. *Transportation Research Part C: Emerging Technologies*, Vol. 111, 2020, pp. 50–71. https://doi.org/10.1016/j.trc.2019.11.023.

24. Barmpounakis, E., and N. Geroliminis. PNEUMA Dataset [Data Set]. On the New Era of Urban Traffic Monitoring with Massive Drone Data: The PNEUMA Large-Scale Field Experiment. *Transportation Research Part C: Emerging Technologies*, Vol. 111, 2024, pp. 50–71. https://doi.org/10.5281/zenodo.10491409.

25. Minderhoud, M. M., and P. H. L. Bovy. Extended Time-to-Collision Measures for Road Traffic Safety Assessment. *Accident Analysis & Prevention*, Vol. 33, No. 1, 2001, pp. 89–97. https://doi.org/10.1016/S0001-4575(00)00019-1.

26. Oikonomou, M. G., A. Ziakopoulos, A. Chaudhry, P. Thomas, and G. Yannis. From Conflicts to Crashes: Simulating Macroscopic Connected and Automated Driving Vehicle Safety. *Accident Analysis & Prevention*, Vol. 187, 2023, p. 107087. https://doi.org/10.1016/j.aap.2023.107087.

27. Kamla, J., T. Parry, and A. Dawson. Analysing Truck Harsh Braking Incidents to Study Roundabout Accident Risk. *Accident Analysis & Prevention*, Vol. 122, 2019, pp. 365–377. https://doi.org/10.1016/j.aap.2018.04.031.

28. Ali, G., S. McLaughlin, and M. Ahmadian. Quantifying the Effect of Roadway, Driver, Vehicle, and Location Characteristics on the Frequency of Longitudinal and Lateral Accelerations. *Accident Analysis & Prevention*, Vol. 161, 2021, p. 106356. https://doi.org/10.1016/j.aap.2021.106356.