Advances in Artificial Intelligence and Machine Learning; Research 5 (2) 3975-3987 Received 29-03-2025; Accepted 21-06-2025; Published 28-06-2025

Traffic Accident Predictive Model for Efficient Resource Allocation in Qatar: A Novel Transformer-Based Approach

Mansoor G. Al-Thani malthani@moi.gov.qa
Technical Affairs Department, Ministry of Interior
Doha-Qatar

Corresponding Author: Mansoor G. Al-Thani

Abstract

Qatar's rapid urbanization and population growth have led to a significant increase in vehicle ownership and traffic accidents, creating challenges for public safety, emergency response, and urban planning. This study proposes TrafficTransformer, a deep learning model designed to predict traffic accident occurrences across 98 zones using traffic data collected from January 2017 to October 2023. The dataset, sourced from police traffic reports, includes hourly accident logs with features such as time of day, weather conditions, and spatial area codes. TrafficTransformer leverages self-attention and multi-head attention mechanisms to capture dynamic spatiotemporal patterns more effectively than traditional CNN and LSTM models. Model performance was evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE), based on a 75%-25% training-testing split. Results demonstrate substantial improvements over baseline models. This research provides actionable insights for intelligent traffic management, targeted interventions, and resource optimization in emergency response systems.

Keywords: Traffic accidents, Predictive modeling, Deep learning, Transformer, Resource allocation; Road safety, AI applications

1. INTRODUCTION

Due to the fast urbanization of modern civilizations, the number of cars on the road has increased exponentially. Urban regions now experience higher levels of pollution, traffic congestion, and traffic accidents [1, 2]. The World Health Organisation (WHO) estimated that 1.35 million people die in traffic accidents annually, while up to an additional 50 million get non-fatal injuries [3].

In Qatar, due to fast economic expansion, a significant rise in car ownership has caused traffic jams and an increase in traffic accidents. In the last ten years, the number of registered cars in Qatar has more than doubled, with the largest growth occurring in metropolitan regions. Traffic accidents have become a serious public safety issue that has a big influence on both social cohesion and economic stability. Hence, the need for an efficient traffic accident prediction model is of paramount importance for better resource allocation and for saving lives and welfare in the country.

Citation: Mansoor G. Al-Thani. Traffic Accident Predictive Model for Efficient Resource Allocation in Qatar: A Novel Transformer-Based Approach. Advances in Artificial Intelligence and Machine Learning. 2025;5(2):224.

With the help of huge traffic data, people are now able to make prompt decisions to avoid traffic by choosing less congested routes and roads [4, 5]. Drivers have been informed of such directions through the use of digital signs that are updated regularly with each road's condition. Similarly, a practical way to forecast the likelihood of traffic accidents is offered by current advancements in big data and deep learning. Predicting and informing drivers in the vicinity of such information will warn them to avoid more dangerous roadways. That being said, predicting traffic accidents accurately is a highly difficult task because it depends on several factors, such as the variations in traffic accident rates between different regions and hazardous weather conditions.

To gather different kinds of traffic data, numerous sensor designs have been created during the last fifty years. Typically, trip duration, speed, and traffic density are among the traffic statistics that are gathered for research purposes. Some of the technologies created to record traffic status data include loop detectors, radar, video detection systems, Bluetooth sensors, and wireless sensors, among others [6].

Another important aspect in this field is the ability to identify traffic accident triggers that would cause or lead to a possible traffic accident. The literature has reported on a vast number of articles that address this issue [4, 7, 8]. Such factors include extreme weather, particular sports events, peak times, and temperatures.

In order to combine the data with the triggers to develop and implement an efficient traffic accident prediction model, several algorithms have been reported. Deep learning models have become increasingly popular for traffic accident forecasting [9, 10]. Marcillo et al. recently reviewed the literature and determined which accident prediction methods were most frequently used [11]. They reported that 30% of the cases employed neural network versions, 15% employed the Support Vector Machine technique, and 12% employed Bayesian network prediction models, according to their findings [11].

Recurrent Neural Networks (RNNs) are also frequently used as a core component in traffic prediction models, especially for estimating traffic speed, time, and flow [12]. Moreover, models based on the Seq2Seq architecture have been employed for forecasting traffic state sequences [13]. Numerous predictive models also focus on identifying connections within traffic networks. In graph-based data, Graph Convolutional Networks (GCNs) are normally used to analyze spatial correlations [14]. However, Convolutional Neural Networks (CNNs) are also excellent at extracting spatial characteristics from data in Euclidean space [15]. However, these methods frequently focus on either spatial or temporal aspects and find it challenging to characterize dynamic spatiotemporal interactions simultaneously.

Advances in deep learning have significantly improved the ability to uncover complex patterns in high-dimensional data, particularly in areas like traffic flow analysis, intelligent transportation systems, and accident risk prediction. For instance, Park et al. created a traffic prediction system using logistic regression and k-means clustering based on data from highway accidents in Seoul [16]. Similarly, Sameen et al. developed a model using six years of accident data from Malaysia based on RNN, achieving 72% accuracy in predicting injury severity [17]. Ren et al. used traffic data from 2016 to 2017 to forecast traffic accident risks in Beijing based on RNN, too. Their research showed that accidents are more frequent in city centers, especially during peak hours [1]. Zhao et al. addressed the restraints of conventional machine learning models by proposing a CNN-based algorithm for traffic accident prediction that extracts key features from large datasets in vehicular ad hoc networks, enabling real-time collision warnings and enhancing vehicle safety [18].

An Artificial Neural Network (ANN) was employed by Chakraborty et al. to forecast fatal pedestrian accidents at Indian crossroads. Vehicle speed and daily traffic volume were shown to be the main factors impacting deaths in their study [19]. Thaduri et al. highlighted the drawbacks of conventional techniques in India, including small sample sizes and data noise. They developed a CNN-based model that considers factors like traffic flow, weather, and lighting, outperforming traditional neural network approaches [20]. Alkheder et al. analyzed 5,740 traffic accident reports from Abu Dhabi, UAE, to identify factors linked to injury severity. They tested four algorithms (ANN, RF, SVM, and BN) and found that accident type and seatbelt usage were the most critical factors. The SVM model provided valuable insights into traffic accidents in the UAE and demonstrated strong predictive performance [21].

In their investigation of factors influencing accident severity, De Oña et al. classified outcomes as severely injured, died, or slightly injured using a Bayesian network. Their findings demonstrated that Bayesian networks can graphically depict complicated systems and are useful for factor prediction without the need for presumptions [22]. For the analysis of huge datasets, Simoncic also discovered that Bayesian networks were more effective than regression models [23]. Regression analysis, including Logistic Regression models, has been used to investigate relevant factors and forecast the severity of accidents. After comparing Bayesian networks and regression models, Zong et al. concluded that Bayesian networks are more appropriate for simulating the severity of accidents. This realization has significant ramifications for enhancing road safety and preventing accidents [24].

Gu et al. developed a Support Vector Machine-based model for predicting traffic fatalities, emphasizing the importance of parameter selection and mutation optimization. Their model achieved high precision with minimal errors [25]. Panwar et al. proposed an SVM-based model for predicting road accidents in India, achieving 67% accuracy using the Kernel scheme combined with a minority over-sampling technique [26]. Hussain et al. explored the performance of data mining algorithms, including Multi-layer Perceptron (MLP), for predicting accident-related factors such as crash causes, locations, and vehicle types. The MLP classifier achieved 85% accuracy, showcasing its effectiveness [27]. Using variables including weather and road layout, Roland et al. created an MLP-based model to pinpoint accident hotspots in a Tennessee city. The approach was used as a live service to improve emergency response and law enforcement [28].

Nevertheless, these approaches often focus on either temporal or spatial features and struggle to model spatiotemporal relationships simultaneously. In this work, the dynamic spatiotemporal correlations in traffic data are captured by the new Transformer-based deep learning model, Traffic-Transformer. The suggested approach shows better predictive accuracy than conventional LSTM and CNN models.

2. MATERIALS AND METHODS

2.1 Transformer Model Development

CNN and LSTM are powerful models in AI applications like time series forecasting and natural language processing. To extract certain features, CNN's structure accurately recognizes the layered structure of incoming data, simulating how the human visual system observes the outside environment. CNN uses convolutional filters, which go over the feature map at a predetermined pace and share filter weights throughout the feature map, in contrast to fully connected neural networks. This significantly reduces the model's parameter count. To further guarantee attention to local features, each neuron in a CNN layer is exclusively coupled to a subset of neurons in the layer above. CNN may progressively learn more intricate and global features by stacking many layers. CNN outperforms fully connected neural networks in feature extraction, thanks to its hierarchical learning mechanism.

In contrast, the LSTM structure is made up of a unit called a cell. This cell contains three different kinds of gate structures: input, output, and forget. The gradient vanishing and explosion problems that are typical of conventional RNNs can be avoided by using these gates to enable the LSTM to selectively recall and forget information. LSTM has been widely used in sequence modeling jobs due to its capacity to model lengthier input sequences and extract sequential abstract information. They are also especially well-suited for activities requiring sequential information over extended periods because of their flexible memory capacity [29]. Even though CNN has many positive aspects, its operating model nevertheless has serious flaws. The convolutional traversal of kernels over the spatial range of feature maps predicts the CNN feature extraction approach, which limits the receptive scope. A linear superposition of numerous layers is required for the design to successfully navigate and absorb long-term dependencies in sequential data, which multiplies the number of parameters [30].

Degradation, a reversal in network performance, could be caused by this architectural depth. LSTM, on the other hand, with its complex gated units, is better at interpreting dependencies over larger periods than CNN can handle. However, due to its inherent recurrent architecture, each LSTM cell must wait in turn for the computational results of its predecessor, which limits the possibility of processing data in parallel. This type of sequential dependency results in observable inefficiencies when applied to the analysis of large-scale traffic data. Furthermore, even though the LSTM is more sophisticated than RNN and CNN at modeling long-term dependencies, it still suffers from reduced performance when handling large sequential inputs, indicating a recurring difficulty in modeling long temporal sequences.

We introduce the TrafficTransformer, a custom framework created to predict traffic accidents to address the drawbacks of the existing traffic prediction techniques. The TrafficTransformer's composite construction is depicted in **FIGURE 1**. The multivariate, irregular interval data of raw traffic incidents is transformed into a structured sampling of temporal sequences that may be sequence forecasted at the Traffic Transformer's input junction using a temporal discretization technique. The combination of input embedding and positional encoding, which produces representational vectors for every element of the input sequence, improves this transformation even further. The TrafficTransformer receives the matrix made up of these vectors as input [30].

The TrafficTransformer uses the Seq2Seq architecture, which consists of an encoder and a decoder, because traffic accident prediction is basically a time series forecasting problem. To find and document the hidden associations between input sequences, the encoder must convert the time series data into fixed-size context vectors. Then, using the context vectors it received from the encoder and the outcome of the previous temporal step, the decoder gradually builds the complete output sequence by forecasting forthcoming temporal components. Each of the series of identical layers that make up the encoder and decoder has a completely connected layer, multi-head attention, simultaneous residual connections [31], and layer normalization [32]. The self-attention mecha-

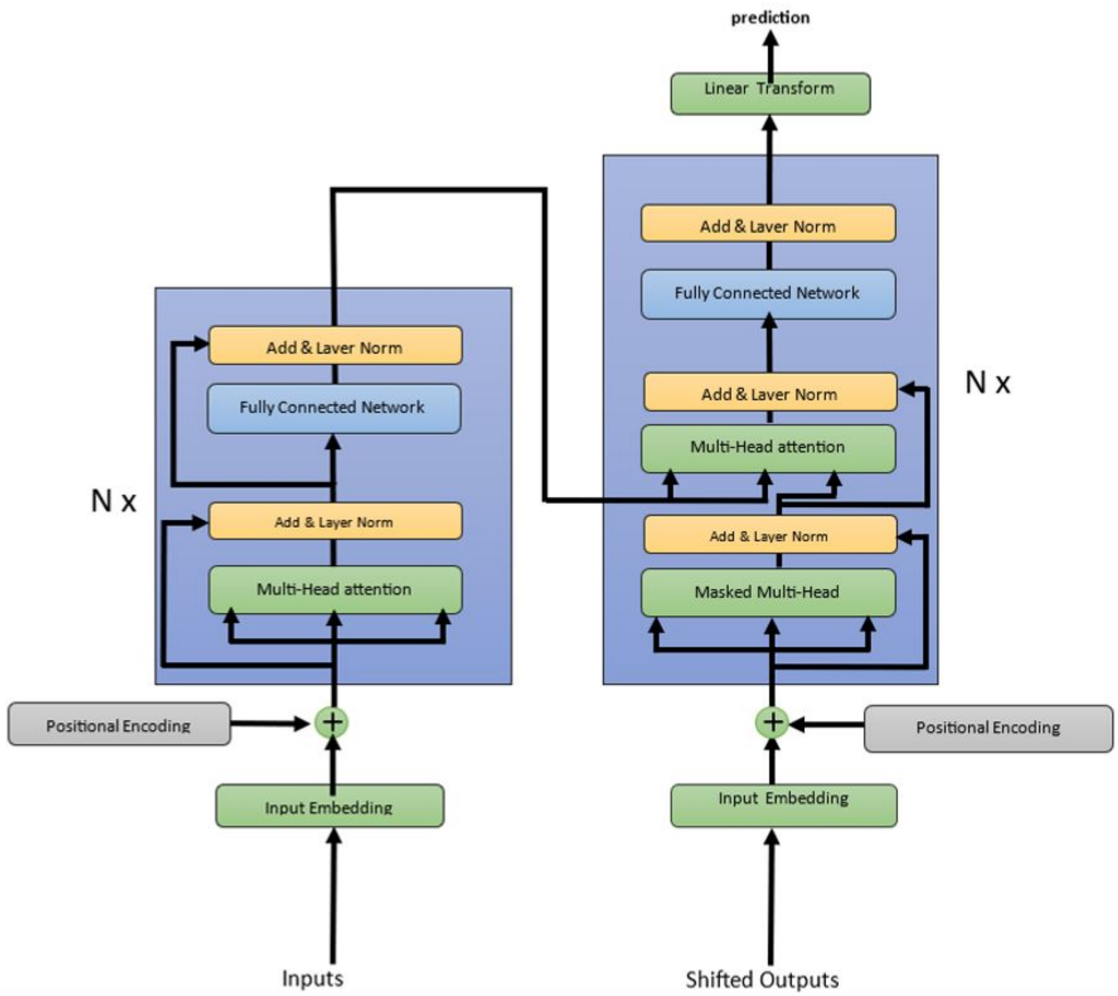


Figure 1: The overall architecture of TrafficTransformer, Source (the authors) [30].

nism, the TrafficTransformer's main component, makes it possible for the model to identify the link between any two items in the input sequence, allowing it to recognize long-term connections spanning arbitrary intervals.

It is evident that the TrafficTransformer architecture we suggest stands out for its creative substitution of a self-attention mechanism for the recurrent and convolutional structures often present in LSTM and CNN. This important modification allows any two elements in the input sequence to be directly correlated, allowing the model to evaluate each element while considering the entire sequence. As a result, the TrafficTransformer's global receptive field effectively captures long-term associations across sequences of any length without information loss, greatly overcoming the inherent shortcomings of CNN and LSTM. Furthermore, by visualizing attention weights, the TrafficTransformer facilitates comprehension of the model and helps to clarify the logic underlying its predictions. It also possesses parallel processing capabilities. These advantages provide the TrafficTransformer a distinct advantage over other well-known models already in use for traffic incident prediction.

Using 12 steps of historical traffic data, the TrafficTransformer model is tested for multi-step forecasting by predicting traffic accidents for 1, 3, and 6 future steps. To accelerate convergence, the model uses the Adam optimizer, which has an initial learning rate of 0.005 and a weight decay of 1×10^{-4}. After much testing, it was determined that the best configuration for the TrafficTransformer balanced efficiency and performance with two layers in the encoder and decoder and four attention heads in the multi-head attention module. Performance can be enhanced by adding more layers, but doing so uses more processing power. Using Pytorch 1.10.2+cu113 and Python 3.7, the model and comparison tests were constructed and executed on an Intel I7-12700F CPU and NVIDIA RTX 3070 GPU [30].

2.2 Experimental Settings

In this study, we decided to train and evaluate the model using the traffic dataset from Qatar. Qatar is home to about 2.68 million people and is located on the Arabian Peninsula's eastern coast. The country's highly updated transportation infrastructure, government-collected high-quality data, geographically limited region, and relatively tiny population make its traffic accident dataset extremely relevant for analysis and predictive modeling. These characteristics, along with Qatar's distinct environment and substantial technological investment, make the dataset an excellent option for predicting traffic accidents, which helps to improve model accuracy and the efficacy of preventative actions. Traffic accidents that happened in Qatar between January 2017 and October 2023 are documented in the dataset. There are 332,766 legitimate records in all after samples with missing values have been eliminated. Each record includes information about the accident time, the weather, and the area codes. Qatar was divided into 98 subareas, coded 1-98, to make forecasts across the various subareas. Subareas 45, 55, and 56, which contained 6,257, 7,725, and 8,315 samples, respectively, were chosen at random for the experiment. The data set was divided into 75% used for training, and the remaining for testing.

We introduced the LSTM-S2S-AM, which improves the LSTM-S2S with an attention mechanism, and the LSTM-S2S with a Seq2Seq structure. These are two well-liked models for forecasting traffic accidents. These LSTM-based comparator models offer a significant amount of benchmarking value and are the most widely used neural network designs for traffic accident prediction. Each model employed the same hyperparameter setups used during the tests with a batch size of 32, an Epoch of 50, a learning rate of 0.005, and weight decay of $1\text{x}10^{-4}$. Comparative analyses were performed on datasets from different regions of Qatar, and tests were undertaken over a range of forecast horizons to determine the models' multi-step forecasting capabilities. A consistent forecast horizon of six steps is used while making predictions over several different locations to guarantee experiment fairness and consistency. In addition, region 45 is always used as the dataset for multi-step forecasting.

To ensure the robustness of the model across diverse geographical contexts, validation experiments were conducted across areas with differing characteristics, including urban, suburban, and industrial zones. The impact of temporal factors, such as time of day (peak vs. off-peak hours) was also considered during model training through explicit feature engineering. However, weather conditions had no impact on the prediction as the weather is normally clear and sunny throughout the year in Qatar. These additions enabled the model to better capture spatiotemporal variations in accident patterns across different regions and times.

3. RESULTS AND DISCUSSION

Comparative experiments in areas coded 45, 55, and 56 are shown in **FIGURE 2**, and it is easily seen that, even with the same Seq2Seq architecture, the LSTM-based LSTM-S2S and its attention-augmented counterpart, LSTM-S2S-AM, perform significantly worse than our suggested Traffic-Transformer on several evaluation metrics. In region 56, where it performs the least well, Traffic-Transformer maintains a marginal error of just 0.26 when taking the MAE measure into account. This represents an error reduction of 88% and 84%, respectively, when compared to LSTM-S2S and LSTM-S2S-AM [30].

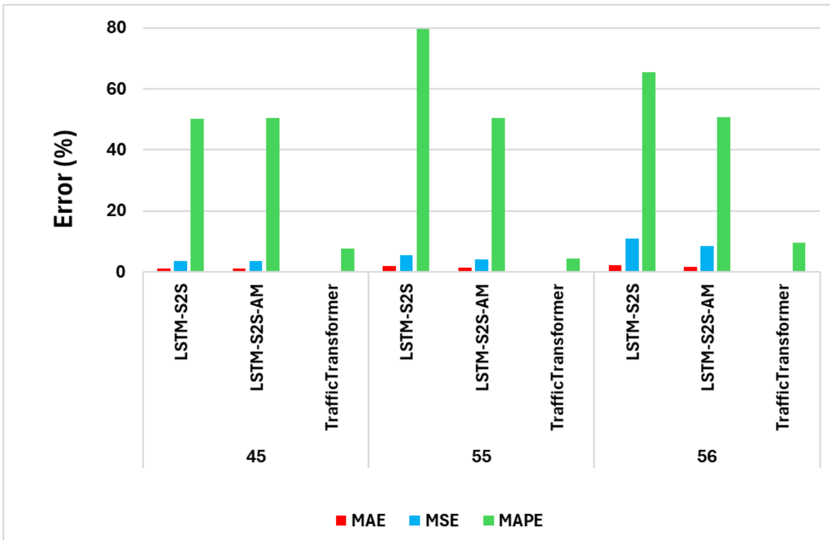
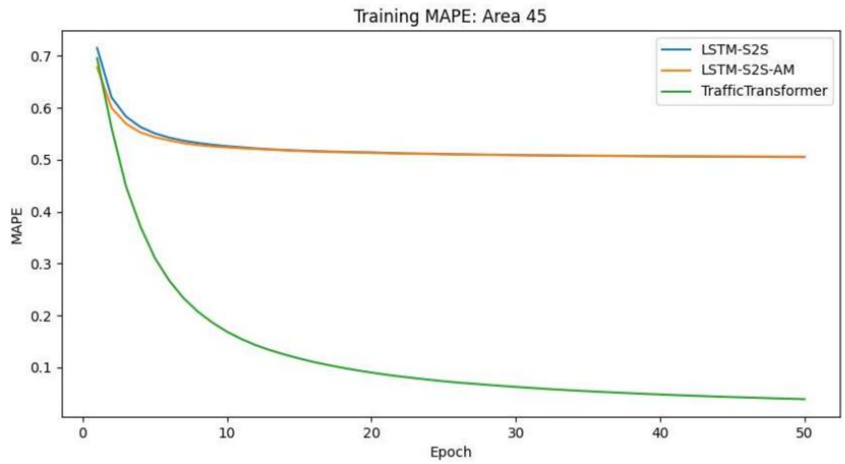


Figure 2: Performance in terms of error of each model in three area zones (45, 55, and 56).

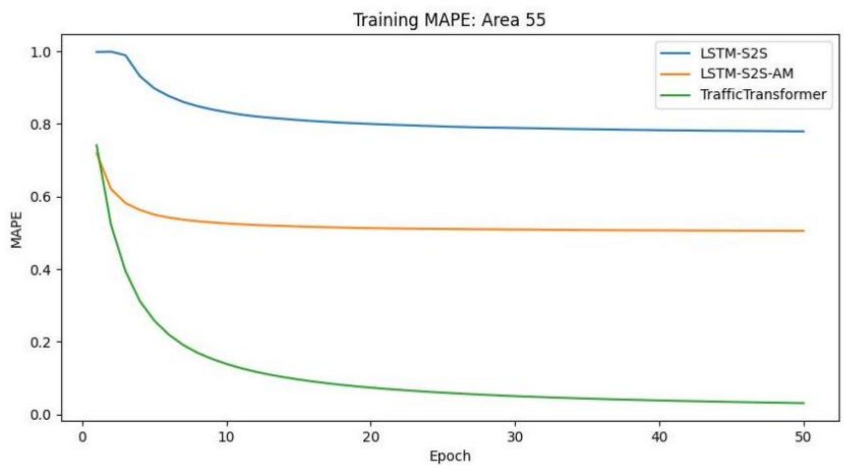
This demonstrates a notable increase in forecast accuracy, particularly a more trustworthy mean-level projection. TrafficTransformer's tolerance to noise or outliers is seen by the notable improvement in the MSE value. Given that MSE gives greater errors more weight, the model's clear superiority in this parameter implies that the improvement in prediction accuracy extends beyond average levels and covers the whole error range. [30].

Finally, the model's prediction error as a percentage of actual values is represented by the MAPE metric, which is displayed in **FIGURE 3**. TrafficTransformer performs noticeably better in every region, but particularly well in area 55, where its error is just 4.43%. This suggests that the Transformer model better reflects the trend changes in the data in terms of relative prediction accuracy. This is particularly crucial for traffic accident forecasting since developing successful preventative and control plans requires a detailed grasp of the trend of incidents [30].

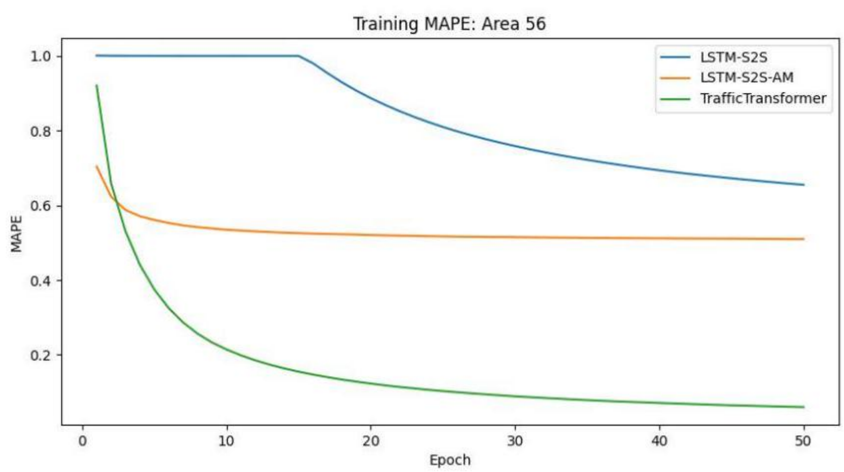
The experimental results show that the proposed TrafficTransformer model outperforms both the basic and attention mechanism-augmented LSTM models based on the Seq2Seq architecture in Qatar's multi-area traffic accident prediction task. The unique self-attention method of TrafficTransformer, which captures long-term dependencies across the input sequence, is the primary source of the performance disparity. However, even the LSTM-S2S-AM model, with its attention mechanism, requires improvement to understand and simulate complex spatio-temporal interactions because



(a)

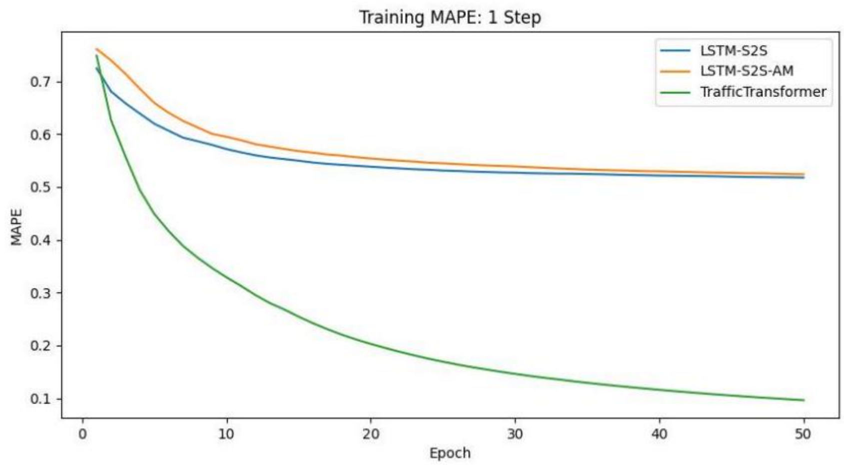


(b)

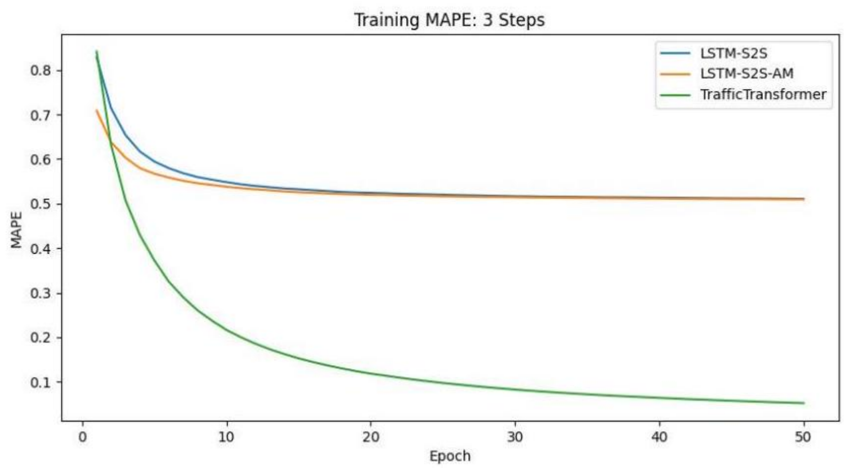


(c)

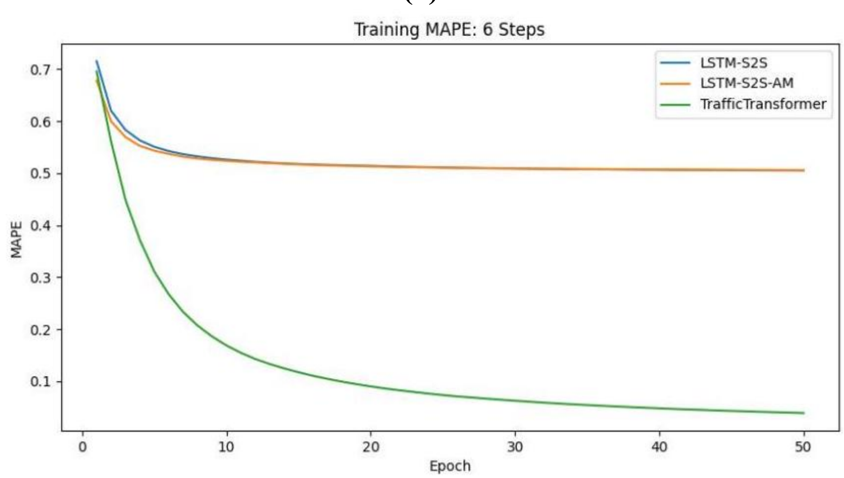
Figure 3: Variation in MAPE during model training across different area zones (a) 45, (b) 55, and (c) 56. The source: the authors [30].



(a)



(b)



(c)

Figure 4: MAPE variation trends during model training over different forecast horizons, (a) 1 step, (b) 3 steps, and (c) 6 steps. The source: the authors [30].

of its inherent sequential data processing approach. The results in **FIGURE 4**, demonstrate the TrafficTransformer's excellent performance throughout a range of predicting intervals of 1, 3, and 6 steps. An analysis of the attention weights within the TrafficTransformer model revealed that the most influential factors contributing to accurate traffic accident predictions were the time of day, weather conditions, and the density profile of the geographical zone. Time-of-day features, especially peak morning and evening periods, were consistently associated with higher predicted accident risks. Incorporating weather and area-specific characteristics significantly enhanced the model's ability to capture spatiotemporal accident patterns

4. CONCLUSIONS

Due to the growing disparity between the demand for mobility and the ability of urban infrastructure to deliver it, information technology solutions (ITS) must be developed quickly. Predicting traffic accidents while thoroughly taking into account intricate external elements is one method to help ITS develop. Traffic accidents rank among the most significant public safety issues since they are the leading cause of mortality for individuals of all ages.

Weather and road conditions, and lighting are examples of environmental elements that can increase the risk. Generally speaking, the objective of traffic accident prediction is to use statistics, analysis, and processing of relevant data from past incidents to produce reliable and scientific forecasts about upcoming traffic accidents. Predicting traffic has both advantages and disadvantages. Traffic prediction is now more possible thanks to the transportation industry's growth and the deployment of rich data from information gathering devices on roads. In many respects, however, predicting traffic accidents is a difficult task.

Accurate traffic accident prediction is crucial for both property and life protection as well as urban growth. This study recommends TrafficTransformer for precise multidimensional and multi-step traffic accident predictions. TrafficTransformer substitutes a complex multi-head attention mechanism for the conventional recurrent architecture included in mainstream models to provide multifarious parallel modelling of the global information encoded in input sequences. Experiments conducted on the Qatari traffic incident dataset repeatedly demonstrate the TrafficTransformer's enhanced predictive capability, outperforming earlier models in terms of accuracy and durability across a range of forecast horizons and locales. Furthermore, the TrafficTransformer's lightweight and efficient inference phase makes it well-suited for real-time or near-real-time deployment in operational traffic management and emergency response systems

5. LIMITATIONS AND COMPUTATIONAL CONSIDERATIONS

Although the TrafficTransformer model demonstrated strong predictive performance across various zones and periods within Qatar, several limitations should be acknowledged. The findings are based on a dataset specific to Qatar's road network, urban design, and driving culture, and may not be directly generalizable to other countries with different environmental, infrastructural, and socio-economic characteristics. Additionally, the possibility of biases in the dataset, such as under-reporting of minor accidents or inconsistencies in environmental factors like weather, may affect the robustness of the predictions. From a computational perspective, while the model's training phase requires moderate computational resources due to the use of deep attention mechanisms, the inference phase is highly efficient. This efficiency allows the model to be deployed in near-real-time applications, such as emergency response centers and smart city traffic monitoring systems.

6. PATENT

This work has been patented as: Chaojie Li, Yin Yang, Mansoor Al-Thani, An Interpretable Graph-based AI System for Traffic Accident Prediction, Patent No. US2023140289, filed in 2023.

6.1 Funding

This research received no external funding.

6.2 Data Availability Statement:

Data is unavailable due to privacy restrictions.

6.3 Acknowledgments

The authors would like to acknowledge the Ministry of Interior for sharing the traffic data.

6.4 Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Ren H, Song Y, Wang J, Hu Y, Lei J, Ren, H., Song, Y., Wang, J., Hu, Y., & Lei, J. (2017). A Deep Learning Approach to the Citywide Traffic Accident Risk Prediction. 2018 21st International Conference on Intelligent Transportation Systems (ITSC). IEEE. 2018:3346-3351.

[2] Rahman S. Development of an Accident Prediction Model for Intersections of Dhaka. Int J Comput Appl. 2012;47:10-16.

[3] https://www.who.int/publications-detail/global-status-report-on-road-safety-2018.

[4] Polson NG, Sokolov VO. Deep learning for short-term traffic flow prediction. Transp Res C. 2017;79:1-7.

[5] https://www.proquest.com/openview/c6332b653d33fb0d5a19954f64f26c4e/1?pq-origsite=gscholar&cbl=2026366

[6] https://www.openarchives.gr/aggregator-openarchives/edm/IKEE_AUT/000097-297413?language=en

[7] Wibisono A, Jatmiko W, Wisesa HA, Hardjono B, Mursanto P. Traffic big data prediction and visualization using fast incremental model trees-drift detection (FIMT-DD). Knowledge-Based Systems. 2016;93:33-46.

[8] Khokale R, Ghate A. Data Mining for Traffic Prediction and Analysis using Big Data. Int J Eng Trends Tech. 2017:48.

[9] Ke J, Zheng H, Yang H, Chen X. Short-Term Forecasting of Passenger Demand Under On-Demand Ride Services: A Spatio-Temporal Deep Learning Approach. Transp Res C. 2017;85:591-608.

[10] Lv Y, Duan Y, Kang W, Li Z, Wang F. Traffic Flow Prediction With Big Data: A Deep Learning Approach. IEEE Transactions on Intelligent Transportation Systems 2014;16:865-873.

[11] Marcillo P, Valdivieso Caraguay ÁL, Hernández-álvarez M. A Systematic Literature Review of Learning-Based Traffic Accident Prediction Models Based on Heterogeneous Sources. Appl Sci. 2022;12:4529.

[12] Duan Y, Lv Y, Wang F. Travel Time Prediction With LSTM Neural Network. IN 2016 IEEE 19th International Conference on Intelligent Transportation Systems. IEEE. 2016:1053-1058.

[13] Li Y, Yu R, Shahabi C, Liu Y. Diffusion Convolutional Recurrent Neural Network. Data-Driven Traffic Forecasting. In: Proceedings of The ICLR. 2018:1-16.

[14] Kipf TN, Welling M. Semi-supervised Classification With Graph Convolutional Networks. In: Proceedings of the ICLR; 2017:1-14.

[15] Ma X, Dai Z, He Z, Ma J, Wang Y, et al. Learning Traffic as Images: A Deep Convolutional Neural Network for Large-Scale Transportation Network Speed Prediction. Sensors. 2017;17:818.

[16] Park Sh, Kim Sm, Ha Yg. Highway Traffic Accident Prediction Using Vds Big Data Analysis. J Supercomput. 2016;72:2815-2831.

[17] Sameen MI, Pradhan B. Severity Prediction of Traffic Accidents With Recurrent Neural Networks. Appl Sci. 2017;7:476-493.

[18] Zhao H, Cheng H, Mao T, He C. Research on Traffic Accident Prediction Model Based on Convolutional Neural Networks in VANET. In: Proceedings of the 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD). IEEE. 2019:79-84.

[19] Chakraborty A, Mukherjee D, Mitra S. Development of Pedestrian Crash Prediction Model for a Developing Country Using Artificial Neural Network. Int J Inj Contr Saf Promot. 2019;26:283-293.

[20] Thaduri A, Polepally V, Vodithala S. Traffic Accident Prediction Based on CNN Model. In: Proceedings of the 5th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE. 2021:1590-1594.

[21] AlKheder S, AlRukaibi F, Aiash A. Support Vector Machine (Svm), Random Forest (Rf), Artificial Neural Network (Ann), and Bayesian Network for Prediction and Analysis of Gcc Traffic Accidents. J Ambient Intell Human Comput. 2023;14:7331-7339.

[22] Oña de J, Mujalli RO, Calvo FJ. Analysis of Traffic Accident Injury Severity on Spanish Rural Highways Using Bayesian Networks. Accid Anal Prev. 2011;43:402-411.

[23] Simoncic M. A Bayesian Network Model of Two-Car Accidents. J Transp Stat. 2004;7:13-25.

[24] Zong F, Xu H, Zhang H. Prediction for Traffic Accident Severity: Comparing the Bayesian Network and Regression Models. Math Probl Eng. 2013;2013:1-9.

[25] Gu X, Li T, Wang Y, Zhang L, Wang Y, et al. Traffic Fatalities Prediction Using Support Vector Machine With Hybrid Particle Swarm Optimization. J Algorithms Comput Technol. 2017:1-10.

[26] Panwar N, Rani R. Road Accident Prediction by Support Vector Machine Using Machine Polynomial Kernel and Synthetic Minority Oversampling Technique (SMOT). International Journal of Research in Electronics and Computer Engineering. 2021;9:31-38.

[27] Hussain S, Muhammad LJ, Ishaq FS, Yakubu A, Mohammed IA. Performance Evaluation of Various Data Mining Algorithms on Road Traffic Accident Dataset. In: Satapathy, S., Joshi, A. (eds) Information and Communication Technology for Intelligent Systems . Smart Innovation, Systems and Technologies Singapore: Springer Singapore. 2018;106:67-78.

[28] Roland J, Way PD, Firat C, Doan TN, Sartipi M. Modeling and Predicting Vehicle Accident Occurrence in Chattanooga, Tennessee. Accid Anal Prev. 2021;149:105860.

[29] Yu B, Yin H, Zhu Z. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. In: Proceedings of the the 27th International Joint Conference on Artificial Intelligence. ACM. 2018:1-7.

[30] https://www.proquest.com/openview/a7af87b530fdb8734691d45df1c26e94/1?pq-origsite=gscholar&cbl=2026366&diss=y

[31] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: Proceedings of the the Ieee Conference on Computer Vision and Pattern Recognition. IEEE. 2016:770-778.

[32] Ba JL, Kiros JR, Hinton GE. Layer Normalization. 2016. Arxiv preprint: https://arxiv.org/pdf/1607.06450