scientific reports

Check for updates

OPEN

Multi-stage classification of abnormal traffic events using a multi-head + LSTM

Pratik Jadhav[1,2], Abderrahim Benslimane[3], Deepali R. Vora[4✉] & Shruti Patil[5]

Traffic congestion, anomalies and incident significantly impact urban transportation efficiency and road safety. Accurate detection and classification of such events are crucial for effective traffic managements, emergency response and infrastructure planning. Traditional approaches based statistical and conventional machine learning models often struggle to generalize across the dynamic and complex traffic patterns which evolves around the time. To address these limitations, we proposed a multi-head + LSTM model in a multistage classification framework. This proposed framework systematically detects anomalies using isolation forest, classifies the congestion into low, medium, high using K-means clustering and determine whether an incident caused an anomaly using a spatial threshold-based approach (1.5 km). the model is trained on 15 days of PeMS traffic data integrated with weather information to enhanced predictive accuracy. Through hierarchical classification the proposed model captures temporal dependencies, integrates contextual weather information and ensures robust anomaly detection, congestion classification and incident identification. Experimental results demonstrates that the multi-Head model significantly outperforms existing methods achieving higher precision, recall, f1-score and ROC-AUC across all classification stages. The results highlight the potential of deep learning-based traffic analysis for intelligent transportation system (ITS) enabling data-driven decision making for urban traffic management.

Keywords Traffic flow prediction, Congestion classification, Anomaly detection, Incident identification, Multi-head attention, LSTM, Intelligent transportation systems (ITS)

Traffic congestion, anomalies and incidents are critical factors which highly influence the urban traffic transportation efficiency and road safety[1]. Identifying such events like congestions, incident caused anomalies accurately is essential for traffic management, emergency response and infrastructure planning[2]. Traditional methods for traffic anomaly detection and congestion analysis often rely on the statistical models or conventional machine learning techniques which may struggle to generalize across complex and dynamic traffic pattern[3]. Therefore, deep learning-based model often given promising results which enhance the accuracy and efficiency of the traffic events classification models[4].

In this study we developed a multistage deep learning base classification framework to systematically detect the anomalies, classify the congestion levels for the detected anomalies and incident classification which determine whether an incident has caused anomaly or not. Our approach is built on the PEMS California traffic dataset, for our work we have considered 15 days of data and integrate it with externa weather and incident data. The classification is structured into three main stages:

- *Anomaly detection*: Identifying the irregular traffic patterns using isolation forest algorithm and based on it labelling each data point into anomaly (Yes/No)
- *Congestion classification*: Categorizing Road congestion into three classes Low, Medium and High using k-means clustering based on the traffic flow characteristics.

[1]Artificial Intelligence and Machine Learning Department, Symbiosis Institute of Technology, Symbiosis International (Deemed) University, Pune 412115, India. [2]The Computer Science Education and Research Center (CERI), University of Avignon, Avignon, France. [3]Computer Science and Engineering, University of Avignon, Avignon, France. [4]Symbiosis Institute of Technology Pune, Symbiosis International (Deemed University), Pune, India. [5]Symbiosis Artificial Intelligence Institute, Symbiosis Center for Applied Artificial Intelligence , Symbiosis International (Deemed University), Pune, India. ✉email: deepali.vora@sitpune.edu.in

- *Incident classification*: Determining whether an incident triggered an anomaly or not by using a threshold-based selection criterion where we have taken into account 1.5 km spatial threshold associating incident with nearby anomalies.

To achieve accurate results for classification task, we have built a multi-head LSTM-inspired model trained in a multistage format where in each stage builds upon previous prediction, allowing for structure and hierarchical classification process. The proposed model integrates past and contextual information in a timely manner for effective classification of traffic anomalies and congestion conditions.

Related work
Anomaly detection in traffic networks

Anomaly detection is road networks is a crucial task in intelligent transportation system (ITS). Anomaly detection is mainly used to identify irregular patterns in the traffic networks such as congestion anomalies, sensor error or accidents[5]. Traditional anomaly detection methods heavily rely on the statists technique and rule-based systems but with the advance machine learning and deep learning-based methods more advance models have emerged which can analyses the large-scale traffic data in real time[6]. This approach leverages structured traffic data, computer vision techniques and time series analysis to enhance anomaly identification accuracy.

The growing availability of traffic sensor data, CCTV footage and vehicular communication networks (VANET's) which has led to good development of novel data-driven anomaly detection models[7]. These models mainly focus on detecting the sudden deviations in the traffic patterns using clustering techniques, neural networks and hybrid deep learning architectures[8]. In recent studies unsupervised and semi-supervise learnings techniques have gained prominence due to their ability to generalize across different traffic conditions without any extensive labelled datasets[9]

Djenouri et al.[10] proposes a deep learning-based frameworks for detecting anomalies in urban traffic networks. Their method employs a Convolutional Neural Networks (CNN) with decomposition strategy to partition traffic flow data into clusters with similar patterns by training CNN separately on each cluster model is able to achieve good results for detecting anomalies by focusing on the local patterns rather that global patterns in traffic trends. Due to the decomposition-based approach model outperforms traditional anomaly detection methods particularly in large scale urban environments. The challenges this method faces are decompositions process which increases computational overhead and due to that method struggles with generalization when traffic patterns change dynamically.

Lei et al.[11] propose a spectral decomposition-based anomaly detection model for time series traffic data. Their method decomposes traffic data into the categories trends, seasonal trends and residual components using discrete Fourier transform (DFT) before implementing LSTM and CNN based predictive models. By reconstructing the original time series and identifying deviations this model has effectively shown good results for detecting anomalies in real time traffic monitoring systems. The use of unsupervised learning enables the model to generalize across different urban roads networks. Due to the reliance on frequency domain decomposition makes model sensitive to high frequency noise and due to its unsupervised nature model can lead to false positive in low variance traffic conditions.

Aboah et al.[12] proposes a vision-based anomaly detection systems which leverages deep learning for real time analysis on the traffic camera footage. aboah approach combines YOLOv5 object detection with decision tree which is a machine learning classification model to detect and classify anomalies in traffic videos. The proposed systems estimate the road background; extract foreground objects apply a decision tree-based filtering mechanism to identify anomalies such as stalled vehicles or accidents. Their method has given effective results for detecting anomalies in real time traffic monitoring applications. Even though model gives effective results but it heavily depends on the video quality and camera angles making it prone to failure in poor visibility conditions like (rain, fog, snow, nigh time scenarios).

Laanaoui et al.[13] presents an intelligent traffic anomaly detection system which is designed for real time traffic load balancing. The study mainly utilizes vehicular and hoc networks (VANET's) and real time sensor data to identify the anomalies and adjust vehicles routes dynamically. By implementing the big data analytics and machine learning their system optimizes traffic flow which helps in reducing congestion and incident related disruptions. The model demonstrates high precision in detecting traffic anomalies while maintaining low latency which make it suitable for deployment in smart cities application. As their approach reliance in VANET's due to that model poorly performs in the areas where there are low congestions limiting its effectiveness in rural or in less technological advance regions.

Traffic congestion in traffic networks

Urban mobility is perhaps one of the most pressing problems of traffic congestion that, besides increasing travel time, fuel consumption, and environmental pollution, also contributes to many road accidents[14]. The classification of congestion levels plays an important role in helping traffic management to mitigate its adverse effects[15]. In traditional congestion classification method is mainly relies on the rule-based systems and statical modelling which often fail to generalize across diverse traffic conditions[16]. In recent advancements in machine learning and deep learning have enable more robust, scalable and accurate classification approaches which leverages real-time traffic data from sensors, GPS and vehicular communications networks.

Modern congestion classification techniques utilize traffic flow, speed and density meters to categorize congestion levels[17]. These techniques can be broadly categorized into vision-based methods, sensor-based classification and ML driven approaches. However, challenges such as data heterogeneity, real-time scalability and the need for extensive labelled datasets remain significant hurdles.

Bereczki and Simon et al.[18] has used a hybrid ML-based congestion detection systems that compares both supervise and unsupervised learning models. The system process and compares both V2X enable vehicles and infrastructure sensors enabling the real time congestion detection. The proposed work consists use of multiple ML models such as Linear regression, Random Forest and Extreme Gradient Boosting to identify congestion pattern with high accuracy additional their model also forecast congestions in upcoming road segments which make suitable for smart city applications as model relies mainly on V2X communications model struggle in the areas where there is low vehicle connectivity and data sparsity in less densely populated area.

Shah et al.[19] propose ML driven congestion classification for vehicular ad-hoc Networks (VANET's). Their study utilizes traffic flow and speed data to estimate vehicular density and classifies the traffic into three states free-flow, Dense and Congested. The system evaluates multiple ML classification model such as Decision Tree, K-Nearest Neighbor, Random Forest and Multilayer perceptron with this ensemble soft-voting classifier demonstrate superior performance across multiple congestion scenarios. The spatial and temporal variation in vehicular density affects model accuracy and their systems requires a fine tuning for different urban and highway environments.

Liu et al.[20] proposes a graph-based congestion classification framework that models spatial–temporal dependencies in traffic networks. Their approach employs Spatial–Temporal Variational Graph Auto-Encoders (ST-VGAE) which helps to extract road networks embedding and classify the congestions levels. By implementing historic accidents and traffic flow data the system given a holistic view of congestion patterns. The proposed work has some limitations as computational complexity of graph-based model is high and model struggle to give good results with incomplete or missing sensor data.

Anomaly classification in traffic networks

Traffic incident classification is a vital component of intelligent transportations systems (ITS) enabling real-time detection of accidents, vehicle breakdown and other disruptions that impact traffic flow[21]. Effective classification of the incidents improves traffic managements, emergency response and congestion mitigation. Traditional incident classification relies on statistical methods and threshold-based selection rule which often result in high false alarm rate and limited adaptability in recent advancements in machine learning and deep learning have enable more accurate and scalable approaches leveraging sensors data time series analysis and deep neural networks[22].

Lu et al.[23] propose a hybrid traffic incidents detection model combining generative advisal networks (GAN's) and transformers to address dataset imbalance and improve detection accuracy. GANs generate synthetic incident samples to balance datasets, while transformers capture temporal dependencies for enhanced classification. Their proposed work reliance on GAN-generated synthetic data which introduces biases and transformers model which requires high computations power limiting their real time application.

Xie et al.[24] propose an ensemble learning based incident detection model using Random Forest and Random Subspace K-Nearest neighbors (RSKNN). The study involves using of SASYNO oversampling to mitigate dataset imbalance and enhance classification reliability. Model heavily depends on the features selection techniques and performance of the model varies on the road network type which requires specific dataset tunning.

Gurusamy et al.[25] developed an LSTM-SE based incident detection classification framework that leverages stacked autoencoders (SAE) for the feature extraction and LSTM for capturing temporal patterns the model demonstrates superior accuracy in detecting incidents from real time traffic data. Computational overhead of LSTM makes it difficult to deploy in low latency environments and therefore model struggle with highly dynamic traffic conditions.

Zhu et al.[26] proposes a network lasso based decentralized traffic incident detections frameworks which optimizes classification models in a distributed traffic monitoring system. This approach reduces network bandwidth usage and improves detection across multiple traffic regions the complex optimization process leads to longer training times and model performance is highly dependent on the quality of decentralized data.

Methodology

This section outlines the methodology implemented in the study with respect to data collection process, labelling strategies and model framework for traffic anomaly detection, congestion classification and incident identification. Our approach integrates multi-source traffic data, machine learning-based labelling and deep learning classification to create a robust and scalable classification system.

Data collection and pre-processing

We have used the PEMS California dataset covering January 1–15, 2024, which consists of station data, station meta data, incident reports and weather data. Station data consists of timestamp traffic observation such as flow, occupancy and speed where station meta data consists of geographical details like latitude, longitude and freeway information. Incident data record traffic incident data such as incident ID, timestamp, location and incident type like (e.g. Accident, breakdown or roadwork). We have also integrated weather data obtained via API include temperature, wind speed, precipitations and humidity which gives environmental context for traffic analysis. These datasets are merged using spatiotemporal alignment forming a compressive foundation for anomaly detection, congestion classification and incident identification for our multi-stage classification framework.

Figure 1 shows outlier detection and removal using IQR method to detect and remove the outliers in latitude and longitude values.

To ensure data consistency, redundant columns such as lane-specific statistics were removed and missing values in traffic flow, occupancy and speed were filed using station level and global mean imputation. The timestamp column was converted to datetime format like day, hour and minute these features were extracted
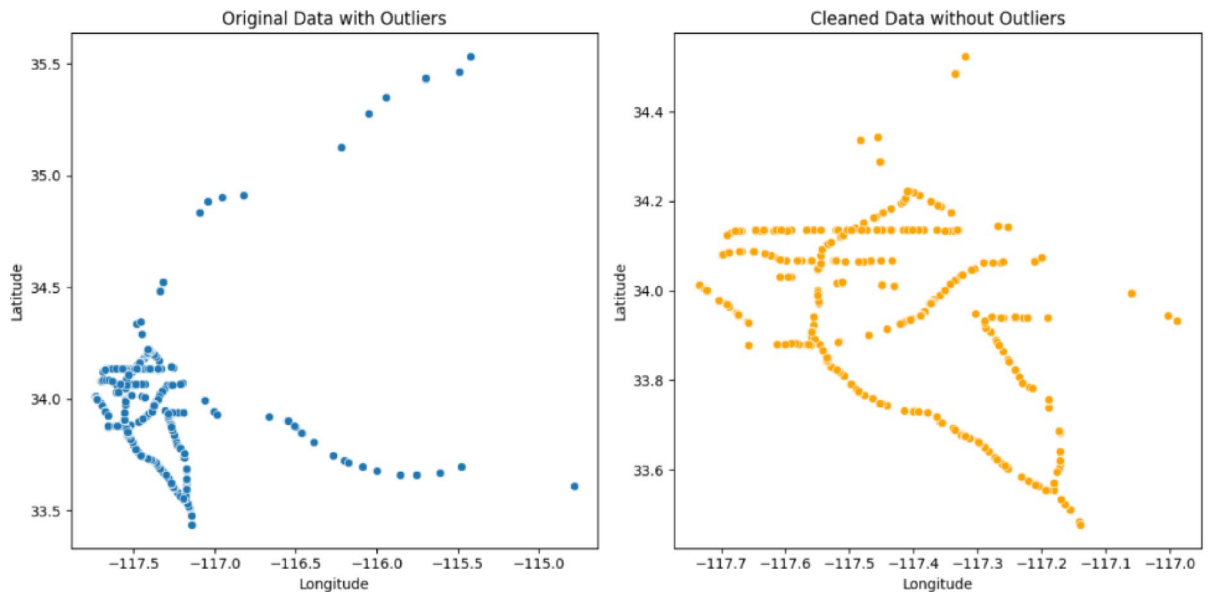
Fig. 1. Station data cleaning using IQR.

with sin cosine transformation applied to capture the cyclical time patterns. Latitude and longitude outliers were filtered as shown in Fig. 1 using IQR method and station metadata was merged to provide geospatial context.

Historic weather data (temperature, wind speed, precipitation and humidity) was retrieved via API and aligned with the traffic records based on timestamp and location. Categorical variables like travel direction and lane type were encode and numerical features were standardize using z-score normalization.in last sliding windows (steps = 6) were applied to capture spatial relationships, optimizing the dataset for multistage classification.

To address the issue of the class imbalance SMOTE (synthetic minority over-sampling technique) was applied on anomaly and incident class, ensuring sufficient representation of the minority class. Unlike naive up sampling, SMOTE generates synthetic samples based on the features space similarities rather than random duplication, preserving the temporal structure of the data. For congestion classification no up sampling was performed as class distribution was balanced additionally model evaluation was conducted on the unseen csv file ensuring no data leakage while up sampling which maintaining the integrity of time-series dependencies during testing.

Labelling strategy

- Anomaly detection labelling strategy

Anomalies in traffic data were labelled using isolation forest algorithm with a 1% contamination rate which means 1% of the data was classified as an anomaly. This threshold was selected through iterative testing and statistical validation using methods like IQR and Z-score analysis to ensure anomalies significantly deviated from normal traffic patterns. The model's predictions were further cross checked with historical reports and visual inspection to enhance reliability. Anomalies were labelled as 1 and 0 anomaly and no anomaly. This approach effectively captures irregularities and detects accurate congestion spikes, sudden speed drops and other unexpected behaviors crucial for smart traffic management. Methods like IQR/Z-score and One-Class SVM where not considered over Isolation Forest as these techniques assumes normality in data and struggles with scalability whereas isolation forest effectively detects anomalies in high dimension like real world traffic data without requiring labelled samples.

- Congestion classification labelling strategy

Traffic congestion was labelled using means clustering, an unsupervised machine learning approach that segmented the data into three cluster corresponding to Low, Medium, High congestions levels. The clustering was performed using the key features like traffic flow, average occupancy and average speed which were normalized before applying to the model. Once the cluster were formed their centroids were analyzed to derive thresholds-based rule for classification. Then based on this threshold the congestion classes were labelled this labelled congestion levels provide valuable insights for traffic predictions and management, enabling better decision making for congestion mitigation road planning. K-means is computationally efficient. Interpretable and naturally identifies distinct congestion patterns as compare to other methods like DBSCAN which is sensitive to noise and GMM (Gaussian Mixture Models) which assumes distribution.

- Incident classification labelling strategy

Incident labels were assigned by matching reported incidents with nearby traffic station which lies within the threshold of 1.5 km radius using spatial querying with cKD Tree. If an incident occurred within this threshold and aligned with the stations timestamp it was labelled as an incident (anomaly caused by an incident) if the incident id does not fall within threshold, then it was marked as no incident (anomaly was not caused by an incident) this method ensures accurate association between traffic disruptions and detected anomalies allowing for better incident classification and analysis. cKD tree provides fast and precise nearest neighbor searches for accurate incident-station joining as compare to other traditional method like brute force distance calculation are computationally expensive.

Although fixed thresholding (e.g. contamination, 1.5 km radius) were validated for PeMS data, their transferability to other networks may be limited. Future extension will consider adaptive or learning based thresholding strategies to enhance generalization.

The sequential three-stage designed reflects the natural progression of traffic events anomalies are first detected as irregularities, then categorized into congestion levels, and finally linked to nearby incidents. This ordering simplifies classification by filtering out normal traffic early, thereby reducing unnecessary computation.

Multi-stage model

- Architecture of the hybrid multistage model

Figure 2 shows the multi-head with LSTM model pipeline which is a multistage traffic classification, progressing through anomaly detection, congestion classification and incident identification. The pipeline begins with data pre-processing where from multiple data sources like station data, station meta data, weather data from API were integrated and transformed using feature engineering techniques including temporal encoding and spatial embedding. As data was not labelled, we have used different labelling strategies to label the data for anomaly we have used solation forest, for congestion classification label were generated based on the k-means threshold value and incident labels were decided based on the spatial distance threshold. The multi-stage + LSTM model first determines whether an anomaly is present in the traffic flow or not if yes then model classifies the anomaly into the congestion class and finally the model predicts weather a congestion was caused by incident or not. Multi-
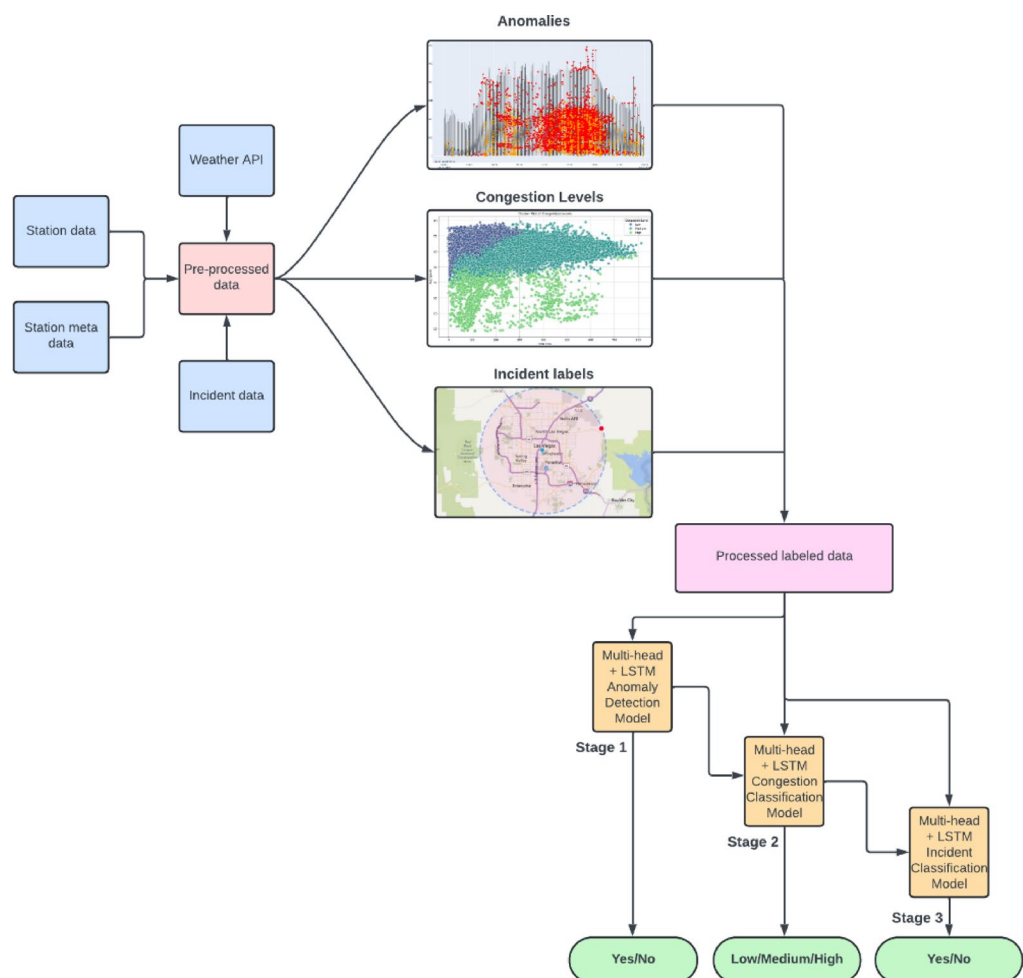

Fig. 2. Proposed multi-head with LSTM model pipeline architecture.

head attention captures complex dependencies in sequential data while LSTM layers are enhanced temporal feature extraction ensuring robust classification of each stage with improved contextual awareness.

This integration differs from prior hybrid approaches (e.g. CNN + LSTM or Transformers-based models) as it explicitly balances long term dependencies through multi-head attention and short- term temporal dynamics via LSTM. Embedding this within a staged classification pipeline (anomaly—congestion—incident) constitutes the novelty of our approach.

Unlike prior single-task models, our innovation lies in integrating anomaly detection, congestion classification and incident identification into sequential multi-stage pipeline, enriched with weather context and incremental training which provides a unique contribution to ITS research.

The multi-head attention with LSTM architecture shown in Fig. 3 combines a self-attention mechanism with LSTM-based sequential modelling for enhanced feature extraction in traffic network classification. First the model takes sequential input data which in (6-time step, 31 features per step) and passes to further through multi-head attention block with 128-dimension representation which help to capture a complex dependency across different time steps. The normalization and feed forward dense layer which are present from it the data is been pass further which refines feature space which are followed by dropout layer to prevent overfitting. A stack of LSTM units then processes the refined features which leverages temporal dependencies to enhance the classification accuracy. Finally, the global average pooling reduces the dimensionality before passing the output through fully connected layers which culminates it in a single unit output layer with an activation function for anomaly and incident the output neuron is sigmoid activation function with binary cross entropy loss and for congestion it SoftMax activation function with sparse categorical cross entropy. Muti-head attention with LSTM able to capture both short term dependencies and long-term dependencies where short term are captured by the LSTM and long terms is captured by multi-head attention. The multistage approach ensures stepwise, interpretable prediction process which helps to improve accuracy by sequentially refining anomaly, congestion and incident classification other models like CNN where it lacks temporal awareness while in other hand model like transformers which requires excessive data and computational power due to which proposed model gives best balance of efficiency and performance.

- Model training

The training process is multi-stage starts with anomaly detection then congestion classification and then incident identification to refine predictions step by step. Class imbalance is handled through SMOTE up sampling, ensuring balance of training samples without any data leakage by reserving unseen csv file for testing. The model is trained on 15 days csv and then tested on the 16th csv from the dataset which was unseen to the model. The model training uses the Adam optimizer with a learning rate of 0.001 and early stopping has been applied to.

prevent overfitting. For evaluation metrics such as precision, recall and f1-score were logged during training. Each stage refines its prediction before passing enriched feature representation to the next which enables a hierarchical and robust classification framework.

Experimental steup

- Hardware and software environment

The model was trained on a GPU P100 compute units to handle large scale traffic dataset efficiently. The implementation was done using TensorFlow and Kera's for deep learning with an additional tool like Scikit-learn, NumPy, pandas and optuna for data pre-processing, feature engineering and hyperparameter tuning.
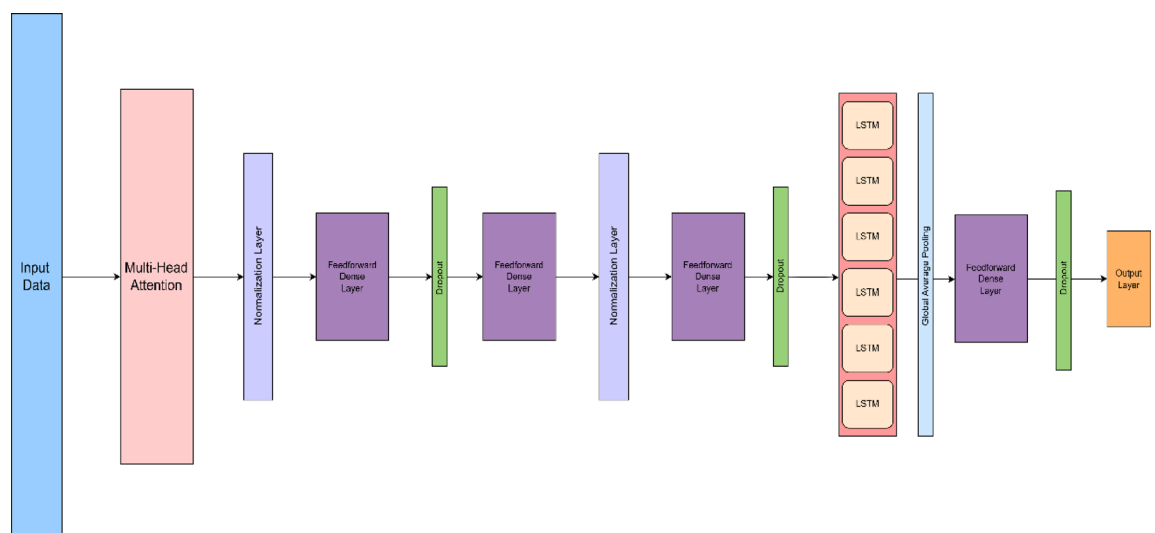

Fig. 3. Multi-head attention with LSTM model architecture.

- Performance metrics

The optimal hyperparameters values were considered using a Bayesian Optimization using optuna. To evaluate the model performances, we have considered confusion matrix, Precision, Recall, F1-Score and Matthews Correlation Coefficient (MCC) where True Positive (TP) represents correctly detected anomalies, congestion levels or incident where True Negative (TN) are correctly identified normal conditions. False Positives (FP) indicates incorrect detection where no issue exits and False Negatives (FN) occur when actual anomalies, congestion or incidents are missed by the model.

(1) Precision

$$\textbf{Precision} = \frac{TP}{TP + FP} \tag{1}$$

Precision (Eq. 1) measures the predictions which are crucial for minimizing false alarms in incident detection

(2) Recall (sensitivity)

$$\textbf{Recall} = \frac{TP}{TP + FN} \tag{2}$$

Recall (Eq. 2) captures how well the model has detected actual incidents or congestion preventing missed events.

(3) F1-score

$$\textbf{F1} - \textbf{Score} = 2 \times \frac{Precision * Recall}{Precision + Recall} \tag{3}$$

F1-Score (Eq. 3) Calculates the harmonic mean of precision and recall ensuring balancing between them.

(4) Matthews correlation coefficient (MCC):

$$\textbf{MCC} = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{4}$$

MCC (Eq. 4) is used for evaluating the imbalance class present in the dataset which gives a robust measure of classification performance.

Results and discussion
Model was trained on the 15 days csv traffic data and was tested on the 16th day csv below are the results and the comparison with existing work are discussed with respective to each objective of the multi-stage model.

Anomaly detection results
Figure 4b shows result of the anomaly detection model in form of confusion matrix which represent the prediction of the samples which correctly predicted vs incorrectly predicted. Figure 4a represents the roc curve of true positive rate vs false positive rate with an AUC score.

The anomaly detection model has achieved high performance with an accuracy score of 99.46% with an precision of 99.98 for normal traffic and 0.74 for anomalies and an F1-score of 0.81 for anomalies. The Matthews correlation coefficient of 0.81 indicates strong predictive performance. The ROC curve further confirms with an AUC score of 0.98 which highlights the model abilities to distinguish between normal and anomalous traffic effectively. The confusion matrix shown in fig a that the model correctly identified 508,711 normal instance and 6027 anomalies with a minimal false positive and negatives.

The above plots are the training logs as the model was trained on multiple days csv file each red horizontal dotted line represent new csv given to the model incremental learning manner. Figure 4c shows the training and validation accuracy trends over epochs. Figure 4d shows model training and validation loss across epochs. Figure 4e demonstrates how precision and recall evolved throughout training. Figure 4f tracks the F1-score progression across training epochs. The training logs illustrate the model learnings progression showing stable accuracy, loss reduction and strong trends of precision, recall and f1-score across epochs. These shows model results are well generalized where anomaly detection model effectively captures traffic patterns.

Model comparison with existing work
Table 1 shows the comparison with the existing work for anomaly detection, proposed multi-stage with LSTM model shows good results which outperform existing approaches across all the key metrics. Unlike existing works rely on CNN, YOLOs or machine learning based systems, the multi-stage with LSTM model effectively captures temporal dependencies and integrates external weather factors achieving the highest precision (99.53%), recall (99.46%) and F1_score (99.48). Additionally, AUC-ROC (99.08) and true negative rate (0.9959) demonstrates strong anomaly detection capabilities, surpassing traditional models that lack a multistage sequential learning approach.
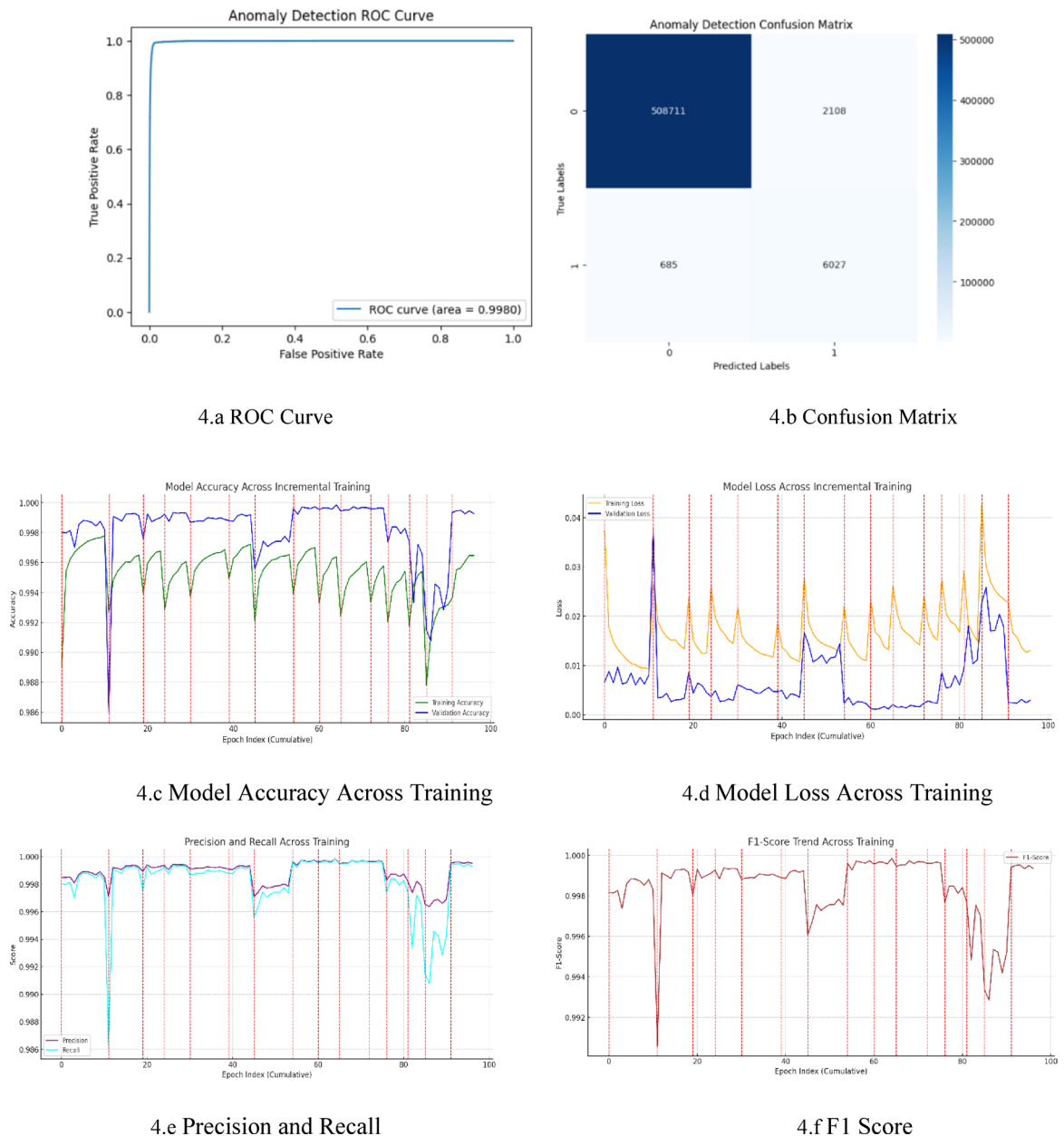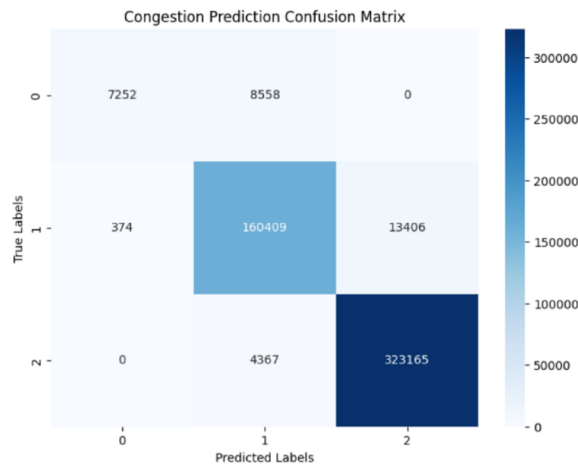
7

4.a ROC Curve

4.b Confusion Matrix

4.c Model Accuracy Across Training

4.d Model Loss Across Training

4.e Precision and Recall

4.f F1 Score

Fig. 4. Performance evaluation of the proposed model.

| First author (year)	Model	Dataset	Precision	Recall	F1-score	AUC-ROC	TPR	TNR
Youcef Djenouri (2022)[10]	CNN with decomposition	Urban traffic flow data (odense)	N/A	N/A	N/A	0.72	0.89	0.94
Armstrong Aboah (2021)[12]	YOLOv5 + decision tree	NVIDIA AI city challenge	N/A	N/A	0.8571	N/A	N/A	N/A
My Driss Laanaoui (2024)[13]	Machine learning-based system	Real-time traffic data	0.93	0.92	0.925	N/A	N/A	N/A
Tianyang Lei (2023)[11]	LSTM + CNN + spectrum decomposition	Traffic flow time series (Pems)	0.89	0.7	0.78	0.98	N/A	N/A
Proposed work	**Multi-head + LSTM**	**PeMS + weather**	**99.53**	**99.46**	**99.48**	**99.8**	**0.8979**	**0.9959**

Table 1. Comparison of the anomaly detection with existing work. Significant values are in bold.
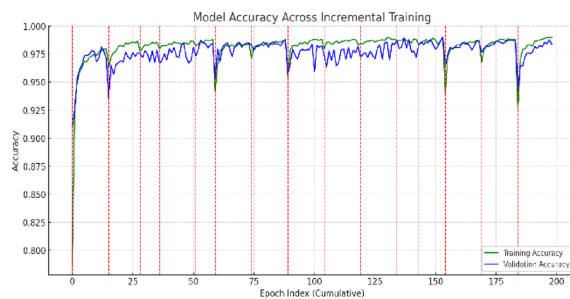
Congestion classification results

Figure 5a shows the result of the Congestion Classification detection model in form of confusion matrix which represents the prediction of the samples which correctly predicted vs incorrectly predicted. Figure 5b represents the roc curve of true positive rate vs false positive rate with an AUC score for each class.
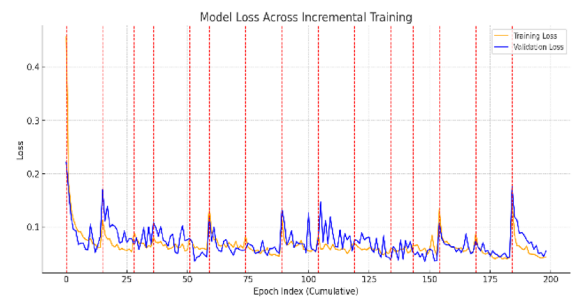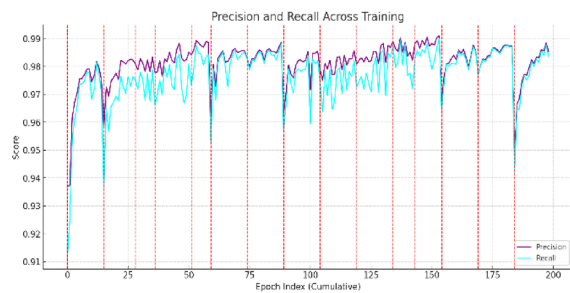
5.a Confusion Matrix
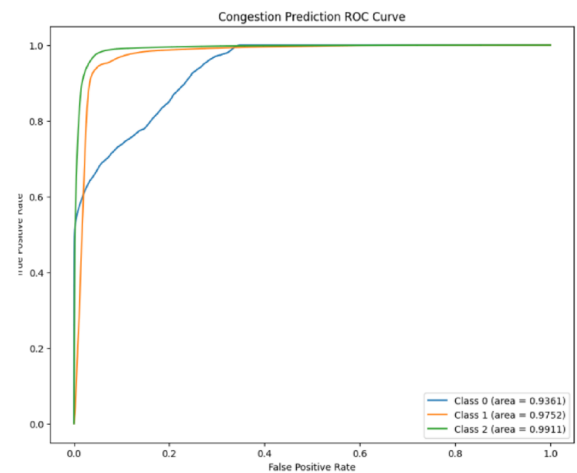
5.b ROC Curve

5.c Model Accuracy Across Training

5.d Model Loss Across Training

5.e Precision & Recall Across Training

5.f F1-Score Across Training

Fig. 5. Performance evaluation of the proposed model.

The congestion classification model has gained good results with an overall accuracy of 94.84% with a high F1-score of 94.55% and MCC of 0.8922 which demonstrates its effectiveness in classifying different congestion levels. The confusion matrix shown above in Fig. 5a indicates that the model correctly classifies most instances with a minimal misclassification across the three congestion levels. The ROC-AUC score of 0.9675 suggests excellent discrimination ability between the classes. The high recall values for medium and high congestion level are 0.92 and 0.99 respectively which indicates that the model is effectively able to classify the significant congestion scenarios. However, the lower recall of 0.46 for low congestion suggests misclassification is potentially due to overlapping of the features.

The above plots are the training logs as model was trained on multiple days csv file each red horizontal dotted line represent new csv given in incremental learning manner. Figure 5c shows the training and validation accuracy trends over epochs. Figure 5d shows model training and validation loss across epochs. Figure 5e demonstrates how precision and recall evolved throughout training. Figure 5f tracks the F1-score progression across training epochs. The training and validation loss plots demonstrate stable convergence across epochs with no significance of overfitting while the accuracy trends show consistent improvements in the precision-recall and f1-score trends to reinforce the model ability to maintain strong generalization.

Model comparisons with existing work

Table 2 shows the comparison with existing work for congestion classification. The proposed multi-Head + LSTM model achieves a competitive accuracy of 94.84% which outperforms the random forest-based models and demonstrates a good recall of 95% compared to existing work ensuring effective congestion detection. The ROC-AUC score of 98.41% indicates a strong discriminatory capability which significantly surpasses the spatial–temporal graph auto-encoder model AUC of 0.62 which highlights robustness. While the TAP model archives higher accuracy of 96% our approach balances precision, recall, f1-score making it highly reliable congestion classification model.

Incident classification results

Figure 6a shows the result of the incident detection model in form of confusion matrix which represent the prediction of the samples which correctly predicted vs incorrectly predicted. Figure 6b represents the roc curve of true positive rate vs false positive rate with an AUC score for each class.

For the incident classification task, the model demonstrates exceptional performances achieving good classification results as shown in Fig. 6. The confusion matrix shown in Fig. 6b indicates minimal misclassification with only 57 false positive and 5 false negative which highlights the model's ability to accurately detect the incidents. The classification report showcases a good high precision, recall and f1-score which shows that model is robust for identifying the incident class. The ROC-AUC curve in Fig. 6a gives a perfect score of 1 demonstrating the models' strong discriminatory power in separating the incidents for non-incidents.

The above plots are the training logs as model was trained on multiple days csv file each red horizontal dotted line represents new csv given to the model in incremental learning manner. Figure 6c shows the training and validation accuracy trends over epochs. Figure 6d shows model training and validation loss across epochs. Figure 6e demonstrates how precision and recall evolved throughout training. Figure 6f tracks the F1-score progression across training epochs. Logs indicate the improvements in accuracy and loss trends over epochs with validation metrics aligning closely with the training metrics which indicates that the model generalize well to unseen data. The gradual convergence of loss values and the stability of the precision, recall, and f1-score throughout the training.

Model comparison with existing work

Table 3 shows the comparison with existing work for incident classification. The proposed model outperforms existing methods in incidents classification achieving a good accuracy of 99.99% with precision, recall, and f1-scores at 0.99. compared to prior works like SAE 92.72 accuracy and network lasso 0.97 AUC-ROC of 1.0 and significantly lower false alarm rate 0.011 which highlights the model exception ability to minimize the misclassifications while maintaining high detection reliability.

Conclusion and future work

This study proposes a multi-head with LSTM model for anomaly detection, congestion classification and incident classification which is trained on PeMS traffic data with an integration of weather data. The model consistently outperformed existing methods achieving higher precision, recall, f1-score and AUC-ROC in all three objectives. The incremental training strategy helped maintaining stable learning while preventing overfitting. For anomaly detection the model accurately identified irregularities with optimized loss curve. Congestion classification has achieved 94.84% accuracy with an f1-score of 95% surpassing previous work like random forest and graph Auto-encoders. Similarly incident classification model has also achieved an exceptional good results with an accuracy of 99.99% outperforming the models such as Network Lasso, GAN-Transformers and SAE based approaches. The results demonstrate the importance of integrating traffic and environmental data for accurate real-time traffic monitoring making the proposed model a strong candidate for intelligent transportation system.

The proposed multi-head with LSTM model faces some key challenges some of them which are data imbalance and where for anomalies and incident are rare which leads to potential bias in prediction Limited weather data diversity is a key limitation as the study considered only 15 days of data. Consequently, the model has not been exposed to longer-term seasonal or holiday variations which could lead to concept drift over time. As a future extension we plan to incorporate multi-month and multi-seasonal datasets to validate the robustness of the proposed framework under diverse traffic conditions. Model interpretability as deep learning model function as black box makes it difficult to trust the model predictions. To overcome these adaptive rebalancing techniques can be used such as Temporally Constrained SMOTE (TCSMOTE) or GANs based data augmentation will be explored in future work, as these approaches better respect sequential dependencies compare to vanilla SMOTE. Online learning strategies, self-adjusting anomaly detection threshold and lighter architectures can be used for real-time performance and scalability for deployment in diverse traffic conditions. Future work will also extend

First author (year)	Model	Dataset	Accuracy (%)	Precision	Recall	F1-score	AUC
Syed Ammad Ali Shah (2024)[19]	Random forest,	Caltrans PeMS	93.33	93.34	93.33	N/A	**0.99**
Zhi liu (2023)[20]	Spatio-temporal variational graph auto-encoders	pems 8	**96**	**0.95**	93	94.4	0.62
Norman Bereczki[18]	xgb	Simulation dataset	0.9	0.84	0.88	0.86	0.96
Proposed Work	**Multi-head + LSTM**	PeMS + weather	94.84	95	**95**	**95**	98.41

Table 2. Comparison of the congestion classification with existing work. Significant values are in bold.

a.   ROC Curve   b.   Confusion Matrix

c.   Model Accuracy Across Training   d.   Model Loss Across Training

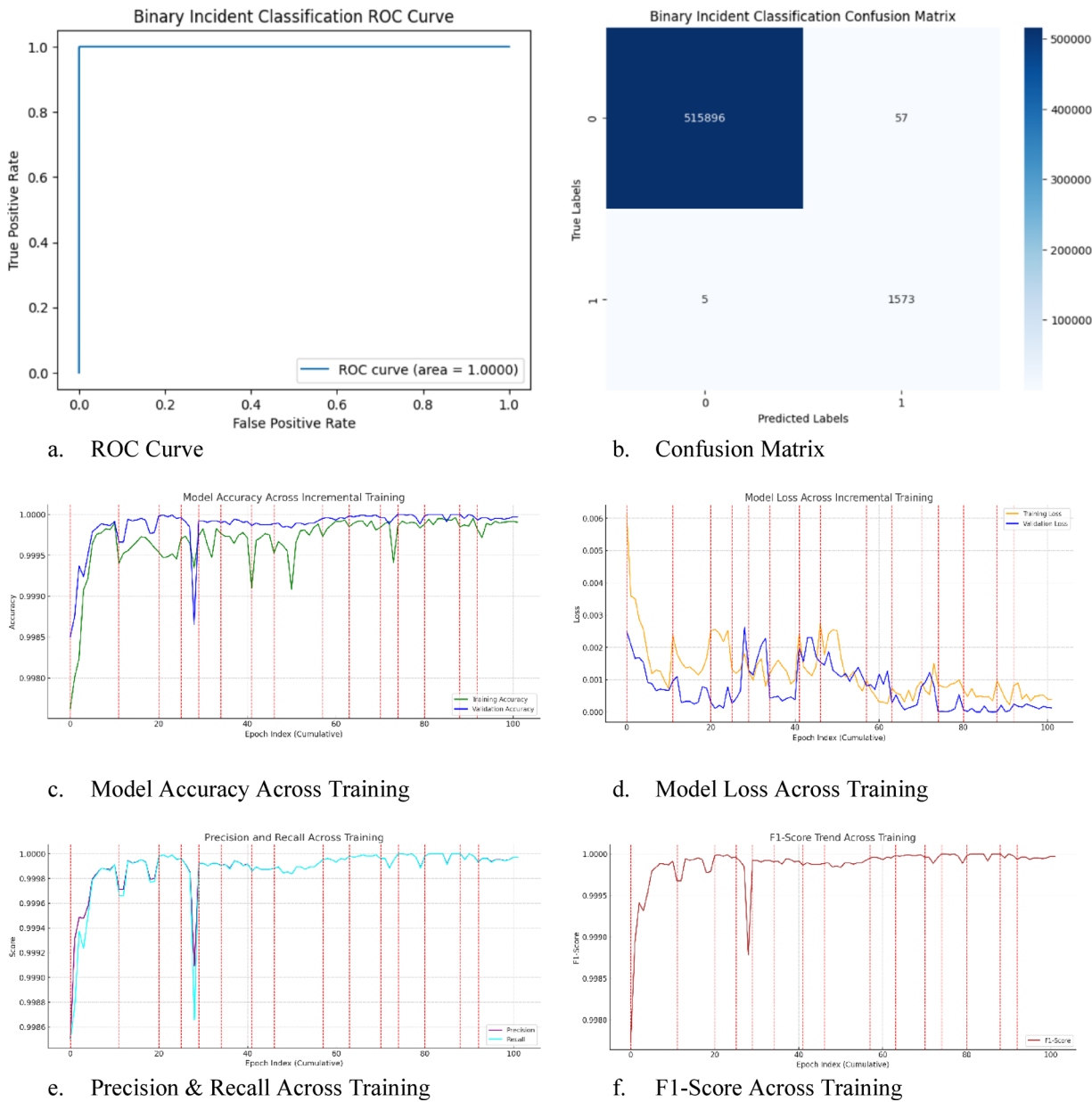e.   Precision & Recall Across Training   f.   F1-Score Across Training

Fig. 6. Performance evaluation of the proposed model.

First author (year)	Model	Dataset	Accuracy	Precision	Recall	F1-score	AUC-ROC	DR	FAR	CR	MCC
Qiyuan Zhu (2024)[26]	Network Lasso	Custom traffic data	0.97	N/A	N/A	0.79	0.94	0.88	N/A	N/A	N/A
Xinying Lu (2024)[23]	GAN + Transformer	PeMS	N/A	N/A	N/A	N/A	0.99	0.98	0.021	0.96	N/A
Tian Xie (2022)[24]	SASYNO-RF-RSKNN	PORTAL Highway I-205highway.	N/A	0.94	N/A	0.96	N/A	0.97	0.061	N/A	0.91
Rajalakshmi Gurusamy (2023)[25]	SAE	Caltrans PeMS	92.73	0.98	0.89	0.85	0.94	N/A	N/A	N/A	N/A
Proposed work	**Multi-head +LSTM**	PeMS + weather	**99.99**	**0.99**	**0.99**	**0.99**	**1**	**0.9968**	**0.011**	**0.99**	**0.98**

Table 3. Comparison of the incident classification in with existing work. Significant values are in bold.

the comparison to include advanced Transformer-based models to benchmark our framework against the latest SOTA architectures.

Data availability

References

1. Adewopo, V. A. & Elsayed, N. Smart city transportation: Deep learning ensemble approach for traffic accident detection. *IEEE Access* (2024).
2. Abraham, A., Zhang, Y. & Prasad, S. Evacuation management framework towards smart city-wide intelligent emergency interactive response system. *arXiv preprint* arXiv:2403.07003 (2024).
3. Almukhalfi, H., Noor, A. & Noor, T. H. Traffic management approaches using machine learning and deep learning techniques: A survey. *Eng. Appl. Artif. Intell.* **133**, 108147 (2024).
4. Azfar, T. et al. Deep learning-based computer vision methods for complex traffic environments perception: A review. *Data Sci. Transp.* **6**, 1 (2024).
5. Fährmann, D. Martín, L. Sánchez, L. & Damer, N. Anomaly detection in smart environments: A comprehensive survey. In *IEEE Access.* **12**, 64006–64049, (2024).
6. Iqbal, A. & Amin, R. Time series forecasting and anomaly detection using deep learning. *Comput. Chem. Eng.* **182**, 108560 (2024).
7. Kong, X. et al. Mobile trajectory anomaly detection: Taxonomy, methodology, challenges, and directions. *IEEE Internet Things J.* **11**, 19210 (2024).
8. Kun Wang, YuFu., Duan, X., Liu, T. & Jianqiao, Xu. Abnormal traffic detection system in SDN based on deep learning hybrid models. *Comput. Commun.* **216**, 183–194 (2024).
9. Van Engelen, J. E. & Hoos, H. H. A survey on semi-supervised learning. *Mach. Learn.* **109**(2), 373–440 (2020).
10. Djenouri, Y., Belhadi, A., Chen, H. C. & Lin, J. C. W. Intelligent deep fusion network for urban traffic flow anomaly identification. *Comput. Commun.* **189**, 175–181 (2022).
11. Lei, T. et al. A novel unsupervised framework for time series data anomaly detection via spectrum decomposition. *Knowl. Based Syst.* **280**, 111002 (2023).
12. Aboah, A. A vision-based system for traffic anomaly detection using deep learning and decision trees. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. (2021).
13. Laanaoui, M. D., Lachgar, M., Mohamed, H., Hamid, H., Villar, S. G. & Ashraf, I. Enhancing urban traffic management through real-time anomaly detection and load balancing. In *IEEE Access*, vol. 12, pp. 63683–63700, (2024).
14. Butler, L., Yigitcanlar, T. & Paz, A. Smart urban mobility innovations: A comprehensive review and evaluation. *IEEE Access* **8**, 196034–196049 (2020).
15. Xu, S., Sun, C. & Liu, N. Road congestion and air pollution-analysis of spatial and temporal congestion effects. *Sci. Total Environ.* **945**, 173896 (2024).
16. Buiya, M. R. et al. Detecting IoT cyberattacks: advanced machine learning models for enhanced security in network traffic. *J. Comput. Sci. Technol. Studi.* **6**(4), 142–152 (2024).
17. Kumar, K., Kumar, M. & Das, P. Traffic congestion forecasting using multilayered deep neural network. *Transp. Lett.* **16**(6), 516–526 (2024).
18. Bereczki, N. & Simon, V. A novel machine learning based traffic congestion recognition system. In *2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH)*. IEEE (2023).
19. Shah, S. A. A., Fernando, X. & Kashef, R. Improved vehicular congestion classification using machine learning for VANETs. In *2024 IEEE International Systems Conference (SysCon)*. IEEE (2024).
20. Liu, Z. et al. TAP: Traffic accident profiling via multi-task spatio-temporal graph representation learning. *ACM Trans. Knowl. Discov. Data* **17**(4), 1–25 (2023).
21. Olugbade, S., Ojo, S., Imoize, A. L., Isabona, J. & Alaba, M. O. A review of artificial intelligence and machine learning for incident detectors in road transport systems. *Math. Comput. Appl.* **27**(5), 77 (2022).
22. Remil, Y., Bendimerad, A., Mathonat, R. & Kaytoue, M. Aiops solutions for incident management: Technical guidelines and a comprehensive literature review. *arXiv preprint* arXiv:2404.01363 (2024).
23. Lu, X., Zhang, D. & Xiao, J. A hybrid model for traffic incident detection based on generative adversarial networks and transformer model. *arXiv preprint* arXiv:2403.01147 (2024).
24. Xie, T., Shang, Q. & Yang, Yu. Automated traffic incident detection: Coping with imbalanced and small datasets. *IEEE Access* **10**, 35521–35540 (2022).
25. Gurusamy, R., Mahalakshmi, M., Divya Bharathi, M. & Harishmaashree, G. Incident detection using LSTM and SAE. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol. IJSRCSEIT* **9**(8), 138–145 (2023) (**ISSN : 2456-3307**).
26. Zhu, Q., Qin, A. K., Abeysekara, P., Dia, H. & Grzybowska H. "Decentralised Traffic Incident Detection via Network Lasso," *arXiv preprint* arXiv:2402.18167 (2024).

Acknowledgements

Author contributions

1. Pratik Jadhav—Conceptualization, Methodology; Project administration; Supervision; Roles/Writing—original draft; Writing—review & editing. 2. Abderrahim Benslimane—Data Curation, Formal Analysis, Investigation, Validation; Visualization; Roles/Writing—original draft 3. Dr. Deepali Vora—Data Curation, Formal Analysis, Methodology; Project administration; Supervision; Writing—review & editing. 4. Dr. Shruti Patil—Methodology; Project administration; Resources; Validation; Visualization; Writing—review & editing.

Funding

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to D.R.V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.