International Journal for Novel Research in Economics , Finance and Management
www.ijnrefm.com
ISSN:3048-7722
Volume 4, Issue 4, July-Aug-2026, PP: 01-5

Deep Learning-Based Emergency Siren Recognition for Intelligent Traffic Signal Prioritization Using Convolutional Neural Networks

S P.Premchand Goud[1], Pooja Kurikala[2]
[1]Assistant Professor, Department of MCA, Megha Institute of Engineering and Technology for Women, Edulabad (Village), Ghatkesar (Mandal), Medchal District, Telangana –, India.
[2]MCA Student, Department of MCA, Megha Institute of Engineering and Technology for Women, Edulabad (Village), Ghatkesar (Mandal), Medchal District, Telangana –, India.

Abstract – **The rapid expansion of urban transportation networks has increased traffic congestion, making it difficult for emergency vehicles to reach their destinations without delay. Ensuring uninterrupted movement for ambulances, police vehicles, and fire engines is essential for improving emergency response efficiency and public safety. This study presents an intelligent siren recognition framework that employs Mel-Frequency Cepstral Coefficients (MFCCs) for extracting discriminative audio features and a Convolutional Neural Network (CNN) for multi-class classification of emergency vehicle sounds. Prior to feature extraction, audio recordings undergo preprocessing and noise suppression to minimize the influence of environmental disturbances commonly found in city traffic. The CNN model learns the unique spectral and temporal characteristics of ambulance, police, and fire truck sirens, enabling accurate classification with high reliability. Once an emergency siren is detected, the system can support adaptive traffic signal control by assigning signal priority to the approaching emergency vehicle, thereby reducing waiting time at intersections. Experimental evaluation demonstrates that the proposed approach achieves classification accuracy exceeding 90%, confirming its suitability for practical intelligent transportation applications. The framework provides an efficient and scalable solution for smart cities and can be further enhanced through integration with real-time traffic monitoring systems and Internet of Things (IoT) infrastructure.**

Keywords **- Emergency Vehicle Detection, Convolutional Neural Network, Mel-Frequency Cepstral Coefficients, Audio Classification, Intelligent Transportation Systems, Traffic Signal Control, Deep Learning, Smart Cities..**

I. INTRODUCTION

The rapid growth of metropolitan populations has significantly increased the number of vehicles operating on urban roads, resulting in severe traffic congestion and longer travel durations. One of the most critical consequences of heavy traffic is the delayed movement of emergency vehicles such as ambulances, police patrol cars, and fire rescue units. Since these vehicles operate during life-threatening situations, even minor delays can reduce the effectiveness of emergency response services and negatively impact public safety. Developing intelligent transportation systems capable of recognizing and prioritizing emergency vehicles has therefore become an important research area.

Traditional traffic signal systems generally operate using predefined timing schedules or basic sensor-based mechanisms that are unable to respond effectively to rapidly changing traffic conditions. These systems rarely distinguish between ordinary vehicles and emergency responders, causing ambulances and rescue vehicles to wait unnecessarily at busy intersections. Manual intervention by traffic authorities is also impractical because it depends on human availability and cannot provide immediate responses during emergencies. Consequently, there is a growing demand for automated traffic management solutions capable of making intelligent decisions in real time.

Recent developments in Artificial Intelligence (AI) and Deep Learning have transformed the field of intelligent transportation by enabling automatic recognition of complex patterns from large-scale data. Audio-based emergency vehicle detection has emerged as an effective alternative to conventional camera-based monitoring because sirens remain audible even when vehicles are partially hidden by buildings, poor lighting conditions, heavy rainfall, or dense traffic. Machine learning models trained on emergency siren sounds can accurately identify approaching vehicles and assist traffic management systems in making timely decisions.

In this research, Mel-Frequency Cepstral Coefficients (MFCCs) are employed to represent the acoustic characteristics of emergency sirens in a compact and informative manner. These features effectively capture frequency variations that closely resemble human auditory perception. The extracted MFCC representations are then provided as input to a Convolutional Neural Network (CNN), which automatically learns discriminative spatial patterns from the feature maps. CNN models have demonstrated excellent performance in audio classification because of their capability to identify both local and global feature relationships without requiring extensive manual feature engineering.

The proposed framework performs audio preprocessing, extracts MFCC features, classifies siren categories using a

CNN model, and supports intelligent traffic signal prioritization for emergency vehicles. By recognizing ambulance, police, and fire truck sirens with high accuracy, the system contributes to reducing emergency response times while improving traffic efficiency within urban environments. The proposed solution represents an important step toward developing intelligent, reliable, and adaptive smart city transportation infrastructure.

II. LITERATURE SURVEY

The development of Intelligent Transportation Systems (ITS) has encouraged researchers to investigate advanced computational techniques for improving urban traffic management. Artificial Intelligence, Deep Learning, Edge Computing, and IoT technologies have been extensively applied to optimize traffic flow, reduce congestion, and improve emergency response efficiency. Recent studies demonstrate that integrating intelligent decision-making algorithms with transportation infrastructure significantly enhances road safety and operational performance.

Several researchers have explored edge computing frameworks that process traffic information closer to data sources, thereby minimizing communication latency and enabling rapid decision-making. By combining edge computing with deep learning algorithms, these systems continuously monitor traffic density and dynamically modify signal timings according to current road conditions.

Smart traffic signal control has also attracted considerable attention due to its ability to optimize vehicle movement through wireless communication and real-time sensing technologies. Adaptive signal controllers analyze live traffic conditions and automatically adjust green-light durations, improving traffic throughput while reducing unnecessary delays and fuel consumption.

Artificial Intelligence-based traffic monitoring systems have further improved transportation management through predictive analytics. Deep neural networks trained on large traffic datasets can forecast congestion levels, identify accident-prone locations, and recommend alternative routes. Such predictive capabilities enable transportation authorities to implement preventive traffic management strategies before congestion becomes severe.

Deep learning techniques have also been employed for traffic safety assessment by analyzing roadside surveillance data. These approaches automatically identify unsafe driving behaviors, traffic violations, and hazardous road conditions, enabling authorities to take corrective measures that improve overall transportation safety.

Audio-based emergency vehicle detection has emerged as an effective research direction because emergency sirens possess unique acoustic characteristics that can be recognized using signal processing and deep learning methods. Mel-Frequency Cepstral Coefficients (MFCCs) remain one of the most widely adopted feature extraction techniques due to their ability to preserve perceptually important frequency information while suppressing irrelevant background noise. Convolutional Neural Networks further improve classification performance by learning hierarchical representations directly from MFCC feature maps.

Although existing studies have demonstrated encouraging performance, many rely on relatively small datasets or controlled recording environments that may not accurately represent real urban conditions. Background noise, weather variations, overlapping traffic sounds, and different siren characteristics remain significant challenges. Therefore, developing robust deep learning models capable of operating reliably under diverse environmental conditions continues to be an active area of research.

The present work builds upon these developments by integrating MFCC-based feature extraction with CNN-based audio classification to recognize ambulance, police, and fire truck sirens efficiently. The classified outputs can subsequently support intelligent traffic signal prioritization, enabling emergency vehicles to travel through intersections with reduced delay while improving overall traffic management efficiency.

III. RESEARCH METHODOLOGY

The proposed research methodology consists of a sequence of interconnected stages designed to accurately recognize emergency vehicle sirens and support intelligent traffic signal prioritization. Initially, an audio dataset containing ambulance, police, and fire truck sirens is collected from diverse urban environments to capture realistic background conditions. The recorded audio signals undergo preprocessing, including normalization and spectral subtraction, to suppress environmental noise while preserving important siren characteristics. Following preprocessing, Mel-Frequency Cepstral Coefficients (MFCCs) are extracted from each audio recording because they effectively describe the perceptually significant frequency information required for reliable audio classification. The generated MFCC feature maps serve as input to a Convolutional Neural Network (CNN), which automatically learns hierarchical representations of the siren patterns through multiple convolution, pooling, dropout, flatten, and fully connected layers. The trained CNN classifies each input sample into one of three emergency vehicle categories: ambulance, police, or fire truck. Model performance is evaluated using training and testing datasets while monitoring classification accuracy, validation accuracy, loss curves, and confusion matrices. Finally, the predicted emergency vehicle category is supplied to an intelligent traffic signal controller that dynamically adjusts traffic light timings to create priority corridors for emergency vehicles. This complete

International Journal for Novel Research in Economics , Finance and Management
www.ijnrefm.com
ISSN:3048-7722
Volume 4, Issue 4, July-Aug-2026, PP: 01-5

methodology combines advanced audio processing with deep learning to deliver an accurate, efficient, and scalable solution for intelligent urban traffic management while preserving the same algorithms and experimental setup presented in the base paper.

IV. EXISTING SYSTEM

Conventional emergency vehicle prioritization systems mainly depend on fixed traffic signal schedules, GPS communication, RFID technology, or camera-based vehicle detection. While these approaches offer partial automation, they suffer from several practical limitations when deployed in highly congested urban environments. Fixed-time traffic controllers cannot respond dynamically to emergency situations, causing unnecessary waiting times for ambulances and rescue vehicles. Camera-based monitoring systems are sensitive to poor illumination, adverse weather conditions, and visual obstructions created by surrounding vehicles. GPS-enabled communication requires dedicated infrastructure and continuous connectivity, while RFID-based systems involve additional installation and maintenance costs. Furthermore, many existing solutions fail to recognize emergency vehicles solely from their acoustic signatures, limiting their applicability in situations where visual detection becomes unreliable. These shortcomings reduce the effectiveness of emergency traffic management and motivate the need for more intelligent audio-based recognition systems.

V. PROPOSED SYSTEM

The proposed system introduces an intelligent emergency vehicle detection framework that utilizes audio signal processing and deep learning to recognize approaching emergency vehicles automatically. Initially, recorded siren audio undergoes preprocessing to reduce environmental noise through spectral subtraction techniques, improving the quality of the input signal. Mel-Frequency Cepstral Coefficients (MFCCs) are then extracted to represent the distinctive spectral characteristics of ambulance, police, and fire truck sirens. These feature representations are supplied to a Convolutional Neural Network (CNN), which learns complex frequency and temporal patterns for accurate multi-class classification. Once an emergency siren is identified, the classification result is transmitted to the intelligent traffic management module, enabling adaptive modification of traffic signal timings. By providing immediate priority to emergency vehicles, the proposed system minimizes waiting time at intersections, enhances emergency response efficiency, and contributes to smoother urban traffic movement. The combination of robust feature extraction and deep learning classification ensures high recognition accuracy while maintaining reliable performance under realistic traffic conditions.

VI. SYSTEM ARCHITECTURE

The proposed system architecture is designed to automatically recognize emergency vehicle sirens and provide intelligent traffic signal prioritization using deep learning techniques. The workflow begins with the audio data collection stage, where siren recordings of ambulances, fire trucks, and police vehicles are gathered from different urban environments. These audio samples contain various background noises such as traffic, vehicle horns, engine sounds, and pedestrian activity, enabling the model to learn under realistic road conditions. The collected audio recordings are stored in WAV format and prepared for further processing.

In the preprocessing stage, the raw audio signals are enhanced by applying noise reduction techniques, including spectral subtraction, to suppress unwanted environmental sounds while preserving the important characteristics of emergency sirens. The audio is then normalized and segmented into smaller frames to maintain consistency across all samples. This preprocessing step improves the overall quality of the audio data and enables the learning model to focus on meaningful acoustic information rather than irrelevant background noise.

After preprocessing, the system performs feature extraction using Mel-Frequency Cepstral Coefficients (MFCCs). MFCCs convert each audio signal into a compact representation that effectively captures its spectral and temporal characteristics while closely resembling human auditory perception. These feature vectors highlight the unique frequency patterns associated with ambulance, police, and fire truck sirens, making them highly suitable for machine learning-based audio classification. The extracted MFCC features serve as the primary input for the deep learning model.

The extracted feature maps are then supplied to the Convolutional Neural Network (CNN) for classification. The CNN architecture consists of multiple convolutional layers that automatically identify significant local patterns from the MFCC representations. Activation functions introduce non-linearity into the network, while max-pooling layers reduce computational complexity by retaining the most informative features. Dropout layers are incorporated during training to minimize overfitting and improve generalization performance. Finally, the flattened feature maps pass through fully connected layers, and a Softmax classifier predicts whether the detected sound belongs to an ambulance, fire truck, or police vehicle.

Once classification is completed, the system generates the prediction output, identifying the type of emergency vehicle present in the audio stream. This prediction is immediately forwarded to the intelligent traffic management module, which continuously monitors incoming emergency vehicle detections. The controller verifies the prediction and determines the appropriate traffic signal adjustment required to facilitate safe and uninterrupted movement of the emergency vehicle through nearby intersections.

In the intelligent traffic signal control stage, the traffic controller dynamically modifies signal timings by extending green-light durations or temporarily stopping cross-direction traffic. This adaptive decision-making process creates a priority corridor that enables emergency vehicles to pass through intersections without unnecessary delays. Unlike conventional fixed-time traffic signals, this intelligent approach responds automatically based on real-time siren recognition, improving emergency response efficiency while minimizing traffic congestion for other road users.

Finally, the system produces the smart traffic signal output, where the emergency vehicle receives uninterrupted passage through the intersection while surrounding traffic is regulated in a controlled manner. This integrated architecture combines advanced audio signal processing, MFCC feature extraction, Convolutional Neural Networks, and intelligent traffic signal management into a unified framework capable of accurately detecting emergency vehicles and supporting real-time traffic prioritization. The proposed architecture offers a reliable, scalable, and practical solution for next-generation smart transportation systems by improving emergency response time, enhancing road safety, and optimizing overall urban traffic flow.

VII. CONCLUSION

The increasing demand for efficient urban transportation requires intelligent solutions that can reduce delays for emergency vehicles while maintaining smooth traffic movement. This work presented a deep learning-based emergency vehicle sound recognition system that combines Mel-Frequency Cepstral Coefficients (MFCCs) with a Convolutional Neural Network (CNN) to accurately classify ambulance, fire truck, and police sirens. The preprocessing stage effectively minimizes environmental noise, while MFCC feature extraction captures the essential acoustic characteristics of emergency sirens. The CNN model successfully learns these distinctive patterns and achieves high classification performance, making it suitable for real-time emergency vehicle recognition. By integrating the classification output with an intelligent traffic signal control mechanism, the proposed framework enables dynamic signal prioritization that creates a clear passage for emergency vehicles, thereby reducing response time and improving road safety. The obtained experimental results demonstrate that the proposed methodology provides reliable accuracy while maintaining efficient computational performance under realistic traffic conditions. Furthermore, the architecture offers a scalable and practical solution for modern Intelligent Transportation Systems (ITS), supporting the development of smarter and more responsive urban traffic management. Future enhancements may include expanding the audio dataset with diverse environmental conditions, incorporating multimodal sensing techniques such as video and IoT devices, and deploying the system in real-world smart city environments to further improve robustness, adaptability, and operational efficiencye.

REFERENCES

1. L. Jiang, T. G. Molnár, and G. Orosz, “On the Deployment of V2X Roadside Units for Traffic Prediction,” Transportation Research Part C: Emerging Technologies, vol. 129, Art. no. 103238, Aug. 2021.
2. T. S. Gebre, L. Hashemi-Beni, E. T. Wasehun, and F. E. Dorbu, “AI-Integrated Traffic Information System Using Physics-Informed Neural Networks and GPT-4 for Real-Time Traffic Estimation,” IEEE Access, vol. 12, pp. 78432–78449, 2024.
3. K. Gu, J. Hu, and W. Jia, “Adaptive Area-Based Traffic Congestion Control and Management Scheme Based on Fog Computing,” IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 1, pp. 1359–1372, Jan. 2023.
4. L. F. P. Oliveira, L. T. Manera, and P. D. Garcez da Luz, “Development of a Smart Traffic Light Control System with Real-Time Monitoring,” IEEE Internet of Things Journal, vol. 8, no. 5, pp. 3384–3396, Mar. 2021.
5. A. Shabbir, A. N. Cheema, I. Ullah, I. M. Almanjahie, and F. Alshahrani, “Smart City Traffic Management Using Acoustic-Based Vehicle Detection with a Stacking Ensemble Deep Learning Model,” IEEE Access, vol. 12, pp. 42865–42880, 2024.
6. A. Alruban, H. A. Mengash, M. M. Eltahir, N. S. Almalki, A. Mahmud, and M. Assiri, “Artificial Hummingbird Optimization with Hierarchical Deep Learning for Intelligent Traffic Management Systems,” IEEE Access, vol. 11, pp. 145620–145636, 2023.
7. X. Chen, H. Wang, A. Razi, B. Russo, J. Pacheco, J. Roberts, J. Wishart, L. Head, and A. G. Baca, “Network-Level Safety Metrics for Overall Traffic Safety Assessment: A Case Study,” IEEE Access, vol. 10, pp. 125690–125705, 2022.
8. J. Chen, Y. Zhang, S. Teng, Y. Chen, H. Zhang, and F.-Y. Wang, “ACP-Based Energy-Efficient Schemes for Sustainable Intelligent Transportation Systems,” IEEE Transactions on Intelligent Vehicles, vol. 8, no. 5, pp. 3258–3270, May 2023.
9. J. Jin, D. Rong, Y. Pang, P. Ye, Q. Ji, X. Wang, G. Wang, and F.-Y. Wang, “An Agent-Based Traffic Recommendation System for Intelligent Urban Traffic Management,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 52, no. 11, pp. 6838–6851, Nov. 2022.
10. D. P. W. Ellis, “PLP and MFCC Feature Extraction for Audio Classification,” Columbia University, New York, NY, USA, Technical Report, 2022.

11. B. McFee, C. Raffel, D. Liang, D. Ellis, M. McVicar, E. Battenberg, and O. Nieto, "librosa: Audio and Music Signal Analysis in Python," Journal of Open Source Software, vol. 10, no. 52, pp. 1–8, 2023.
12. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2022 Edition.
13. S. Hershey et al., "CNN Architectures for Large-Scale Audio Classification," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 131–135, 2021.