Intelligent Traffic Signal Prioritization for Emergency Vehicle Diversion in Urban Environments Using Multi-Modal Deep Learning

Santosh D. Pandure[1], Pravin L. Yannawar[2]

[1,2]Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Chhatrapati Sambhaji Nagar, Maharashtra, India.

Abstract:
This research work introduces an innovative approach for diverting traffic in urban areas by prioritizing traffic signals for emergency vehicles. The proposed method leverages deep learning techniques to identify and prioritize emergency vehicles, thereby enabling the manipulation of signal sequences at local traffic intersections along their path. The identification of emergency vehicles is achieved through the integration of multiple artificial intelligence(AI)-based algorithms, combining audio and video data analysis to enhance accuracy and reliability in urban settings. This research contributes to the enhancement of urban traffic management, reducing response times for emergency services, and improving overall road safety.

Keywords: Emergency Vehicle(EV), artificial intelligence(AI), Emergency services, Urban.

I. Introduction

In urban areas, efficient traffic management is crucial to ensure timely response of emergency services. Delays in the movement of emergency vehicles can have critical consequences. This research addresses the issue of optimizing traffic signal prioritization for emergency vehicles in urban environments. The primary objective of this research is to develop a comprehensive system that employs deep learning techniques to identify and prioritize emergency vehicles, thereby allowing the adjustment of traffic signals at local intersections along their route. This system is designed to significantly improve the efficiency of emergency response in urban areas.

II. Literature Review

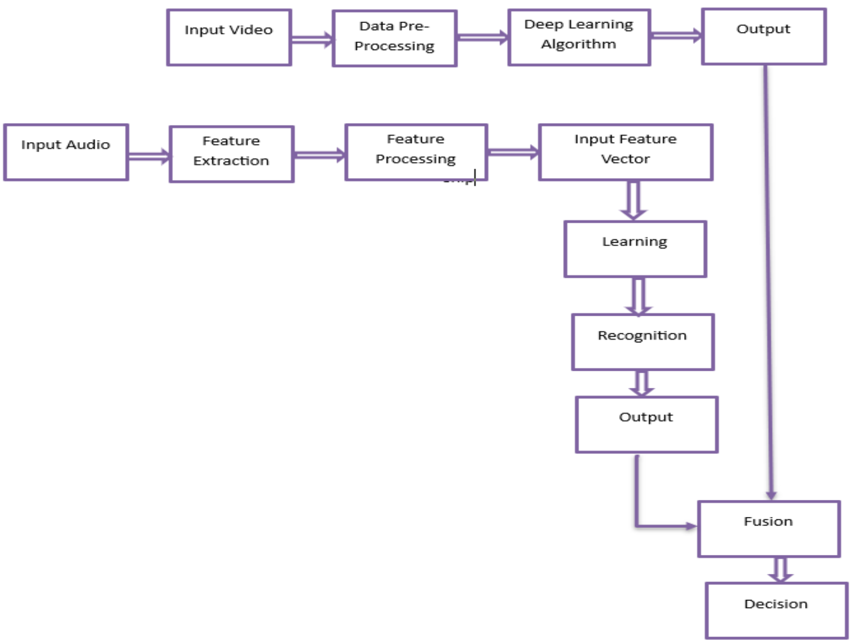
The research presents a Deep ConvNet2D-based surveillance system for identifying emergency vehicles in congested traffic, aiming to improve response times. It highlights the potential of audio-based analysis with Extreme Learning Machine (ELMs) for real-time traffic monitoring and event detection [1].The review discusses the rise of camera-based traffic monitoring systems in response to the growing number of vehicles in urban areas. It categorizes shadow detection methods into four types and emphasizes the importance of robust shadow detection for accurate vehicle segmentation in such systems [2].The review underscores the significance of audio recognition in smart city applications, addressing challenges posed by unstructured urban sounds. It emphasizes the importance of Deep Learning architectures, data augmentation, and feature selection for achieving high accuracy in sound classification models, while recognizing the complexity of distinguishing sounds in real-world urban environments[3].The paper introduces a framework for alerting sound event detection in driving scenarios, enhancing smart vehicle capabilities in understanding urban sounds. By employing noise-removal techniques based on gammatonegrams and STFT, it achieves improved classification accuracy, with future prospects for place-dependent soundscape models and multi-modal detection [4].This paper delves into object detection using Convolutional Neural Networks (CNNs) and compares the performance of SSD with MobileNetV1 and Faster-RCNN with InceptionV2 models. The findings highlight a trade-off between speed and accuracy, making them suitable for different applications[5]. This paper presents a comprehensive survey of deep learning-based object detection methods, categorizing them into anchor-based, anchor-free, and transformer-based approaches[6]. Augmentation improves performance, and class-specific data augmentation is recommended for further enhancement [7]. The study evaluates bird sound classification using deep learning models and fusion approaches, yielding a balanced accuracy of 86.31% and a weighted F1-score of 93.31% [8].This paper presents Fast R-CNN, an efficient object detection method that outperforms previous approaches in terms of speed and accuracy. It highlights the potential of sparse object proposals to improve detector quality[9].

This paper highlights the significance of automatic audio analysis for applications like public safety and wildlife monitoring. It introduces an audio event detection system based on deep learning with features like MFCC and log Mel spectrogram, showing promising results and emphasizing the potential for larger datasets and scalability in the future[10].Deep learning-based approaches, such as Convolutional Neural Networks and transfer learning, show promise in addressing the challenge of real-time emergency vehicle detection in heavy traffic scenarios. These technologies aim to enhance response times and safety by accurately identifying emergency vehicles and generating green corridors[11].The study outlines a supervised machine learning approach for alarm sound detection, achieving high accuracy and low false positive rates, making it suitable for real-time audio monitoring applications in noisy environments. The methodology incorporates feature selection and pre-processing techniques to enhance performance and robustness[12].The study introduces a highly effective acoustic digital signal processing algorithm for automatic medical alarm identification, demonstrating exceptional accuracy and robustness in both quiet and noisy conditions. This approach holds great promise for enhancing patient safety and addressing alarm fatigue issues in healthcare settings[13]. The paper introduces an intelligent traffic control system driven by deep learning, addressing the issues of traffic congestion and emergency vehicle prioritization. By dynamically adjusting traffic light schedules based on vehicle density and the presence of emergency vehicles, this system offers a promising solution to optimize urban traffic management [14]. The Ambulance Detection system employs machine learning and image processing to swiftly identify ambulances in congested traffic. This innovative approach has the potential to save lives by expediting ambulance passage during emergencies[15].This research presents advanced EVD systems that combine vision and audio approaches for detecting emergency vehicles. YOLO-EVD provides strong detection accuracy for visual data, while WaveResNet offers robust audio-based detection, and the integrated AV-EVD system shows promising results for real-world applications[16].The research presents a novel training criterion for deep neural networks that combines cross-entropy and M3 CE, resulting in enhanced performance on image classification tasks. CNNs outperform other classifiers in terms of accuracy, making them a suitable choice for image classification applications[17].The paper presents an automated system for detecting emergency vehicles in heavy traffic using deep convolutional neural networks. This technology has the potential to improve response times and save lives in congested urban areas[18]. This paper presents an automated system for ambulance detection in traffic using computer vision. By training a Convolutional Neural Network (CNN) and implementing a detection phase, it helps clear the way for ambulances at traffic signals, potentially saving lives[19].The research presents a deep learning-based ensemble model for efficient emergency vehicle detection using siren sounds, achieving an impressive accuracy of 98.7%. This technology has the potential to enhance traffic management and emergency response systems[20].The paper demonstrates the feasibility of using deep learning and convolutional neural networks, particularly AlexNet and GoogLeNet, for environmental sound classification, which has applications in mobile devices. The investigation evaluates different sound representations like Spectrograms, MFCC, and CRP for classification accuracy[21].

The article presents a sensor combining magnetometer and microphone for traffic monitoring, with an emphasis on detecting emergency vehicles through acoustic signal analysis. It highlights the sensor's sensitivity, classification potential based on vehicle length and magnetic field changes, and plans for expanding the sensor network for traffic monitoring and algorithm development[22].

III. Proposed Methodology

a. Emergency Vehicle Identification using video and audio



Detecting emergency vehicles in videos involves gathering diverse data and meticulously labeling emergency vehicle presence. Augmenting data enhances model adaptability, aiding in choosing appropriate deep learning models like YOLO or SSD. Training, fine-tuning, and evaluating the model refine its accuracy and performance metrics. Post-processing steps enhance results before deploying the model for real-time inference, ensuring continuous monitoring for adaptability and sustained accuracy.

Detecting emergency vehicles through audio involves collecting a diverse dataset of emergency vehicle sounds. Preprocessing extracts key features, like frequency and amplitude, from the audio data. A chosen machine learning or deep learning model, such as CNNs or RNNs, is trained on labeled data to identify emergency vehicle patterns. Evaluation metrics assess the model's accuracy, precision, and recall. Once trained, the model can be deployed for real-time inference, contributing to audio-based emergency vehicle detection systems.

b. Traffic Signal Prioritization

Decision fusion aims to leverage the strengths of multiple modalities to enhance overall accuracy and reliability in complex tasks like emergency vehicle detection. Regularly assess and refine the fusion strategy based on real-world performance and changing conditions.

IV. Experimental Results

1. Data Collection and Preprocessing:

We utilized two separate datasets in our research project. The first dataset, comprising 2352 images, was sourced from Kaggle and was employed for identifying vehicles based on images. Additionally, we incorporated a second dataset for proposed research, which consists of 3-second .wav format audio files. This dataset includes audio recordings of Emergency Vehicles such as Ambulance and Firetruck sirens. Another category within this dataset is labeled "Traffic," which contains 3-second .wav format audio files capturing ambient traffic sounds. Each category contains 200 samples.

2. Model Selection and Training:

2.1 Data Preprocessing:

Data preprocessing involves transforming raw data into a more usable and informative format, often achieved through machine learning algorithms, mathematical modeling, and statistical techniques. Automation is a key aspect of this process. In the context of image data, it's common to encounter artifacts or undesired regions that need removal. To address this, it's crucial to identify and categorize these regions for effective elimination. To facilitate this, we utilized the Numpy module and the PIL image library, enabling the storage of pixel values in arrays for efficient operations. Additionally, when dealing with images, noise can be a disruptive element. Employing techniques like v2.fastNIMeansDenoisingColored() with parameters such as 'h' (filter strength) helps in effectively reducing noise. This step contributes to enhancing the overall quality of images in the dataset.

2.2 Model Creation:

The architecture of CNNs encompasses various layers designed for specific tasks. Introduction to the Network: Neural networks, inspired by the human brain, include diverse types, and one such variant is the Convolutional Neural Network (CNN). Specialized in image processing, CNNs excel in tasks like object identification, segmentation, and image classification.

2.3 Pooling Layer:

This reduction in dimensionality not only optimizes computational efficiency but also minimizes the demand for computing power during data processing.

2.4 Fully Connected Layer:

Leveraging the output of a convolutional layer, the Fully Connected Layer instructs the model to learn non-linear features. Transforming this output into a column vector facilitates multi-perception, sending it through a feed-forward neural network with backpropagation during each training cycle.

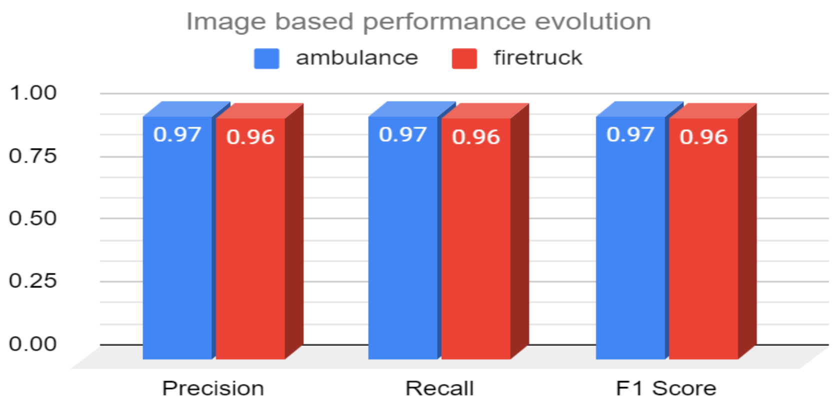
2.5 Loading Data into Batch:

Batch processing emerges as an effective method for consistently handling large datasets. By utilizing computational resources, the batch technique enables automated data processing with minimal human intervention. This approach proves particularly beneficial when dealing with extensive datasets in machine learning applications.

3. Decision Fusion Strategies:

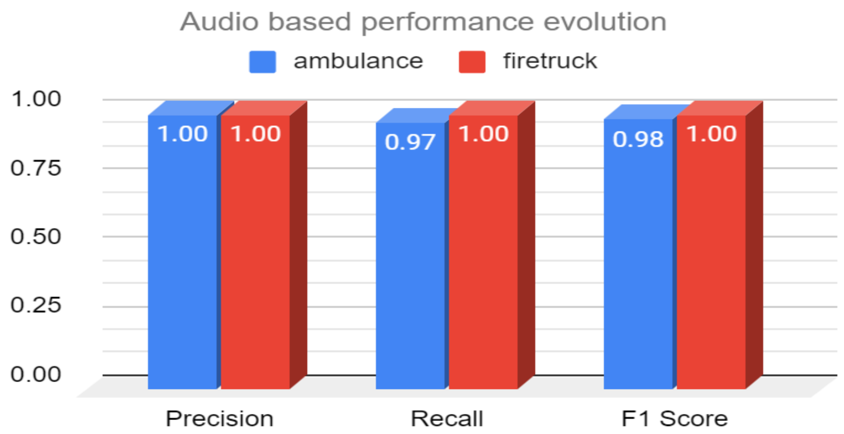
Decision Fusion Strategies, particularly late fusion, were employed in our approach. In late fusion, the outputs of two distinct machine learning models were considered independently. In the final decision-making process, if either one of the models detects the presence of an emergency vehicle, the overall conclusion is that an emergency vehicle is detected. This means that the final output is determined based on the collective results of both models, and the system doesn't require a consensus from both models to identify the presence of an emergency vehicle. The advantage of this approach lies in its flexibility and robustness, allowing the system to be effective even if one of the models performs exceptionally well in certain scenarios or under specific conditions.

4. Performance Evaluation:



Graph 1 Classification report on image based emergency vehicle detection

Graph No 1 represents the strong performance across all classes, achieving precision, recall, and F1 scores of 0.97 for ambulances, 0.96 for firetrucks. The overall accuracy stands at 96%, showcasing the model's ability to correctly classify instances. In summary, the model demonstrates robust and accurate classification across diverse emergency vehicle categories.



Graph 2 Classification report on audio based emergency vehicle detection

The model showcases exceptional precision, achieving a perfect score of 1.00 for both "ambulance" and "firetruck" classes, indicating precise identification of positive instances. Recall is consistently high, with perfect scores for "firetruck" and "traffic" (1.00) and 0.97 for "ambulance," demonstrating the model's ability to capture relevant instances effectively. F1 scores are robust across all classes, ranging from 0.98 to a perfect 1.00, highlighting a harmonious balance between precision and recall. The overall accuracy of 99% underscores the model's remarkable performance in accurately classifying instances. In summary, the model excels in precision, recall, and F1 score, resulting in outstanding accuracy and consistent performance across diverse emergency vehicle classes.

V. Conclusion

This research presents a cutting-edge solution for enhancing urban traffic management by prioritizing traffic signals for emergency vehicles. The integration of deep learning, multi-modal data analysis, and adaptive signal control promises to mitigate traffic congestion and optimize emergency response times. Future work will focus on real-world deployment and the fine-tuning of the system for broader urban implementation.

VI. Acknowledgement

The authors gratefully acknowledge support by the Department of Computer Science and Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Chhatrapati Sambhaji Nagar (MS) India.

References

1. Shankar Ka, Madhavan Va, Muhammad Tariq P Ra, and Sanjay Subramanian Ka,"Emergency Vehicle Detection in Heavy Traffic using Deep ConvNet2D and Computer Vision", Proceedings of the International Conference on Innovative Computing & Communication (ICICC) 2022.
2. Pablo Barcellosa, Vitor Gomesa, and Jacob Scharcanskia, "Shadow Detection in Camera-based Vehicle Detection: Survey and Analysis ", Journal of Electronic Imaging, May 2016.
3. Ana Filipa Rodrigues Nogueira, Hugo S. Oliveira, José J. M. Machado and João Manuel R. S. Tavares, " Sound Classification and Processing of Urban Environments: A Systematic Literature Review", Sensors 2022.
4. Letizia Marchegiani and Ingmar Posner, "Leveraging the Urban Soundscape: Auditory Perception for Smart Vehicles", IEEE International Conference on Robotics and Automation (ICRA) Singapore, May 29 - June 3, 2017.
5. Reagan L. Galvez, Argel A. Bandala, Elmer P. Dadios, Ryan Rhay P. Vicerra, and Jose Martin Z. Maningo, "Object Detection Using Convolutional Neural Networks", TENCON 2018 - 2018 IEEE Region 10 Conference, Jeju, Korea (South), 2018, pp. 2023-2027, doi: 10.1109/TENCON.2018.8650517.
6. Ayoub Benali Amjoud and Mustapha Amrouch, "Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review", IEEE, 2023.
7. Justin Salamon and Juan Pablo Bello, "Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification", IEEE Signal Processing Letters, 2016.
8. Jie Xie, Kai Hu, Mingying Zhu, Jinghu Yu, and Qibing Zhu, "Investigation of Different CNN-Based Models for Improved Bird Sound Classification ", IEEE, 2019.
9. Ross Girshick, "Fast R-CNN", IEEE International Conference on Computer Vision, 2015.

10. Sophiya E., and Jothilakshmi S., “Audio event detection using deep learning model”, Int. J. Computer Aided Engineering and Technology, Vol. 16, No. 3, 2022.
11. Hari Vignesh K, Musharraf U., and Balaji A,”Emergency Vehicle Detection using Deep Learning”, IJARIIE-ISSN(O)-2395-4396, Vol-9 Issue-2, 2023.
12. Dean Carmel, Ariel Yeshurun, and Yair Moshe, “Detection of Alarm Sounds in Noisy Environments”, 25th European Signal Processing Conference (EUSIPCO), 2017.
13. Paul Potnuru , Richard H. Epstein , Richard McNeer , and Christopher Bennett, “Development and Validation of an Algorithm for the Identification of Audible Medical Alarms”, Article in Cureus, November 2020.
14. Aiesha Roshan, Manukrishnan, Sneha Mohandas, and Sreejith PR, “Traffic Density Control Using Deep Learning”, IJRES, Vol 10 Issue 6, 2022.
15. K Agrawal, M K Nigam, S Bhattacharya, and Sumathi G, “Ambulance detection using image processing and neural networks”, Journal of Physics: Conference Series, 2021.
16. Van-Thuan Tran, and Wei-Ho Tsai, “ Audio-Vision Emergency Vehicle Detection”, IEEE SENSORS JOURNAL, VOL. 21, NO. 24, DECEMBER 15, 2021.
17. Mingyuan Xin, and Yong Wang, “Research on image classification model based on deep convolution neural network”, J Image Video Proc. 2019.
18. Shuvendu Roy, and Md. Sakif Rahman, “Emergency Vehicle Detection on Heavy Traffic Road from CCTV Footage Using Deep Convolutional Neural Network”, International Conference on Electrical, Computer and Communication Engineering (ECCE), 7-9 February, 2019.
19. Aishwarya W, Mrs. Sheelavathi A, Jhanani R S, Karthika D R, and Keerthana P, “Detection of Ambulance Using Computer Vision”, IJRASET, Volume 10 Issue VI June 2022.
20. Usha Mittal, and Priyanka Chawla, “Acoustic Based Emergency Vehicle Detection Using Ensemble of Deep Learning Models”, Elsevier, 2023.
21. Venkatesh Boddapati, Andrej Petef, Jim Rasmusson, and Lars Lundberg, “Classifying environmental sounds using image recognition networks”, International Conference on Knowledge Based and Intelligent Information and Engineering, 2017.
22. Karpis, and Ondrej, “System for Vehicles Classification and Emergency Vehicles Detection”, IFAC Proceedings Volumes, 2012.