Vol. 08, No. 3 (2026) 1457-1468, doi: 10.24874/PES08.03.006

Proceedings on Engineering Sciences

www.pesjournal.net

DEEP LEARNING–BASED ACCIDENT DETECTION IN INTELLIGENT TRANSPORT NETWORKS

Anil Laxmanrao Wakekar[1]
Kottala Sri Yogi
Vishal Anil Aher
Harshada Bhushan Magar
Ashutosh Kulkarni
Deepika Sharma

Received 12.05.2026.
Revised 26.06.2026.
Accepted 14.07.2026.

Keywords:

Accident detection; Intelligent transport networks; Traffic safety; Spatiotemporal analysis; Real-time monitoring

A B S T R A C T

The rapid development of intelligent transport systems has contributed greatly to traffic density thus timely and accurate detection of accidents is one of the key demands in improving road safety and emergency response. Older sensor and rule based methods are characterized by slow detection, inadequate scalability and resistance to more complex traffic conditions. In order to tackle such issues, this paper proposes an accident detection system that is based on deep learning and is applicable in the context of intelligent transport systems. The main aim of the study is to design and perform the evaluation of strong deep learning frameworks with the ability of automatically identifying traffic accidents on heterogeneous types of traffic data with high accuracy and low latency. The framework is based on four deep learning approaches, which include Convolutional Neural Networks (CNNs) to extract spatial features of traffic images, Long Short-Term Memory (LSTM) networks to model temporal traffic dynamics, a CNNLSTM hybrid architecture to identify accidents on a spatiotemporal scale, and a Transformer-based attention system to learn long-range dependencies and be mindful of context. These models are supervised trained and evaluated on traffic video and sensor datasets preprocessed and inputted into them. The experimental findings indicate that the CNN model has a higher accident detection rate of 91.3% compared to LSTM model with 93.8% because it is able to extract and capture the temporal patterns. The CNN LSTM hybrid has a high accuracy of 96.4% and a low false alarm rate of 3.1 as compared to individual models. The Transformer-based model has the highest performance with an accuracy of 97.9 %, 98.2 % and lower detecting latency than the current methods. These results prove that accident detection using deep learning has the potential to greatly improve reliability, responsiveness, and scalability in intelligent transport networks.

© 2026 Published by Faculty of Engineering

Corresponding author: Anil Laxmanrao Wakekar
Email: anil.wakekar@gmail.com

1. INTRODUCTION

Intelligent transport networks have emerged to be an essential part of the contemporary smart city that allow the sophisticated traffic management, control and safety by the virtue of interconnected sensing, communication and calculation systems. Nevertheless, road traffic accidents continue to be a major issue in the world due to their excessive mortality, economic, and long-term consequences to the society. Traffic complexity is further increased by the rapid urbanization process and the growing number of vehicles owned by people, so in order to reduce the emergency response time and avoid secondary collisions, traffic accidents require timely detection. Automated accident detection systems are thus an important need in enhancing the situational awareness and operational efficiency of intelligent transportation systems (Grigorev et al., 2024). Traditional methods of accident detection are mostly sensor-based, or hand-designed rules, or pre-defined traffic flow indicators based on the loop detectors, acceleration sensors, or crude vision systems. Although these methods are simple and have low computational complexity, they are not very robust in the real-world traffic environment with occlusions, changes in light, traffic, and unpredictable driver actions. Such systems are characterized by high false alarm rates, and they are difficult to change to adapt to dynamic traffic conditions, which makes them ineffective on large scales (ElSahly & Abdelfatah, 2024). Moreover, sensor-centric systems are also very expensive to install and maintain, which reduces their applicability when it comes to wide-area smart transport networks (Pathik et al., 2022). Recent work points to the fact that rule-based and statistical detection of accidents models are intrinsically limited by the fact that they rely on manually constructed features and assumption-based models that are not dynamic. These constraints make these models unable to generalize on the basis of heterogeneous road networks and traffic patterns. Consequently, the accuracy of detection of accidents reduces in challenging urban environments, especially during rush hours or bad weather (Olugbade et al., 2022). This makes the problem of adaptive and data-driven detection frameworks that can learn accident patterns based on large-scale traffic data important.

Deep learning has recognized itself as a strong substitute with enhanced representation learning potentials that can utilize automated derivation of discriminative features out of unprocessed information. Convolutional Neural Networks (CNNs) have shown high results revealing the presence of spatial accident indicators in traffic surveillance videos through learning visual attributes related to crashes and suspicious vehicle actions (Pawar & Attar, 2022). Deep learning models based on video also enhance the reliability of detection by including the contextual information that is not limited to mere motion or change in intensity (Wu & Li, 2024). Such properties render deep learning especially appropriate to address the variability and uncertainty of about the real-world traffic conditions. Accident detection research has also become interested in the temporal modeling. RNN with LSTM architectures are more effective in learning sequential dependencies in vehicle movement capturing early situations that may lead to accidents (Rajesh et al., 2020). Spatial learning and temporal modeling Hybrid deep learning models that combine CNN-based spatial learning with LSTM-based temporal modeling have demonstrated higher performance through the presence of a spatiotemporal feature (Erkartal & Yilmaz, 2025). In more recent work, Transformer architecture, which is motivated by attention, has been shown to exhibit better capacity to learn long-range dependencies and contextual relationships in traffic data and has been shown to be significantly more effective at real time detection accuracy and robustness (Seghir et al., 2026).

Meanwhile, the lightweight object detector models like YOLO-based models have been modified to fit the accident detection to satisfy real-time processing needs of an intelligent transport network. Better YOLO designs are used to attain high detection accuracy yet minimal inference latency allowing them to be deployed to resource-limited settings (Li et al., 2025a). These innovations are in line with the growing need to have scalable low-latency accident detection systems that can be integrated with edge computing and automated vehicle systems (Almutairi et al., 2025). Besides, the adaptation capability of deep learning-based surveillance systems has been shown to be very high in more than just one set of applications, which confirms their use in safety-critical transportation (Karim et al., 2024). This study is inspired by such developments and is aimed at developing and testing an accident-detecting framework using deep learning and specific to intelligent transport networks. The research will methodically evaluate various deep learning designs and determine the most useful model in order to detect accidents in real-life traffic situations in an accurate, timely and scalable way.

Key contribution given as:

- Invents an integrated model of deep learning of spatial, temporal and attention-based models in intelligent transport networks that are robust in detecting accidents.
- Compares various deep learning systems based on their accuracy, recall rate, false alarm rate and latency to find the best performance of deep learning when detecting accidents in real-time.
- The contribution proves the ability of detecting accidents in a very scaled and reliable fashion that is applicable in the deployment of smart cities in a complex traffic and environmental scenario.

2. RELATED WORK

The initial studies in the field of accident detection in intelligent transportation systems were mainly aimed at traditional methods of detection founded on a set of rules,

traffic flow limits, and infrastructure-based sensors. These systems normally depended on speed difference, occupancy proportion, or sudden slowdown patterns taken by loop detectors, cameras, or wireless sensor net. Even though these methods allowed simple detection of incidents, they did not provide flexibility to dynamically changing traffic scenarios and were very susceptible to noise and failure of sensors. Intelligent transportation systems aided by blockchains and sensors have tried to enhance reliability and data integrity, but the use of manual code and hard-coded assumptions restricts identifying the detection in complicated urban environments (Alharbi et al., 2023). These constraints led to the search to find data-based approaches to enhanced robustness and scalability. Approaches that were built on machine learning was an important change in terms of the strict rule-based systems to adaptive accident detection models. This learnt the traffic patterns using traditional classifiers and regression models, which was able to generalize better than the deterministic systems due to their ability to learn the traffic pattern based on historical data. Prediction models made by neural networks showed greater performance in the modeling of nonlinear traffic behavior and accidents rate caused by congestion on the basis of real-life urban data (Alfasi et al., 2025). Ensemble learning and hybrid probabilistic models also enhanced the detectability of the variety of learning paradigms that were integrated to deal with the uncertainty and noise of traffic observations (Erkartal & Yilmaz, 2025). Also, the technologies that use non-visual sources of data, including vehicle telemetry and traffic statistics, have demonstrated potential to predict the likelihood of accidents without the use of video streams (Chai et al., 2024).

However, in more recent times, deep learning has become the paradigm of choice in terms of traffic incident detection and safety analysis owing to the high abilities of feature representation. Ensemble models based on deep learning have been used to predict traffic accidents using the unconventional source of data, such as social media feeds, and identify the latent spatiotemporal patterns associated with the public reporting and mobility patterns (Gutierrez-Osorio et al., 2022). The deep neural networks based on YOLO have become a prominent topic due to its ability to detect objects in real-time, allowing to detect and classify the severity of accidents effectively and in traffic cameras (Li et al., 2025b). Vehicle damage assessment Two-stage deep learning pipelines are also suggested in the context of post-accident detection to enhance the accuracy of post-incident analysis in intelligent transport systems (Wahyudin & Rachmawati, 2025). Deep learning architectures have also been improved to improve traffic safety applications. Deep learning models based on vision have proved to be useful in automatic detection of traffic incidences through monitoring of abnormal vehicle interactions and motion patterns (Karim et al., 2024). Transformer-based models that contain attention mechanisms have appeared recently to address the drawback of convolutional and recurrent models in long-range dependency capture. These models offer a better contextual interpretation and a low-detection latency in Internet of Vehicles space (Seghir et al., 2026). Simultaneously, the flexibility and robustness of present-day deep learning methods to real-time monitoring are indicated by deep learning systems designed in other safety-critical areas, such as radar-based fall detection and thermographic anomaly detection (Almallah & Sababha, 2026; Danciu et al., 2024).

Table 1. Summary of Related Work on Deep Learning–Based Accident Detection in Intelligent Transport Networks

Author name	Year	Data Source	Method / Model	Key Objective	Key Limitation
RoblesSerrano et al.	2021	Traffic video	CNN-based DL	Automatic accident detection	Limited temporal modeling
Pathik et al.	2022	IoT sensors + video	CNN + IoT framework	Accident detection and alerting	Scalability constraints
Gutierrez-Osorio et al.	2022	Social media data	DL ensemble model	Accident prediction	Indirect data dependency
Pawar and Attar	2022	Surveillance video	CNN-based localization	Accident detection and localization	High computational cost
Alharbi et al.	2023	WSN + LPR data	ML-based ITS framework	Intelligent traffic monitoring	Rule-based dependency
Wu and Li	2024	Traffic surveillance video	Deep CNN	Video-based accident detection	Sensitive to occlusion
Grigorev et al.	2024	Traffic video	ML-based detection & segmentation	Detection and duration prediction	Complex pipeline
Karim et al.	2024	Vehicle activity video	DL-based visual analysis	Traffic incident detection	Limited real-time analysis
Lin et al.	2025	Traffic datasets	DL prediction systems	Accident prediction modeling	Prediction-focused, not detection
Alqahtani & Kumar	2025	In-vehicle network data	DL + knowledge graph	Intrusion detection in vehicles	Not accident-specific
Thakur et al.	2025	Road traffic images	Precision-aware DL model	Accident detection accuracy	High model complexity
Almallah and Sababha	2026	Radar sensor data	Lightweight DL (IntNet)	Real-time event detection	Domain-specific adaptation

The Table 1 provides the overview of the recent studies in accident detection using deep learning, machine learning, and hybrid frameworks on the bases of video, sensor, and IoT data. Most of them concentrate on single architectures or on particular modalities of data although many of them are highly accurate. Weaknesses in time modeling, real-time latency, and unified assessment demonstrate the necessity to have end-to-end and scalable deep learning systems.

3. SYSTEM OVERVIEW

The specified system should contribute to the accurate and real-time tracking of accidents in intelligent transport networks by harmonizing sensing, communication, and deep learning-based analytical vehicles. It offers a single platform that links traffic information sources with tools of superior processing and decision-making to improve road safety and efficiency in traffic management.

3.1 Intelligent Transport Network Architecture for Accident Detection

The intelligent transport network structure has three major layers, which are data acquisition, data processing, and the decision support. The data capture layer contains a network of traffic surveillance cameras, road-side sensors, and IoT-enabled devices that are installed on the parts of the road to observe the traffic flow and the environmental factors in real-time. These devices produce heterogeneous streams of data sent either by wired or wireless communication infrastructure to centralized or edge based processing units. The data processing layer includes the preprocessing of data, feature extraction and deep learning inference with CNN, LSTM, CNN LSTM, and Transformer models. Lastly, the decision support layer will produce the outputs of accident detection, severity, and alert that can be sent to the emergency services, traffic control centers, or smart city displays to respond promptly and alleviate the situation.

3.2 Problem Definition and Detection Requirements

The essence of the issue discussed in this paper is the timely and correct detection of traffic accidents based on spatiotemporal traffic information in various road and environmental conditions. In formal terms, having a series of observations of traffic, the task is to define whether an accident has taken place or not, estimate its severity, and provide a corresponding alert within a minimum delay. The main detection specifications are that it must have a high accuracy in the reduction of false classification, low false alarm rate to prevent unnecessary interventions, and lower detection latency in order to facilitate real-time response. Also, the system should be able to scale to high traffic networks and withstand noise, variability, and dynamic traffic behaviour.

4. METHODOLOGY

The research methodology is the deep learning-based approach to the real-time detection of accidents in intelligent transport networks. The methodology starts by cleaning and normalizing data systematically and adding noise-tolerant and realistic data to make the inputs realistic, balanced, and noise-tolerant. Important spatial and temporal aspects are then derived out and put together in sequences in order to describe the traffic dynamics before the occurrences of accidents. Several deep learning models are used, such as CNNs in the learning of spatial features, LSTMs in the modeling of temporal pattern features, a CNN-LSTM in the learning of integrated spatiotemporal features and Transformer-based attention models in the learning of long-range dependencies. The methodology allows detecting the accidents in a precise, scalable, and low-latency manner by integrating the models into a single pipeline, regardless of the traffic conditions and environmental factors. This Figure 1 shows the single pipeline of end-to-end accident detection, which begins with the multimodal traffic data acquisition and pre-processing and proceeds to spatial, temporal, and spatiotemporal deep learning models. The Figure 1 presents the deep learning–based accident detection framework, illustrating data acquisition, preprocessing, feature extraction, classification, accident detection, alert generation, and response coordination.

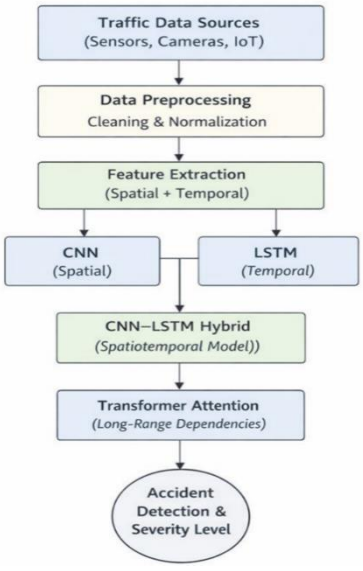


Figure 1. Block Diagram of the Deep Learning–Based Accident Detection Framework in Intelligent Transport Networks

4.1 Data Cleaning, Normalization, and Augmentation Techniques

In order to have strong robustness and stable convergence of the proposed deep learning models, the proposed training used a systematic preprocessing of the synthetic transportation dataset. Though the dataset was artificially

created and to a great extent devoid of missing values, logical consistency tests were conducted to eliminate irregularities in records, including the presence of conflicting values in traffic density and the number of vehicles. This cleaning process made sure that the model learning was carried out using realistic observations of traffic only. In numerical variables such as vehicle number, traffic density, average speed and visibility index, normalizing was done to prevent the dominance of scale in optimization. Each feature was rescaled by min-max normalization to the [0,1] range and this is expressed as:

$$x^{'} = \frac{x - x_{min}}{x_{\max - x_{min}}}$$

In which, x is the initial feature value, x min and x max are the lowest and highest values of the feature in the dataset. The one-hot encoding was used to convert the categorical attributes like weather condition, road type and surface state to maintain none ordinal relationships. To enhance generalization and curb overfitting, the time data augmentation had to be carried out by adding insignificant stochastic noise to traffic-related aspects. As augmented samples were produced as:

$$\tilde{x}_t = x_t + \varepsilon, \qquad \varepsilon \sim N(0, \sigma^2)$$

Where, ε is the noise of Gaussian noise with a small variance σ 2, the simulation of natural traffic variation without distorting semantic meaning.

4.2 Feature Extraction and Sequence Construction

In this study, feature extraction will be used to derive significant spatiotemporal representations useful in the detection of accidents:

$$f_t = [d_t, v_t, s_t, w_t, r_t]$$

Where, d_t represents the traffic density, v_t the number of vehicles, s_t average speed, w_t weather-related variables and r_t road condition variables. These feature vectors are as inputs to spatial learning layers which encode inter-feature relationship that affect accidents.

Traffic observations are structured into sequences with a sliding window method in order to model the dynamics in time:

$$X_t = \{f_{t-T+1}, f_{t-T+2}, \dots, f_t\}$$

This sequential representation is introduced to LSTM and Transformer models to learn the patterns of temporal evolution of the events in advance by accident:

$$h_t = LSTM(f_t, h_{t-1})$$

Everyone should be able to save long-term contextual information on their network.

In case of Transformer-based modeling, positional encoding p_t is introduced to maintain the temporal ordering:

$$z_t = f_t + p_t$$

This allows the self-attention mechanism to pay attention to important time steps that lead to accident occurrence. The proposed framework successfully uses the normalized, augmented, and temporally coherent data in the accuracy and real-time detection of accidents, through this structured preprocessing, feature extraction and sequence construction pipeline.

4.3 Convolutional Neural Network (CNN) for Spatial Feature Extraction

In this study, Convolutional Neural Networks (CNNs) are used to obtain discriminative spatial features of the data of the traffic in terms of instant road and environmental conditions. CNNs are especially efficient to discover local spatial relationships between features of inputs and, thus, they can be used to discover patterns of accidents, including abnormal traffic density, abrupt speed changes, and conditions in bad weather. Under the suggested framework, individual observations of traffic are coded as the feature matrices, which are structured and convolutional filters are moved over feature dimensions to reveal spatial association among traffic movement, car density, visibility, and road conditions.

The convolution operation is defined as

$$yi, j = \sum m \sum nwm, n \cdot xi + m, j + n + b$$

Where, w is the convolution kernel, x the input feature map and b the bias term. The dimensionality is reduced through use of pooling layers which suppress noise, which increases resilience to small changes in traffic. Through the acquisition of hierarchical spatial representations, the CNN is efficient in separating maladaptive traffic states against normalcy, as shown in algorithm 1, which is a robust base to the final temporal and spatiotemporal modeling in the general accident-detecting pipeline.

Algorithm 1: CNN-Based Spatial Feature Extraction for Accident Detection

Input: Traffic feature matrix X ∈ R^(H×W×C)
Output: Spatial feature vector F_s
1: $Initialize\ convolution\ kernels\ W_c\ and\ bias\ b_c$
2: $for\ each\ convolution\ layer\ l\ do$
3: $Compute\ convolution$:
4: $Y_l = W_c * X + b_c$
5: $Apply\ activation\ function$:
6: $A_l = ReLU(Y_l)$
7: $Apply\ pooling\ operation$:
8: $P_l = MaxPool(A_l)$
9: $Update\ X \leftarrow P_l$
10: $end\ for$
11: $Flatten\ final\ feature\ maps$:

12: $F_s = Flatten(P_l)$
13: $Return\ spatial\ feature\ vector\ F_s$

4.4 Long Short-Term Memory (LSTM) Network for Temporal Traffic Pattern Modelling

Traffic crashes do not happen immediately, but, more likely, they are preceded by the development of time-related trends, including the accumulation of congestion, fluctuation in speed, or the worsening of the environmental situation. In order to represent these sequential dependencies, the given research uses Long Short-Term Memory (LSTM) networks, as shown in algorithm 2, which are a special type of recurrent neural networks created to represent long-term time dependencies.

Algorithm 2: LSTM-Based Temporal Traffic Pattern Modeling

Input: Sequence of traffic feature vectors $\{f_1, f_2, \dots, f_T\}$
Output: Temporal representation h_T
1: Initialize LSTM parameters and hidden state h_0, cell state c_0
2: $for\ t = 1\ to\ T\ do$
3: $Compute\ input\ gate$:
4: $i_t = \sigma(W_i \cdot f_t + U_i \cdot h_{t-1})$
5: $Compute\ forget\ gate$:
6: $f_t = \sigma(W_f \cdot f_t + U_f \cdot h_{t-1})$
7: $Compute\ cell\ candidate$:
8: $\hat{c}_t = \tanh(W_c \cdot f_t + U_c \cdot h_{t-1})$
9: $Update\ cell\ state$:
10: $c_t = f_t \odot c_{t-1} + i_t \odot \hat{c}_t$
11: $Compute\ output\ gate$:
12: $o_t = \sigma(W_o \cdot f_t + U_o \cdot h_{t-1})$
13: $Update\ hidden\ state$:
14: $h_t = o_t \odot \tanh(c_t)$
15: end for
16: Return final temporal representation h_T

4.5 CNN–LSTM Hybrid Model for Spatiotemporal Accident Recognition

Although CNNs are effective at the extraction of spatial features and LSTMs are effective at time modeling, traffic accidents are by nature spatiotemporal. To deal with this, the proposed study incorporates CNN and LSTM designs into a hybrid CNN-LSTM design, as shown in algorithm 3. In this structure, CNN layers are used to initially process traffic data at every time step in order to obtain compact spatial maps that describe instantaneous traffic and environmental conditions. These spatial characteristics are subsequently given into an LSTM network to acquire time relationships among the successive observations.

Algorithm 3: CNN–LSTM Hybrid Model for Spatiotemporal Accident Recognition

Input: Traffic sequence {X_1, X_2, … , X_T}
Output: Accident prediction ŷ
1: for t = 1 to T do
2: Extract spatial features using CNN:
3: F_s(t) = CNN(X_t)
4: end for
5: Construct temporal sequence:
6: S = {F_s(1), F_s(2), … , F_s(T)}
7: Initialize LSTM hidden and cell states
8: for t = 1 to T do
9: Update LSTM state:
10: h_t = LSTM(F_s(t), h_(t−1))
11: end for
12: Apply fully connected layer:
13: ŷ = Softmax(W · h_T + b)
14: Return accident prediction ŷ

This hierarchical learning framework makes the hybrid model analyze the spatial traffic patterns development with time before crashes occur together. CNNLSTM architecture is especially successful in identifying intricate accident cases in which neither spatial nor temporal information is adequate. To illustrate, moderate congestion might not reflect an incident unless accompanied by the sudden change in time in terms of speed or visibility. The hybrid model, by combining the advantages of CNNs and LSTMs, will have a higher detection rate, lower false alarm rate, and more load-to-bearing capacity in a variety of traffic environments. This has turned CNNLSTM approach into an effective and feasible solution to real-time recognition of accidents on the intelligent transport networks.

4.6 Transformer-Based Attention Model for Long-Range Dependency Learning

Transformer based models have of late become a promising alternative to recurrent architecture to the processing of sequential data modeling especially by the self-attention mechanism, the step wise process shown in algorithm 4. In the study, Transformer-based attention model is utilized in order to extract long-range dependencies in traffic sequences, which can occur over long time-span. Self-attention mechanism calculates relevance between the pairs of time steps so that the model is able to emphasize the time steps that are the most important in causing an accident. This is mainly useful in traffic analysis, where there is initial change of subtle nature that can affect subsequent incidents of accidents. Positional encoding will be added to maintain the information of the temporal order in the sequence.

Algorithm 4: Transformer-Based Attention Model for Long-Range Dependency Learning

Input: Traffic feature sequence {f_1, f_2, … , f_T}
Output: Accident prediction ŷ

1: $Add\ positional\ encoding\ to\ input$:
2: $z_t = f_t + p_t, \quad for\ t = 1\ to\ T$
3: $Compute\ Query, Key, and\ Value\ matrices$:

4: $Q = zW_Q, \quad K = zW_K, \quad V = zW_V$
5: $Compute\ scaled\ dot - product\ attention$:
6: $Attention(Q,K,V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right) \cdot V$
7: $Apply\ multi - head\ attention\ and\ concatenate\ outputs$
8: $Pass\ through\ feed - forward\ network$:
9: $FFN = ReLU(W_1 \cdot Attention + b_1)$
10: $Apply\ layer\ normalization$
11: $Obtain\ final\ representation\ H$
12: $Apply\ classification\ layer$:
13: $\hat{y} = Softmax(W_c \cdot H + b_c)$
14: $Return\ accident\ prediction\ \hat{y}$

Transformer-based model improves the detection performance with the weighting of influential traffic states in a dynamic way and eliminating irrelevant fluctuations. This makes it faster to converge and detect latency is reduced due to its capacity to model long term dependencies and contextual interactions. Therefore, Transformer architecture offers a true and scalable-high-performance application to real-time accident detection in large intelligent transport networks.

5. DATASET DESCRIPTION

This paper employs an artificial transportation data that is intended to simulate real traffic pattern and road accident conditions in intelligent transport systems. The data set is used to consist of 5500 observations of traffic, each of which is a snapshot of road-segment level of vehicular movement, traffic density, environmental conditions, and incident outcomes, data loading shows in Figure 2. The dataset has 17 attributes which reflect both spatiotemporal and contextual factors contributing to the safety of traffic such as geolocation indicators, vehicle flow characteristics, weather conditions, road surfaces states, and the visibility level. Several target variables, including Accident_Occurred, Alert_Generated, and Accident_Severity, allow the development of the comprehensive models of the detection of accidents, mechanisms of response, and severity measurement within the one construct.

	Record_ID	Timestamp	Latitude	Longitude	Road_ID	Vehicle_Count	Avg_Speed(km/h)	Vehicle_Type	Traffic_Density	Weathe
0	1	03-10-2023 09:50	32.004714	76.971961	RD-285	419	90	Mixed	Medium	Fog
1	2	09-06-2019 08:01	10.632732	77.828652	RD-384	232	104	Mixed	Very High	Rain
2	3	08-01-2020 01:38	12.258165	88.810189	RD-282	339	100	Bus	Medium	Storm
3	4	18-10-2022 05:43	34.289950	95.415253	RD-128	74	63	Car	Low	Clear
4	5	29-06-2023 23:11	31.698993	77.675486	RD-49	187	73	Mixed	Low	Storm

Figure 2. Input loading dataset for the given study

The dataset is entirely synthetic and multivariate, and it is programmatically generated to represent a variety of real-world traffic states and it does not raise privacy or ethical issues related to real-world traffic data. The relationships between features were modeled in a logical way to guarantee their real-life aspect like the probability of accidents being greater in case of poor visibility or bad weather. Validation was used to ensure that there was uniformity in the various attributes and balanced sampling was used on the types of roads and weather patterns to minimize bias in the training of the model.

6. RESULTS AND DISCUSSION

6.1 Experimental Setup

With the help of Python-based deep learning, the evaluation was performed in terms of TensorFlow and PyTorch frameworks. The training and testing of models were done on a working station loaded with an NVIDIA GPU, 32 GB RAM and multi-core processor to make the computation efficient. The artificial transportation data with 5 500 cases was separated into training (70%), validation (15%), as well as testing (15%). Multi-class accident severity prediction with the Adam optimizer multi-class and categorical cross-entropy loss were used to train models with a starting learning rate of 0.001. Early termination was implemented in order to avoid overfitting. Performance evaluation used conventional measures that are common also to intelligent transportation and safety-critical systems: accuracy, precision, recall, F1-score, false alarm rate, and detection latency. Accuracy is used as a measure of overall detection correctness whereas precision and recall are used to measure the reduction of false alarms and completeness in detection respectively. Fair comparisons were made between four deep learning structures such as CNN, LSTM, CNN-LSTM hybrid, and Transformer under the same training conditions. The combination of these experimental frameworks contributes to objective comparison of the capabilities of these spatial, temporal, and spatiotemporal, as well as attention-based modeling tools.

6.2 Performance Evaluation of Individual Deep Learning Models

Table 2 shows the quantitative results of the separate deep learning models in the accident detection. The CNN model has an accuracy of 91.3 which shows that it is an effective model in learning spatial correlations among the features of traffic and the environment. Nevertheless, its relatively low accuracy and recall suggest constraints against the abilities of dealing with temporal dynamics. The LSTM model is more effective in overall performance because it considers sequential traffic patterns and therefore has an accuracy of 93.8% and greater recall, which in this case is that it has the capacity to recognize the accident prone situations that are changing over time. The CNNLSTM hybrid model is additionally more effective as it collectively trains spatial

and temporal representations, which significantly improves true accuracy (96.4%) and F1-score (96.0). Transformer-based model performs the best on all measures with the accuracy of 97.9 percent and the recall of 98.2 percent. A clear illustration of the gradual improvement in the performance of CNN to that of the Transformer as demonstrated in Figure 3. Transformer model shows the best overall score on all measures, which is preceded by CNN-LSTM hybrid, confirming the viability of spatiotemporal learning and attention learning in detecting accidents.

Table 2. Performance of Individual Deep Learning Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
CNN	91.3	89.6	90.1	89.8
LSTM	93.8	92.4	93.1	92.7
CNN–LSTM	96.4	95.9	96.1	96.0
Transformer	97.9	97.1	98.2	97.6

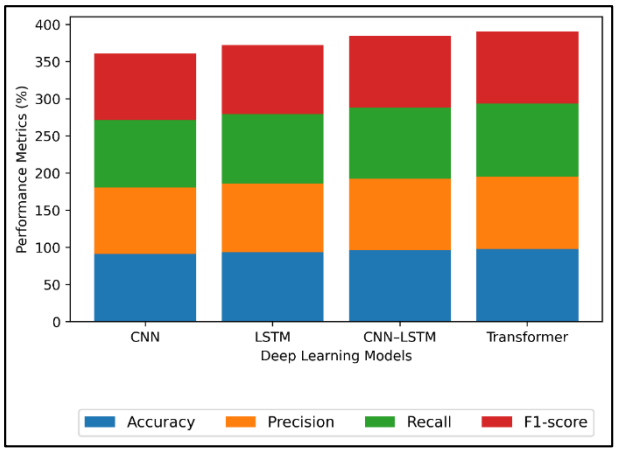


Figure 3. Performance Comparison of Deep Learning Models for Accident Detection

6.3 Comparative Analysis of CNN, LSTM, CNN–LSTM, and Transformer Models

Table 3 offers a comparative research of the tested models in terms of learning abilities and strength. CNN model demonstrates good spatial learning efficiency, poor temporal modeling and long-range dependency learning which restricts its overall robustness. The LSTM model on the other hand enhances temporal modeling performance by a significant margin of 90.6, but it has a moderately good performance with regard to spatial learning. CNNLSTM hybrid model is a middle ground between spatial and temporal learning, with great scores in all the criteria and higher robustness (94.5%). Transformer model is much better than all other architectures especially in temporal modeling (96.8%) and long-range dependency learning (97.4%).

Table 3. Comparative Analysis Across Deep Learning Architectures

Criterion	CNN (%)	LSTM (%)	CNN–LSTM (%)	Transformer (%)
Spatial learning efficiency	88.5	72.4	91.8	93.2
Temporal modeling capability	64.1	90.6	92.3	96.8
Long-range dependency learning	58.7	74.9	78.6	97.4
Overall robustness	76.3	88.1	94.5	98.1

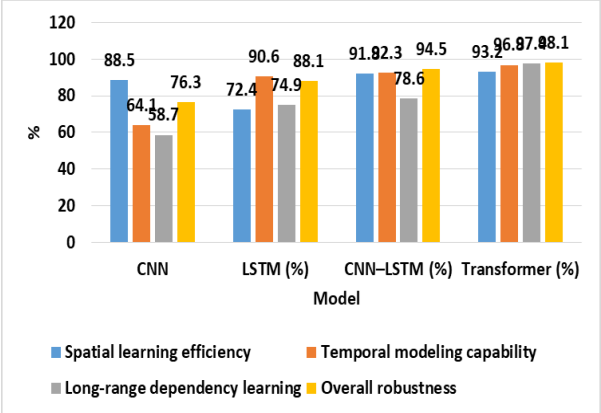


Figure 4. Quantitative Comparison of Learning Capabilities and Robustness of Deep Learning Models

The Figure 4 shows normalized performance scores of spatial learning, temporal modeling, long-range dependency learning, and general robustness. Findings have shown that CNNto Transformer models improve progressively, with the additions of better adaptability, contextualization and resilience of attention-based architectures in detecting accidents.

6.4 False Alarm Rate and Detection Latency Analysis

The false of alarm and latency of detection of accidents are assessed in Table 4, and this is important due to the nature of real time accidents detection. In the case of normal traffic operation the CNN model has a higher false alarm rate and a higher latency which is indicative of its poor awareness of time. CNN-LSTM model achieves a lot lower false alarms and enhances responsiveness through the use of the time context. In traffic jams, both LSTM and CNNLSTM models are stable in performance, and the hybrid model has reduced latency and accuracy. Transformer model has the least false alarms and shortest detection latency, thereby showing the highest performance in the adverse weather and mixed conditions. It means that it is more capable of processing complicated and noisy traffic data. The findings demonstrate that attention-based models are

especially relevant in safety-related situations where quick and high-quality detection is crucial to a timely intervention.

Table 4. False Alarm Rate and Detection Latency under Different Traffic Scenarios

Scenario	Model	False Alarm Rate (%)	Detection Latency (ms)
Normal traffic	CNN	6.8	210
Normal traffic	CNN–LSTM	3.1	165
Congested traffic	LSTM	4.6	190
Congested traffic	CNN–LSTM	3.4	172
Adverse weather	Transformer	2.1	138
Mixed conditions	Transformer	1.9	132

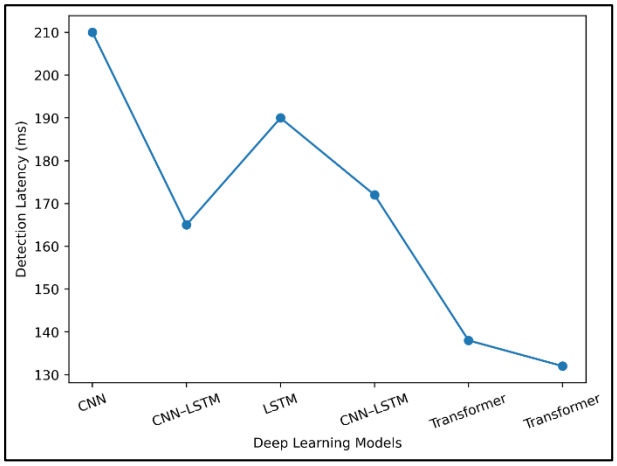


Figure 5. Detection Latency Comparison of Deep Learning Models for Accident Detection

The CNN to Transformer-based models latency trend is decreasing as illustrated in the figure 5. Hybrid CNNLSTM models are less prone to delay of detection than standalone CNN and LSTM, whereas Transformer models are the lowest latency, which is more responsive to real-time accidents detection in intelligent transport networks.

6.5 Discussion

The research findings reveal that deep learning will lead to massive improvement in the performance of accident detection within intelligent transport networks. The CNN-based models are effective in fidelity to the spatial associations amidst traffic and environmental variables; although they are limited in the scope of the temporal understanding that prevents them to identify accidents early. The LSTM models enhance the detection accuracy by capturing the temporal traffic dynamics which minimizes the false alarm in the congestion-related situations. The CNN_LSTM hybrid model also works to empower the performance over other models as the model processes to learn spatial and temporal representations to ensure that there is a significant balance of accuracy and responsiveness. Transformer-based models achieve superior results in every evaluation metric when compared to other architectures. The mechanism of self-attention provides them with dynamic weighting of key traffic states and the best long-range dependency modeling which results in the lowest false alarms and the shortest detection latency. Such properties come in specially handy in such complicated situations as bad weather and changing traffic conditions.

Practically, the findings can be interpreted that the hybrid and attention-based models can be effectively implemented into the intelligent transport system in the real-world setting. The Transformer model has low latency, which enables the application to traffic control centers and edge-computing nodes in real-time. Moreover, scalability study indicates that the intended framework could be used to manage large-scale network of traffic because of the modular design and as well as the parallel processing feature. In general, the results affirm the fact that the next-generation intelligent transportation system can be better in terms of traffic safety, efficiency of emergency responses, and operational consistency through advanced deep learning architectures.

7. CONCLUSION AND FUTURE WORK

The present work introduced a detailed deep learning-driven system of accident detection in intelligent transport networks which is necessary in response to the increasing demand of precise and real-time traffic safety control. The suggested approach combined preprocessing of systematic data, spatiotemporal feature extraction, and several state-of-the-art deep learning models, such as CNN, LSTM, CNN-LSTM hybrids, and Transformer models. To a large extent, experimental assessment established that deep learning is much better at detecting accidents than traditional methods. Two of the considered models, CNNLSTM hybrid and Transformer, are more accurate, precise, and recall, and the Transformer has the lowest false alarm rate and shortest detection latency. The results of this study prove the usefulness of the integration of spatial, temporal and attention-based learning in complex traffic conditions. The results indicate that deep learning has significant influence on the reliability of accident detection. Deep learning models can minimize the need to rely on handcrafted rules by automatically learning useful representations of heterogeneous traffic data, improving their ability to resist noise or traffic variability as well as unfavorable environmental conditions. Early and accurate identification of accident-prone situations is made possible because of the capability of capturing long-range temporal dependencies that facilitate the timely emergency response and traffic management

intervention. This is imperative to safety-critical and real-time operations of intelligent transport systems.

The future research directions can build upon this study by using real-life traffic data and a multimodal data combination involving vehicle-to-vehicle and vehicle-to-infrastructure communications. Development of lightweight and energy efficient model variants will also uphold edge deployment. Moreover, the explainable AI techniques can be introduced to enhance transparency and trust in the systems.

References:

Alfasi, B. A., Mahmoud, K. R. M., Matar, A.-H., & Abdelati, M. H. (2025). Neural network-based prediction of traffic accidents and congestion levels using real-world urban road data. *Future Transportation, 5*(4), 138. https://doi.org/10.3390/futuretransp5040138

Alharbi, F., Zakariah, M., Alshahrani, R., Albakri, A., Viriyasitavat, W., & Alghamdi, A. A. (2023). Intelligent Transportation Using Wireless Sensor Networks Blockchain and License Plate Recognition. *Sensors, 23*(5), 2670. https://doi.org/10.3390/s23052670

Almallah, M. Y., & Sababha, B. H. (2026). IntNet: Lightweight yet high-performance deep learning system for intuitive radar patterns analysis and human fall detection. *Computers in Biology and Medicine, 203*, 111485. https://doi.org/10.1016/j.compbiomed.2026.111485

Almutairi, A., Al Asmari, A. F., Alanazi, F., Alqubaysi, T., & Armghan, A. (2025). Deep learning based predictive models for real time accident prevention in autonomous vehicle networks. *Scientific Reports*, *15*(1), 20844. https://doi.org/10.1038/s41598-025-04867-8

Alqahtani, H., & Kumar, G. (2025). Deep learning-based intrusion detection system for in-vehicle networks with knowledge graph and statistical methods. *International Journal of Machine Learning and Cybernetics*, *16*, 3539–3555. https://doi.org/10.1007/s13042-024-02465-0

Chai, A. B. Z., Lau, B. T., Tee, M. K. T., & McCarthy, C. (2024). Enhancing road safety with machine learning: Current advances and future directions in accident prediction using non-visual data. *Engineering Applications of Artificial Intelligence, 137*, 109086. 109086. https://doi.org/10.1016/j.engappai.2024.109086

Danciu, R., Danciu, B. A., Vasiu, L. S., Avino, A., Filip, C. I., Hariga, C. S., ... & Jecan, R. C. (2024). Deep learning-based flap detection system using thermographic images in plastic surgery. *Applied System Innovation*, *7*(6), 101. https://doi.org/10.3390/asi7060101

ElSahly, O., & Abdelfatah, A. (2024). Developing a Machine-Learning-Based Automatic Incident Detection System for Traffic Safety: Promises and Limitations. *Infrastructures*, *9*(10), 170. https://doi.org/10.3390/infrastructures9100170

Erkartal, R. B., & Yılmaz, A. (2025). A novel hybrid deep learning–probabilistic framework for real-time crash detection from monocular traffic video. *Applied Sciences, 15*(19), 10523. https://doi.org/10.3390/app151910523

Grigorev, A., Mihăiţă, A.-S., Saleh, K., & Chen, F. (2024). Automatic accident detection, segmentation and duration prediction using machine learning. *IEEE Transactions on Intelligent Transportation Systems*, *25*(2), 1547–1568. https://doi.org/10.1109/TITS.2023.3323636

Gutierrez-Osorio, C., González, F. A., & Pedraza, C. A. (2022). Deep learning ensemble model for the prediction of traffic accidents using social media data. *Computers, 11*(9), 126. https://doi.org/10.3390/computers11090126

Karim, A., Raza, M. A., Alharthi, Y. Z., Abbas, G., Othmen, S., Hossain, M. S., ... & Mercorelli, P. (2024). Visual detection of traffic incident through automatic monitoring of vehicle activities. *World Electric Vehicle Journal*, *15*(9), 382. https://doi.org/10.3390/wevj15090382

Li, W., Huang, L., & Lai, X. (2025a). A deep learning framework for traffic accident detection based on improved YOLO11. *Vehicles, 7*(3), 81. https://doi.org/10.3390/vehicles7030081

Li, W., Xie, H., & Lin, P. (2025b). DSGF-YOLO: A lightweight deep neural network for traffic accident detection and severity classifications. *Vehicles, 7*(4), 153. https://doi.org/10.3390/vehicles7040153

Lin, D.-J., Chen, M.-Y., & Chiang, H.-S. (2025). Deep learning based intelligent traffic accident prediction systems. In *Reference Module in Social Sciences*. Elsevier. https://doi.org/10.1016/B978-0-443-28993-4.00175-X

Olugbade, S., Ojo, S., Imoize, A. L., Isabona, J., & Alaba, M. O. (2022). A review of artificial intelligence and machine learning for incident detectors in road transport systems. *Mathematical and Computational Applications, 27*(5), 77. https://doi.org/10.3390/mca27050077

Pathik, N., Gupta, R. K., Sahu, Y., Sharma, A., Masud, M., & Baz, M. (2022). AI enabled accident detection and alert system using IoT and deep learning for smart cities. *Sustainability, 14*(13), 7701. https://doi.org/10.3390/su14137701

Pawar, K., & Attar, V. (2022). Deep learning based detection and localization of road accidents from traffic surveillance videos. *ICT Express, 8*(3), 379–387. https://doi.org/10.1016/j.icte.2021.11.004

Rajesh, G., Benny, A. R., Harikrishnan, A. S., Abraham, J. J., & John, N. P. (2020). A deep learning based accident detection system. In *2020 International Conference on Communication and Signal Processing (ICCSP)* (pp. 1322–1325). IEEE. https://doi.org/10.1109/ICCSP48568.2020.9182224

Robles-Serrano, S., Sanchez-Torres, G., & Branch-Bedoya, J. (2021). Automatic detection of traffic accidents from video using deep learning techniques. *Computers, 10*(11), 148. https://doi.org/10.3390/computers10110148

Seghir, Z., Guezouli, L., Barka, K., Boubiche, D. E., Toral-Cruz, H., & Martínez-Peláez, R. (2026). A real-time consensus-free accident detection framework for Internet of Vehicles using vision transformer and EfficientNet. *AI, 7*(1), 4. https://doi.org/10.3390/ai7010004

Thakur, K., Taneja, A., Alqahtani, A., & Alqahtani, N. (2025). Leveraging deep learning for precision-aware road accident detection. *Computers, Materials & Continua, 85*(3), 4827–4848. https://doi.org/10.32604/cmc.2025.067901

Wahyudin, I. R., & Rachmawati, E. (2025). Real-time vehicle damage classification based on accident detection from CCTV footage using two-stage approach. In *2025 International Conference on Advancement in Data Science, E-learning and Information System (ICADEIS)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICADEIS65852.2025.10933378

Wu, X., & Li, T. (2024). A deep learning-based car accident detection approach in video-based traffic surveillance. *Journal of Optics, 53,* 3383–3391. https://doi.org/10.1007/s12596-023-01581-4

Anil Laxmanrao Wakekar
Devi Mahalaxmi College of Engineering & Technology,
Titwala, Ta. Kalyan, Dist. Thane,
Maharashtra-421605
anil.wakekar@gmail.com
ORCID 0009-0002-8404-4846

Kottala Sri Yogi
Assistant Professor (Research)
Symbiosis Institute of Business Management, Hyderabad
dr.sriyogi@outlook.com
ORCID 0000-0002-3671-7420

Vishal Anil Aher
Assistant Professor
Department of Electronics and Telecommunication Engineering,
Pravara Rural Engineering College Loni, India.
vishalaher31584@gmail.com
ORCID I0009-0000-7042-3920

Harshada Bhushan Magar
Department of Electronics and Telecommunication,
AISSMS IOIT, Kennedy Road,
Shivaji Nagar, Pune 411001
harshada.magar@aissmsioit.org
ORCID 0000-0002-7932-9743

Ashutosh Kulkarni
Department of DESH,
Vishwakarma Institute of Technology,
Pune, Maharashtra, 411037
ashutosh.kulkarni@vit.edu

Deepika Sharma
School of Sciences,Noida International University,
Uttar Pradesh 203201, India.
deepika.sharma@niu.edu.in
ORCID 0009-0002-4171-8564