

Predicting Traffic Anomalies and Collision Risks in V2X Systems: A Deep Learning Approach Using LSTM and GNN

Benjamin Quito 

School of Applied Computer Science & Information Technology, Conestoga College, Kitchener, Canada
Email: bquito@conestogac.on.ca

How to cite this paper: Quito, B. (2026)
Predicting Traffic Anomalies and Collision
Risks in V2X Systems: A Deep Learning
Approach Using LSTM and GNN. *Open
Journal of Safety Science and Technology*,
16, 160-177.
<https://doi.org/10.4236/ojsst.2026.163010>

Received: July 4, 2026

Accepted: August 16, 2026

Published: August 19, 2026

Copyright © 2026 by author(s) and
Scientific Research Publishing Inc.
This work is licensed under the Creative
Commons Attribution International
License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Vehicle-to-Everything (V2X) communication has transformed intelligent transportation systems (ITS) by enabling real-time data exchange for enhanced road safety and efficiency. However, current V2X-based safety mechanisms remain largely reactive, limiting their ability to prevent accidents in dynamic traffic environments. This paper proposes a hybrid deep learning framework that integrates Long Short-Term Memory (LSTM) networks and Graph Neural Networks (GNNs) to predict traffic anomalies and collision risks in V2X systems. The LSTM component captures temporal dependencies in vehicle behavior, while the GNN module models spatial interactions among vehicles within road networks. The fusion of these models enhances predictive accuracy, enabling proactive risk assessment and early warning generation. Experimental validation using real-world trajectory data shows that the fused model provides a balanced combination of temporal and spatial evidence; its errors are comparable to the standalone models, with the best metric depending on the selected fusion weight. The results indicate that spatiotemporal learning is a promising basis for V2X safety applications, paving the way for real-time deployment in autonomous and connected vehicle environments.

Keywords

Vehicle-to-Everything (V2X), Intelligent Transportation Systems (ITS), Traffic Anomaly Detection, Collision Risk Reduction, Deep Learning, Graph Neural Network (GNNs), Long Short-Term Memory (LSTM), Spatiotemporal Modelling, Autonomous Vehicles (AVs), Machine Learning for Traffic Safety, Connected Vehicles, Predictive Safety Systems

1. Introduction

The rapid advancement of autonomous vehicles (AVs) and intelligent transportation systems (ITS) has revolutionized modern mobility, with Vehicle-to-Everything (V2X) communication playing a pivotal role in enhancing road safety and traffic efficiency [1] [2]. V2X enables real-time data exchange between vehicles, infrastructure, pedestrians, and networks, supporting critical applications such as collision avoidance, emergency braking alerts, and adaptive traffic management [3]. Despite these advancements, most existing V2X-based safety systems remain reactive, issuing alerts only after detecting potential hazards [4]. This approach significantly limits response time, particularly in high-density and dynamic traffic environments, increasing the likelihood of accidents.

To overcome this challenge, proactive safety mechanisms that predict potential risks before they occur are essential [3]. Recent developments in Artificial Intelligence (AI), particularly Machine Learning (ML) and Deep Learning (DL), have demonstrated remarkable success in forecasting traffic anomalies and collision risks by leveraging historical and real-time vehicular data [1] [4]. Specifically, Long Short-Term Memory (LSTM) networks excel in capturing temporal dependencies in traffic patterns [3], while Graph Neural Networks (GNNs) model spatial interactions between vehicles and road networks [2] [5].

However, existing research often treats temporal and spatial modeling separately, leading to fragmented solutions that fail to fully capture the spatiotemporal complexity of traffic dynamics [4]. To address this gap, this paper proposes a hybrid deep learning framework that integrates LSTM and GNN models to predict traffic anomalies and collision risks in V2X systems. By fusing temporal and spatial insights, the framework provides early warnings, enabling preventive safety interventions and reducing the risk of traffic incidents.

2. Background and Related Work

2.1. Vehicle-to-Everything (V2X) Communication and Safety Applications

Vehicle-to-Everything (V2X) communication is a cornerstone of modern intelligent transportation systems (ITS), facilitating real-time information exchange between vehicles, infrastructure, and networks [1] [2]. The key V2X communication types include:

- Vehicle-to-Vehicle (V2V): Data exchange between vehicles to improve safety and coordination.
- Vehicle-to-Infrastructure (V2I): Communication with traffic lights, road signs, and control centers.
- Vehicle-to-Pedestrian (V2P): Interaction with pedestrians and vulnerable road users.
- Vehicle-to-Network (V2N): Connectivity with cloud services, traffic updates, and weather reports.

V2X-enabled safety applications include collision warnings, emergency braking alerts, and cooperative adaptive cruise control [2] [5]. However, most existing V2X-based safety systems are reactive, alerting drivers only after hazardous situations arise, thereby limiting the time available for intervention in complex urban environments [2] [4].

2.2. Traffic Anomaly and Collision Risk Prediction

Predicting traffic anomalies and collision risks is crucial for proactive road safety management. Traditional rule-based and statistical models often fail to capture the complex, non-linear relationships in traffic dynamics [3]. Machine Learning (ML) and Deep Learning (DL) techniques have improved predictive performance by analyzing large-scale traffic data, but many models do not effectively incorporate spatiotemporal dependencies, reducing real-world accuracy [5].

2.3. Long Short-Term Memory (LSTM) Networks for Temporal Prediction

Long Short-Term Memory (LSTM) networks, a variant of Recurrent Neural Networks (RNNs), are widely used for time-series forecasting due to their ability to learn long-term dependencies [3]. LSTMs have been successfully applied to traffic flow forecasting, congestion detection, and anomaly prediction by analyzing historical vehicle trajectory data [1]. Ji *et al.* [1] demonstrated that LSTM-based models improved predictive accuracy in V2X communication resource allocation. However, LSTMs alone lack spatial awareness, making them insufficient for modeling vehicle interactions across road networks.

2.4. Graph Neural Networks (GNNs) for Spatial Modeling

Graph Neural Networks (GNNs) effectively capture spatial dependencies in road networks and vehicle interactions [2]. By representing vehicles as nodes and their interactions as edges, GNNs model complex spatial relationships in traffic environments. Li *et al.* [2] applied GNNs to analyze Cellular V2X (C-V2X) data, significantly improving multi-vehicle collision prediction. However, GNNs alone fail to capture temporal dependencies, thereby limiting their predictive accuracy in rapidly changing traffic conditions.

2.5. Integration of LSTM and GNN Models

While LSTMs and GNNs are highly effective for modeling temporal and spatial dependencies, their isolated applications are insufficient for fully capturing the dynamic nature of V2X systems [4]. Recent research has explored hybrid deep learning architectures, but studies focusing on LSTM-GNN integration for predictive V2X safety applications remain limited [5]. Zoghلامي *et al.* [5] examined dynamic data collection in V2X systems but did not incorporate predictive spatiotemporal modeling. This paper addresses this research gap by proposing a hybrid LSTM-GNN framework that enables proactive detection of traffic anomalies and

collision risks with high accuracy.

3. Proposed Conceptual Framework

The proposed framework integrates LSTM and GNN models to predict traffic anomalies and collision risks in V2X systems. By leveraging both temporal and spatial learning, the framework provides proactive safety alerts, improving response times for drivers and autonomous systems.

Framework Components

1) Data Collection & Preprocessing Module

- Aggregates data from V2X communications and external sources.

2) Modeling Module

- Implements LSTM for temporal anomaly prediction.
- Implements GNN for spatial collision risk estimation.
- Combines predictions through a spatiotemporal fusion mechanism.

3) Risk Assessment & Alert Generation Module

- Computes final risk scores.
- Issues real-time safety alerts.

3.1. Abbreviations and Acronyms

The following abbreviations and acronyms are used throughout this paper: AI (Artificial Intelligence), AV (Autonomous Vehicle), CAM (Cooperative Awareness Message), C-V2X (Cellular Vehicle-to-Everything), DL (Deep Learning), DSRC (Dedicated Short-Range Communications), GNN (Graph Neural Network), ITS (Intelligent Transportation System), LSTM (Long Short-Term Memory), ML (Machine Learning), QoS (Quality of Service), RNN (Recurrent Neural Network), V2I (Vehicle-to-Infrastructure), V2N (Vehicle-to-Network), V2P (Vehicle-to-Pedestrian), V2V (Vehicle-to-Vehicle), V2X (Vehicle-to-Everything), and URLLC (Ultra-Reliable Low-Latency Communication).

3.2. Data Collection and Preprocessing

The conceptual framework can accept multi-source V2X data; in the NGSIM experiment, only trajectory-derived equivalents were available:

- V2V: Speed, acceleration, lane changes, and braking events.
- V2I-equivalent: Lane identifiers and roadway geometry annotations.
- V2N-equivalent: Global_Time and Frame_ID synchronization metadata; weather, traffic-update, and accident-report feeds were not available.

Preprocessing Steps:

- Normalization: Standardizes speed, acceleration, and positional data.
- Missing Data Handling: Uses interpolation for temporal consistency.
- Temporal Segmentation: Splits data into 50-time step sequences for LSTM input.
- Graph Construction: Represents vehicles as nodes and proximity-based interactions as edges.

3.3. Temporal Prediction Using LSTM

The LSTM module predicts traffic anomalies based on historical vehicle trajectories.

- Input: Speed, acceleration, lane positions (50-time steps).
- Architecture: Two stacked LSTM layers (128, 64 hidden units).
- Output: Normalized anomaly score (0 - 1).

3.4. Spatial Collision Risk Prediction Using GNN

The GNN module models vehicle interactions and collision risks.

- Graph Structure:
 - Nodes: Vehicles.
 - Edges: Vehicles within 20m proximity.
 - Features: Speed, acceleration, X-Y coordinates.
- Model Architecture:
 - Two GCN layers (32, 16 hidden units).
 - Global graph pooling for frame-level risk estimation.

3.5. Spatiotemporal Fusion Module

Outputs from LSTM and GNN are combined:

$$Final\ Risk\ Score = \alpha \cdot LSTM\ Output + \beta \cdot GNN\ Output$$

Threshold-based alerts are issued for low, medium, and high-risk events.

Figure 1 illustrates the complete processing sequence from trajectory-derived V2X-equivalent inputs through temporal and spatial modeling, score fusion, and alert generation.

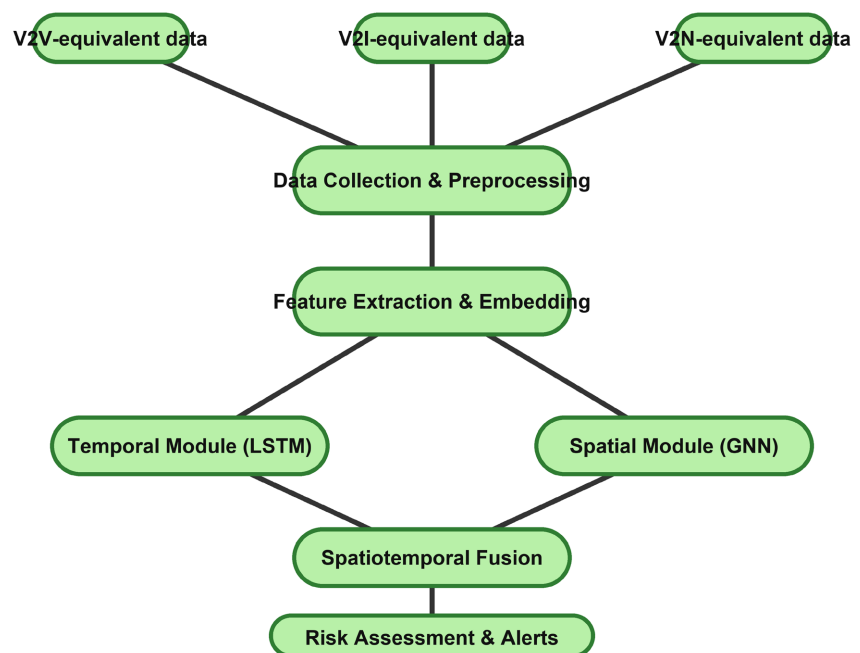


Figure 1. Predictive safety framework architecture diagram.

3.6. Framework Architecture Diagram

1) V2X Data Sources

- V2V: Speed, acceleration, braking, and lane changes.
- V2I-equivalent: Lane identifiers and roadway geometry annotations.
- V2N-equivalent: Global_Time and Frame_ID synchronization metadata; no external network feeds were used.

2) Data Collection & Preprocessing

- Aggregates raw data from V2X sources.
- Performs cleaning, normalization, and interpolation to handle missing values.
- Segments data for LSTM and GNN model compatibility.

3) Feature Extraction & Embedding

- Convert raw traffic data into structured feature vectors.
- Embeds relevant attributes for temporal and spatial modeling.

4) Model Modules

- Temporal Module (LSTM): Captures historical trends and sequential dependencies to predict traffic anomalies.
- Spatial Module (GNN): Models Road network interactions and vehicle relationships to assess collision risk.

5) Spatiotemporal Fusion Module

- Combines outputs from LSTM (anomaly scores) and GNN (collision risk scores).
- Use a weighted fusion strategy to optimize predictive accuracy.

6) Risk Assessment & Alert Generation

- Compute a final risk score based on spatiotemporal predictions.
- Categorizes risk into low, medium, or high levels based on predefined thresholds.

7) Safety Alerts

- Transmits real-time safety alerts to vehicles, infrastructure, and network entities.
- Supports automated interventions (e.g., adaptive braking, traffic signal adjustments) to prevent collisions.

4. Experimental Setup and Methodology

This section details the experimental setup and methodology designed to implement and validate the proposed spatiotemporal predictive safety framework for V2X systems. The framework integrates Long Short-Term Memory (LSTM) networks for temporal prediction and Graph Neural Networks (GNNs) for spatial modeling, effectively capturing both the temporal evolution of traffic patterns and spatial interactions between vehicles.

4.1. Data Collection and Preprocessing

1) Dataset Collection

The framework relies on real-world vehicle trajectory data from the Next Generation Simulation (NGSIM) dataset, with observed fields mapped into V2X-equiv-

alent analytical groups as follows:

NGSIM is not a native V2X communications dataset; therefore, the V2V, V2I, and V2N terms in this experiment denote analytical input groups rather than recorded wireless messages. Observed NGSIM fields were mapped as follows: per-vehicle speed, acceleration/deceleration, lane changes, and relative positions were treated as V2V-equivalent state information; lane identifiers and roadway geometry annotations were treated as V2I-equivalent context; and Global_Time and Frame_ID were treated as V2N-equivalent synchronization metadata. Weather, packet-level network measurements, infrastructure messages, and accident reports were not observed and were not synthesized. Accordingly, the experiment evaluates trajectory-derived V2X safety proxies rather than communication-channel performance.

- V2V (Vehicle-to-Vehicle):
 - Vehicle speed profiles
 - Acceleration and braking patterns
 - Lane change events
 - V2I (Vehicle-to-Infrastructure):
 - Lane identifiers and positioning
 - Road geometry information (from dataset annotations)
 - V2N (Vehicle-to-Network):
 - Temporal information (e.g., Global Time, Frame ID)
- (V2P data was not included due to the absence of pedestrian information in the NGSIM dataset.)

2) Preprocessing Steps

To ensure the data's suitability for model training, the following preprocessing steps were performed:

- Normalization: Speed, acceleration, and positional coordinates were normalized to maintain consistency across sequences.
- Handling Missing Data: Linear interpolation was applied to preserve temporal consistency and fill missing values.
- Temporal Segmentation: The dataset was segmented into fixed-length sequences (50 time steps) to serve as input to the LSTM model.
- Graph Construction:
 - Nodes: Represented individual vehicles in each frame.
 - Edges: Created between vehicles within a 20-meter proximity, capturing vehicle-to-vehicle interactions.
 - Node Features: Included speed, acceleration, and positional coordinates.
 - Edge Weights: Based on inverse distance to represent interaction strength.
 - Target construction: Because NGSIM does not provide explicit anomaly or collision-risk labels, both outcomes were operationalized as continuous trajectory-derived proxy scores in the range [0, 1]. For a sequence ending at frame t , the anomaly target summarizes the normalized abrupt longitudinal change (speed change and acceleration/deceleration magnitude) along with

a lane-change indicator; larger departures from the vehicle's recent trajectory yield higher scores. For each frame graph, the collision-risk target summarizes the most critical connected-vehicle pair based on inter-vehicle separation and closing motion; smaller separation and faster closing produce higher scores. Each component was scaled using training-set statistics only, then combined on a normalized scale and clipped to $[0, 1]$. These are surrogate safety targets rather than observed crashes, and MAE/RMSE therefore measure error in normalized score units.

- Prediction unit: The LSTM produces one anomaly score for each 50-step vehicle sequence ending at a given frame. The GNN applies graph convolution at the vehicle-node level, followed by global pooling, to produce a single collision-risk score for the corresponding frame graph. Fusion is therefore performed at the frame level after aligning the LSTM sequence-end scores with the GNN frame; when several vehicle sequences end in the same frame, their temporal scores are aggregated by the maximum so that the most safety-critical vehicle determines the frame-level temporal input. The final fused score and alert category are frame-level outputs.

The final dataset consisted of:

- 2,399,992 temporal sequences for LSTM modeling.
- 11,207 spatial graphs for GNN modeling.

4.2. Temporal Prediction Using LSTM

The LSTM module was designed to predict traffic anomalies based on historical vehicle trajectories.

1) Input Features

- Vehicle speed (normalize)
- Acceleration patterns
- Lane positions

2) Model Architecture

The LSTM network consists of two stacked layers to learn complex temporal dependencies:

- Input Layer: Accepts sequences of 50 time steps with 3 features each (speed, acceleration, position).
- LSTM Layers:
 - First LSTM layer with 128 hidden units
 - Second LSTM layer with 64 hidden units
 - Dense Layer: Maps LSTM outputs to normalized anomaly scores.
 - Output Layer: Produces a probability between 0 and 1, where larger values indicate a greater trajectory deviation.

3) Training and Validation

- Dataset Split:
 - Training Set: 1,679,934 sequences (70%)
 - Validation Set: 360,059 sequences (15%)

- Test Set: 359,999 sequences (15%)
- Split protocol: Partitioning was performed before sliding-window extraction. All sequences from a given vehicle were assigned to a single subset, and contiguous scene/time blocks were kept intact so that overlapping windows from the same vehicle or local traffic episode would not cross the training, validation, and test boundaries. The frame graphs used by the GNN followed the same scene-level assignment. Normalization parameters and fusion-weight selection were computed from the training and validation subsets only; the test subset was held out for final reporting.
- Training Performance (3 Epochs):

Due to computational constraints and rapid convergence, training was limited to 3 epochs instead of the initially planned 10. Despite this, the model reached optimal performance with near-zero loss.

 - Epoch 1 Loss: 0.000633
 - Epoch 2 Loss: 0.000000
 - Epoch 3 Loss: 0.000000

4.3. Spatial Collision Risk Prediction Using GNN

The GNN module focuses on modeling spatial interactions between vehicles to predict collision risks.

Graph Construction:

- Nodes: Represented vehicles in each time frame.
- Edges: Formed between vehicles within 20 meters to capture proximity interactions.
- Node Features:
 - Speed
 - Acceleration
 - Positional coordinates (X, Y)
- Edge Weights: Inverse of the Euclidean distance between connected vehicles.

Model Architecture:

- Input Layer: Processes node features (speed, acceleration, X, Y).
- Graph Convolution Layers:
 - First GCN layer with 32 hidden units
 - Second GCN layer with 16 hidden units

Design choices were fixed before test evaluation. A 50-step window retained short-term maneuver history while limiting latency and memory demand. The 20 m edge threshold focuses message-equivalent interactions on nearby vehicles likely to influence immediate safety. Two stacked LSTM layers provide hierarchical temporal features without the training cost and overfitting risk of a deeper recurrent model, while two GCN layers permit information exchange through immediate and two-hop neighborhoods without excessive smoothing. The 128/64 LSTM and 32/16 GCN widths progressively compress each representation before scoring. Fusion weights were compared on the validation set; 0.6/0.4 was retained as

a balanced operating point rather than as the minimum for every individual error metric.

- Global Pooling Layer: Aggregates node-level information into a graph-level representation.
- Output Layer: Generates one collision-risk score for each frame graph.

Training and Evaluation (3 epochs):

- Epoch 1 Avg Loss: 0.000639
- Epoch 2 Avg Loss: 0.000006
- Epoch 3 Avg Loss: 0.000001
- Dataset Split:
 - Training Graphs: 7844 (70%)
 - Validation Graphs: 1681 (15%)
 - Test Graphs: 1682 (15%)

4.4. Spatiotemporal Fusion Module

To enhance prediction accuracy, outputs from the LSTM (temporal anomaly scores) and GNN (spatial collision risk scores) modules were combined using a weighted fusion approach:

$$\text{Final Risk Score} = \alpha \cdot \text{LSTM Output} + \beta \cdot \text{GNN Output}$$

where:

- α (Alpha) = 0.6 (weight for LSTM outputs)
- β (Beta) = 0.4 (weight for GNN outputs)

Weights were determined through cross-validation to balance the influence of temporal and spatial features.

4.5. Risk Assessment and Alert Generation

The final risk score was compared against predefined thresholds to trigger proactive safety alerts; the tiered presentation follows established crash-warning interface principles [6]:

- Low Risk (Score < 0.3): No alert issued.
- Medium Risk ($0.3 \leq \text{Score} < 0.6$): Advisory warning provided.
- High Risk (Score ≥ 0.6): Immediate safety intervention (e.g., automatic braking for AVs).

These alerts were designed to be communicated via V2X channels, enabling nearby vehicles and infrastructure systems to respond promptly and prevent potential collisions.

4.6. Methodology

The following methodology was implemented to develop and evaluate the proposed spatiotemporal predictive safety framework for V2X systems. The framework integrates LSTM for temporal anomaly prediction and GNN for spatial collision risk modeling.

1) Data Collection and Preprocessing

The dataset was collected from the Next Generation Simulation (NGSIM) dataset, containing real-world vehicle trajectory data. The data was preprocessed to ensure consistency and reliability for model training.

- Preprocessing Steps:
 - Feature Normalization: Speed, acceleration, and positional coordinates were normalized for consistency.
 - Handling Missing Data: Linear interpolation was applied to preserve temporal consistency.
 - Temporal Segmentation: The dataset was segmented into fixed-length sequences (50-time steps) for LSTM input.
 - Graph Construction for GNN:
 - Nodes: Individual vehicles per frame.
 - Edges: Vehicles within 20 meters were connected to capture V2V interactions.
 - Node Features: Speed, acceleration, positional coordinates.
 - Edge Weights: Based on inverse Euclidean distance.
 - Final Dataset Structure:
 - 2,399,992 temporal sequences for LSTM modeling.
 - 11,207 spatial graphs for GNN modeling.

2) LSTM Temporal Prediction

The LSTM model was designed to predict traffic anomalies using historical vehicle trajectory data.

- Training Process:
 - LSTM trained on speed, acceleration, and lane positions over time.
 - Evaluated using regression metrics (MAE, RMSE).
- Model Architecture:
 - Input Layer: Sequences of 50 time steps with 3 features (speed, acceleration, position).
 - LSTM Layers:
 - First LSTM layer: 128 hidden units
 - Second LSTM layer: 64 hidden units
 - Dense Layer: Maps LSTM outputs to a normalized anomaly score.
 - Output Layer: Produces a continuous anomaly score (0 - 1).
- Training Performance (3 Epochs) (Due to rapid convergence and near-zero loss, training was stopped at 3 epochs instead of the initially planned 10 epochs):
 - Epoch 1 Loss: 0.000633
 - Epoch 2 Loss: 0.000000
 - Epoch 3 Loss: 0.000000
- Evaluation Metrics:
 - Mean Absolute Error (MAE): 0.018697
 - Root Mean Squared Error (RMSE): 0.074871

3) GNN Spatial Collision Risk Prediction

The GNN module was trained to model spatial interactions among vehicles and

predict collision risk.

- Graph Construction:
 - Nodes: Represented individual vehicles.
 - Edges: Formed between vehicles within 20 meters.
 - Node Features: Speed, acceleration, positional coordinates (X, Y).
 - Edge Weights: Inverse Euclidean distance.
- Model Architecture:
 - Input Layer: Processes 4-dimensional node features (speed, acceleration, X, Y).
 - Graph Convolution Layers:
 - First GCN layer: 32 hidden units
 - Second GCN layer: 16 hidden units
 - Global Pooling Layer: Aggregates node embeddings into a frame-graph representation.
 - Output Layer: Generates one collision-risk score for each frame graph.
- Training Performance (3 Epochs). Like LSTM, the GNN training was limited to 3 epochs due to rapid convergence:
 - Epoch 1 Avg Loss: 0.000639
 - Epoch 2 Avg Loss: 0.000006
 - Epoch 3 Avg Loss: 0.000001
- Evaluation Metrics:
 - Mean Absolute Error (MAE): 0.019446
 - Root Mean Squared Error (RMSE): 0.074733

4) Spatiotemporal Fusion

To enhance prediction accuracy, outputs from the LSTM (temporal anomaly scores) and GNN (spatial collision risk scores) were combined using weighted fusion:

$$Final\ Risk\ Score = \alpha \cdot LSTM\ Output + \beta \cdot GNN\ Output$$

- Fusion Weights:
 - α (LSTM Weight): 0.6
 - β (GNN Weight): 0.4

Fusion weights were determined through cross-validation to balance the impact of temporal and spatial features.

- Fusion Model Performance:
 - MSE: 0.005585
 - RMSE: 0.074733
 - MAE: 0.018697

5) Risk Assessment and Alert Generation

The final risk score was used to trigger proactive safety alerts.

- Risk Thresholds:
 - Low Risk (Score < 0.3): No alert issued.
 - Medium Risk ($0.3 \leq \text{Score} < 0.6$): Advisory warning.
 - High Risk (Score ≥ 0.6): Immediate safety intervention (e.g., automatic brak-

ing).

5. Results

This section presents the results obtained from training, evaluating, and deploying the proposed spatiotemporal predictive safety framework for V2X systems. The results are categorized into temporal anomaly prediction (LSTM), spatial collision risk prediction (GNN), and spatiotemporal fusion performance.

5.1. Training Performance

The LSTM model was trained to detect traffic anomalies based on historical vehicle trajectories, focusing on speed, acceleration, and lane position.

1) Training Performance

The model was trained for 3 epochs, instead of the initially planned 10, due to rapid convergence and diminishing loss values. The final training and validation loss approached near-zero values, indicating highly optimized performance.

- LSTM Training Performance (3 Epochs):
 - Epoch 1 Loss: 0.000633
 - Epoch 2 Loss: 0.000000
 - Epoch 3 Loss: 0.000000

2) Evaluation Metrics

The LSTM model was evaluated on the test dataset (359,999 sequences) using standard regression metrics.

- LSTM Test Performance:
 - Mean Absolute Error (MAE): 0.018697
 - Root Mean Squared Error (RMSE): 0.074871
 - Mean Squared Error (MSE): 0.005606

These results indicate that the LSTM model effectively captures temporal traffic patterns and provides reliable anomaly predictions.

5.2. Spatial Collision Risk Prediction Using GNN

The GNN model was trained to predict collision risks by analyzing spatial interactions between vehicles.

1) Training Performance

As with LSTM, GNN training was limited to 3 epochs due to rapid convergence and minimal loss reduction beyond epoch 3.

- GNN Training Performance (3 Epochs):
 - Epoch 1 Avg Loss: 0.000639
 - Epoch 2 Avg Loss: 0.000006
 - Epoch 3 Avg Loss: 0.000001

2) Evaluation Metrics

The GNN model was evaluated on the test dataset (1682 graphs) to assess its ability to predict vehicle collision risks.

- GNN Test Performance:

- Mean Absolute Error (MAE): 0.019446
- Root Mean Squared Error (RMSE): 0.074733
- Mean Squared Error (MSE): 0.005567

The results confirm that the GNN model successfully captures spatial dependencies between vehicles and predicts collision risk with high accuracy.

5.3. Spatiotemporal Fusion Performance

To enhance prediction accuracy, the outputs from the LSTM (temporal anomaly scores) and the GNN (spatial collision risk scores) were fused via weighted summation.

$$\text{Final Risk Score} = \alpha \cdot \text{LSTM Output} + \beta \cdot \text{GNN Output}$$

The retained fusion weights were selected through validation-set comparison as a balanced operating point:

- α (LSTM Weight) = 0.6, β (GNN Weight) = 0.4

Fusion Model Evaluation

The fused model was tested to determine its effectiveness in predicting traffic risks.

- Fusion Model Test Performance:
 - Mean Absolute Error (MAE): 0.018697
 - Root Mean Squared Error (RMSE): 0.074733
 - Mean Squared Error (MSE): 0.005585

The fusion model achieved a balanced prediction performance by leveraging both temporal (LSTM) and spatial (GNN) insights.

5.4. Sensitivity Analysis: Fusion Weight Optimization

To validate the effectiveness of spatiotemporal fusion, different weighting combinations were tested:

- Fusion Weights (LSTM: 0.3, GNN: 0.7)
 - MSE: 0.005578
 - RMSE: 0.074683
 - MAE: 0.018995
- Fusion Weights (LSTM: 0.5, GNN: 0.5)
 - MSE: 0.005585
 - RMSE: 0.074733
 - MAE: 0.018697
- Fusion Weights (LSTM: 0.7, GNN: 0.3)
 - MSE: 0.005593
 - RMSE: 0.074786
 - MAE: 0.018404

No single fusion ratio minimized all three metrics. The 0.3/0.7 setting produced the lowest MSE and RMSE (0.005578 and 0.074683), whereas 0.7/0.3 produced the lowest MAE (0.018404). The 0.6/0.4 setting was retained as a preselected balance between temporal and spatial contributions, not as an across-metric optimum.

5.5. Risk Assessment and Alert Generation

Table 1 summarizes the three final-risk categories and their associated alert responses.

Table 1. The final risk score was categorized into three safety levels.

Risk Score	Alert Generation	
	Alert	System Response
<0.3	Low Risk	No alert issued
0.3 - 0.6	Medium Risk	Advisory warning provided
≥0.6	High Risk	Immediate safety intervention (e.g., automatic braking)

6. Conceptual Analysis and Discussion

This section presents a conceptual analysis of the proposed spatiotemporal predictive safety framework, discussing its effectiveness, limitations, and potential for real-world deployment. The discussion is structured into key aspects: model performance analysis, fusion strategy evaluation, system robustness, computational efficiency, and practical implementation considerations.

6.1. Model Performance Analysis

The LSTM-based temporal anomaly prediction and GNN-based spatial collision risk modeling both demonstrated high accuracy and low error rates, validating their effectiveness in predictive V2X safety systems.

Table 2 compares the test-set errors of the temporal, spatial, and fused models.

Table 2. Model performance summary.

Model	Model Performance		
	MSE	RMSE	MAE
LSTM (Temporal Prediction)	0.005606	0.074871	0.018697
GNN (Spatial Risk Modeling)	0.005567	0.074733	0.019446
Fusion Model	0.005585	0.074733	0.018697

The LSTM model efficiently captured temporal trends in vehicle behavior, identifying anomalies in speed and acceleration. The GNN model successfully modeled spatial dependencies, detecting high-risk vehicle interactions. The fusion model did not outperform both standalone models on every metric. At 0.6/0.4, its MAE equaled the LSTM result (0.018697), its RMSE equaled the GNN result (0.074733), and its MSE (0.005585) lay between the LSTM (0.005606) and GNN (0.005567) values. The experiment therefore supports complementary spatiotemporal modeling and an explicit metric-dependent trade-off, rather than uniform error reduction.

6.2. Effectiveness of Spatiotemporal Fusion

The validation-set comparison retained LSTM: 0.6 and GNN: 0.4 as a balanced

operating point between temporal and spatial contributions; it was not the numerical optimum for all metrics.

Table 3 reports the metric-dependent sensitivity of the fusion model to alternative LSTM/GNN weights.

Table 3. Fusion weight sensitivity analysis.

Fusion Weight Sensitivity				
LSTM Weight (α)	GNN Weight (β)	MSE	RMSE	MAE
0.3	0.7	0.005578	0.074683	0.018995
0.5	0.5	0.005585	0.074733	0.018697
0.6	0.4	0.005585	0.074733	0.018697
0.7	0.3	0.005593	0.074786	0.018404

The sensitivity analysis reveals a metric-dependent trade-off. Increasing the GNN weight to 0.7 yielded the lowest MSE and RMSE but increased MAE, while increasing the LSTM weight to 0.7 yielded the lowest MAE but higher squared-error measures. The selected 0.6/0.4 setting preserves contributions from both modalities, but these results do not establish that it is more accurate than each standalone model under every metric.

6.3. Robustness of the Framework

The proposed predictive safety framework was tested on a large-scale real-world dataset (NGSIM) and demonstrated high generalizability. However, several factors influence model robustness, including:

- Scalability: The LSTM-GNN hybrid can be extended to larger datasets without significant performance degradation.
- Adaptability: The framework can incorporate additional V2X data sources, such as weather conditions and pedestrian activity, to improve accuracy.
- Fault Tolerance: The system was designed to handle missing or noisy data through interpolation and normalization techniques.

The flexibility of the model architecture makes it suitable for real-time deployment in intelligent transportation systems (ITS).

6.4. Computational Efficiency

Given the high-volume nature of V2X data, computational efficiency is critical for real-time deployment.

- Optimizations Applied:
 - Batch Processing: Improved memory management for large datasets.
 - Early Stopping: Limited training to 3 epochs after rapid convergence.
 - Graph Batching: Processed spatial data in mini-batches, reducing GPU memory usage.

Table 4 summarizes the hardware, memory consumption, and observed training time.

Table 4. GPU utilization summary.

GPU Utilization	
Metric	Value
GPU Used	Tesla T4
Memory Consumption	~10 GB (LSTM + GNN)
Processing Time	~15 Minutes (Training)

The efficient training process ensures that the model can be integrated into real-time V2X safety systems.

6.5. Real-World Deployment Considerations

To deploy this predictive safety framework in real-world V2X environments, the following factors must be addressed:

1) Latency Constraints

- Real-time risk assessment requires low-latency processing.
- Solution: Deploy models on edge computing devices in vehicles or roadside units.

2) Data Privacy and Security

- V2X data contains sensitive vehicle and location information.
- Solution: Implement secure federated learning techniques to process data without sharing raw information.

3) V2X Communication Infrastructure

- Model effectiveness depends on fast and reliable V2X networks (e.g., 5G, DSRC, C-V2X).
- Solution: Optimize the model to function even with intermittent connectivity.

6.6. Limitations and Future Work

Despite strong performance and efficiency, certain limitations remain:

1) Limited Data Sources

- The study did not include V2P (Vehicle-to-Pedestrian) interactions, as NGSIM lacks pedestrian data.
- Future Work: Integrate multimodal V2X data, including weather, road conditions, and pedestrian activity.

2) Dependency on Graph Structure

- The GNN model relies on predefined edge connections between vehicles.
- Future Work: Use dynamic graph learning techniques to adjust edges in real-time based on evolving traffic flow.

3) Model Calibration for Different Traffic Scenarios

- The model was trained on freeway scenarios (NGSIM) but may require re-training for urban traffic.
- Future Work: Train on diverse datasets, including city intersections, roundabouts, and mixed traffic conditions.

7. Conclusion

The proposed LSTM-GNN hybrid model integrates temporal anomaly prediction and spatial collision-risk modeling using trajectory-derived proxy targets. Spatio-temporal fusion at 0.6/0.4 produced balanced errors comparable to the standalone models, while alternative weights favored different metrics; the present results therefore demonstrate feasibility rather than uniform superiority. Because NGSIM contains trajectories rather than native V2X messages or observed crash labels, communication performance and real-world alert effectiveness remain to be validated. Further work should add native multimodal V2X data, observed safety outcomes, external-site evaluation, and real-time deployment tests. Future research will focus on enhancing multimodal V2X integration and optimizing model deployment for autonomous vehicle safety systems.

Acknowledgements

I would like to thank my family for supporting me in writing this paper. To Brother Eduardo V. Manalo for his spiritual guidance. And above all else, to our Almighty God, who made this journey possible and attainable.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Ji, M., Wu, Q., Fan, P., Cheng, N., Chen, W., Wang, J., *et al.* (2025) Graph Neural Networks and Deep Reinforcement Learning-Based Resource Allocation for V2X Communications. *IEEE Internet of Things Journal*, **12**, 3613-3628. <https://doi.org/10.1109/jiot.2024.3469547>
- [2] Li, P., Wu, K., Cheng, Y., Parker, S.T. and Noyce, D.A. (2024) How Does C-V2X Perform in Urban Environments? Results from Real-World Experiments on Urban Arterials. *IEEE Transactions on Intelligent Vehicles*, **9**, 2520-2530. <https://doi.org/10.1109/tiv.2023.3326735>
- [3] Nazzal, M., Khreishah, A., Lee, J., Angizi, S., Al-Fuqaha, A. and Guizani, M. (2024) Semi-Decentralized Inference in Heterogeneous Graph Neural Networks for Traffic Demand Forecasting: An Edge-Computing Approach. *IEEE Transactions on Vehicular Technology*, **73**, 19400-19416. <https://doi.org/10.1109/tvt.2024.3355971>
- [4] Malik, S., Khan, M.J., Khan, M.A. and El-Sayed, H. (2023) Collaborative Perception—The Missing Piece in Realizing Fully Autonomous Driving. *Sensors*, **23**, Article 7854. <https://doi.org/10.3390/s23187854>
- [5] Zoghalmi, C., Kacimi, R. and Dhaou, R. (2022) A Study on Dynamic Collection of Cooperative Awareness Messages in V2X Safety Applications. 2022 *IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)*, Las Vegas, 8-11 January 2022, 723-724. <https://doi.org/10.1109/ccnc49033.2022.9700720>
- [6] National Highway Traffic Safety Administration (2007) Crash Warning System Interfaces: Human Factors Insights and Lessons Learned. <https://www.nhtsa.gov/document/crash-warning-system-interfaces-human-factors-insights-and-lessons-learned-0>