Smart Ambulance Traffic Sensing using Artificial Intelligence and Internet of Things

Gopinathan S
Assistant Professor, Department of Artificial Intelligence and Data Science,
Sri Sai Ram Institute of Technology, *Sai Leo Nagar, West Tambaram, Chennai, India.*

Abishek B, Kathiravan G, Roshith P B, Bharath V
B.Tech, Department of Artificial Intelligence and Data Science,
Sri Sai Ram Institute of Technology, *Sai Leo Nagar, West Tambaram, Chennai, India.*

***Abstract* - Living in a world of high-density traffic leads to numerous issues, including significant loss of life due to ambulance delay caused by heavy congestion. To overcome this problem, we must automate the process for clearing traffic junctions as Ambulances arrive at these points and ensure they reach their destination on time. Our main objective is preserving human lives while lowering mortality rates created by tardiness. The implementation of Deep Learning coupled with embedded systems play a critical role here. Audio detectors capture siren sounds. CNN (Convolutional Neural Network) and ANN (Artificial Neural Network) algorithms detect the ambulance in real time from the live footage and classify the sound produced when responding emergency vehicles move through intersections effectively using YOLO (You Only Look Once) algorithm. Then, Raspberry Pi is programmed to integrate input data and the model required towards modifying our still functioning Traffic algorithm hence achieving more precise results. Hopefully, providing assistance from above-described technologies will eventually result in reduced death counts and arising out delayed medical interventions.**

Keywords - Traffic signal controlling, Deep Learning, Embedded systems, Audio Detectors YOLO, CNN, ANN, Raspberry Pi.

I. INTRODUCTION

The prevalence of traffic congestion has become a common occurrence in many areas, resulting in significant consequences for various incidents. One of the most critical impacts is the obstruction of ambulances in traffic jams, leading to the loss of precious lives. To effectively reduce the death rate resulting from traffic jams, it is imperative to implement an automated system that can swiftly clear the traffic using signals easily and efficiently. By adopting automated traffic signal systems, priority can be given to emergency vehicles, particularly ambulances, allowing them to navigate through congested areas without unnecessary delays. Furthermore, implementing sophisticated traffic surveillance tools, including cameras and sensors, enables real-time monitoring of traffic conditions [1]. This data collection allows for dynamic adjustments to the signal timings, optimizing traffic flow and minimizing the likelihood of ambulances becoming trapped in heavy traffic. However, by implementing automated traffic signal systems that prioritize the passage of emergency vehicles, we can effectively reduce the death rate caused by delayed ambulances.

II. RELATED WORKS

We have come across several similar projects related to the problem and they have their own advantages and disadvantages. To state a few related works are the following: They collected Emergency vehicle siren sound and road noise datasets for ambulance acoustic monitoring. Using the dataset they developed a deep learning (MLP-based) model and trained to predict the ambulance presence on the roads [2]. The project includes developing a website" Health Card " for doing registration about the medical history of all citizens and sending data to the particular hospital before the ambulance reaches. In the second phase, they are trying to control signals for the prior ambulance. They use technologies like GPS and RFID. Discussed RF based clearance. The location of the RF transmitter was varied incrementally until the RF receiver was not able to detect the RF transmitter. RF receiver receives signal from RF transmitter. PIC triggers the traffic light to green. If the ambulance passes the junction the traffic light operation goes back to normal else delay is added [3]. Existing traffic light systems are often locally programmed, leading to inefficiencies and difficulty in determining their operational status.This paper proposes a centralized control topology for a wireless network of traffic lights, aiming to improve overall traffic management in smart cities. Munasinghe, S. A. The system aims to detect vehicles and pedestrians, prioritizing emergency vehicles to address congestion and enhance road safety.This approach not only tackles typical traffic issues but also addresses the challenges posed by the ongoing pandemic, emphasizing the importance of efficient ambulance prioritization [4]. This paper proposes a solution focusing on controlling green lights at four-way junctions during peak and non-peak hours.An image processing system using MATLAB analyzes real-time videos from installed cameras, calculating vehicle density. Mansur, A. proposed robust solution based on Vehicle-to-Infrastructure (V2I) technology aims to improve road safety and traffic management.The system utilizes Dashboard Traffic Light (DBTL) sensors in vehicles, communicating with an intersection control station through the V2I network.

979-8-3503-5268-9/24/$31.00 ©2024 IEEE

III. DATA ACQUISITION

a. *DATA PREPROCESSING:*

Data cleansing and filtration are crucial steps in preparing audio and image datasets for deep learning model training. These processes involve the removal of noise and the application of preprocessing techniques to enhance the quality of the data, ultimately making it more suitable for efficient model training. Proper data preprocessing is crucial for the successful creation of a robust model. Some Augmentation techniques are also carried out to increase the quality of the dataset for various circumstances [5].

S.no.	Table 1. Preprocessing techniques	
	Techniques	***Values***
1	Resizing	640x640
2	Contrast adjustment	Auto
3	Flipping	-
4	Rotation	-15° to +15
5	Shearing	±10°
6	Blur & Noise	2.2px

b. *IMAGE DETECTION DATASET:*

For training the image detection model, a dataset consisting of 500 images containing ambulances and other vehicles in various circumstances was utilized. Each image was meticulously annotated using labeling tools to ensure that the model had accurate and relevant information for learning. The labels are split into two categories: 1. ambulance 2. others(which is not ambulance) to improve the performance of classification. Following the annotation process, the model was trained on this labeled dataset [6]. To assess the model's performance and accuracy, a separate set of 100 images was reserved for validation.

c. *SOUND DETECTION DATASET:*

Training the audio detection model required a large and diverse audio dataset. A comprehensive siren dataset was collected from Kaggle, serving as the primary data source. The audio files are then converted to our preferred .wav format for the process of training. During the training phase, the audio data was segmented into discrete segments. These segments were then transformed into spectral representations, comprising discrete values, to facilitate effective model training and testing. The model's accuracy was assessed through rigorous testing procedures.

d. *DATA INTEGRATION OF AUDIO AND IMAGE MODELS:*

To enhance the efficiency of ambulance detection in real-time scenarios, an integrated approach was adopted, combining both the image and audio detection models into a single, unified model. This integration allows the model to classify and detect ambulances with greater accuracy and precision. By combining the capabilities of image and audio detection, the model performs more efficiently than if these were separate entities [7].

IV. EXPERIMENTAL APPROACHES

A.) *SYSTEM ARCHITECTURE:*

The Architecture *Fig. 1* involves various steps which are mentioned below. The flow of the system is given as the System Architecture which gives a clear workflow of the entire process. The flow consists of an input detection module, a server module where the deep learning modules infer the input data, and the controller module which controls the traffic signal according to the conditions.

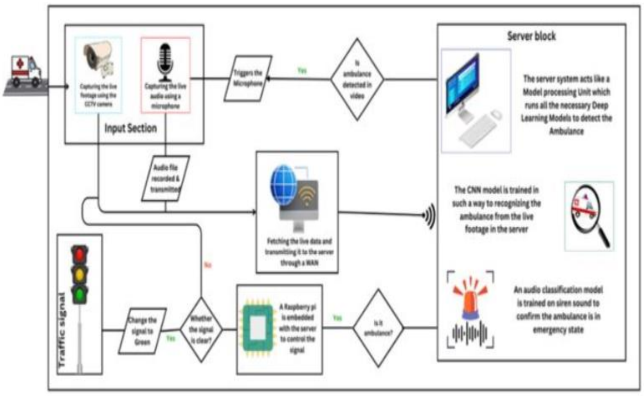

Figure 1: *System Architecture/Flow Diagram*

B.) *AMBULANCE DETECTION BY VIDEO:*

A Deep learning model is trained in a way to detect the ambulance incoming. Various ambulance data sets are given as input and the model is trained using the Python module open CV in which an algorithm called YOLO which is based on the CNN deep learning algorithm is used for detecting the ambulance in a live format. In this project the latest version of YOLO which is YOLO v8 was used for better performance and thus it widely enhances the usage of real-time data [8]. A CCTV camera is fixed at the signal poles which is responsible for monitoring the ambulance entry. The live footage is directly transferred to a server which processes the video for detecting ambulances using the pre-installed YOLO model. In the case of an ambulance detected, the microphone is triggered to receive the audio. The model consists of various layers to classify the image.

b.1 YOLO architecture

YOLO stands for You Only Look Once is a single-shot object detection algorithm that was specifically designed to solve the problem of real-time detection of objects and instances. It is an algorithm built on the architecture of fully convolutional neural networks(CNN) to process the input images and learn features to further process object detection [9]. At the initial stage of designing the YOLO algorithm involves 20 convolutional layers which performed a low result.

To improve the performance of the algorithm various versions of it were introduced from the timeline of creation. So far, eight versions of YOLO have been designed which is depicted in *Fig. 2*, the latest version of YOLO v8 outperforms all other versions which almost contain 53 convolutional layers [10]. These version's algorithms have different performance levels.

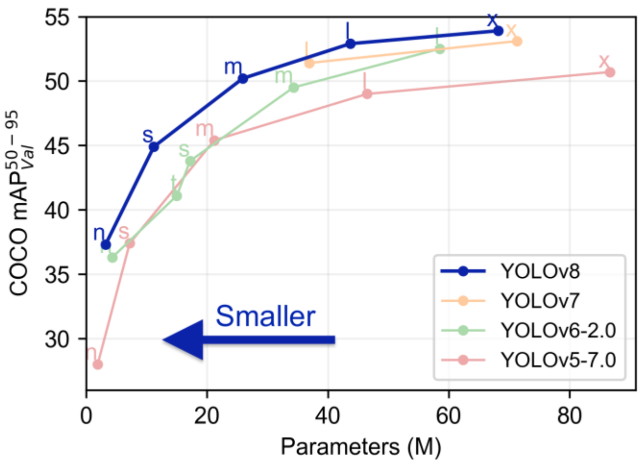

Figure 2: *YOLO performance mapping*

b.2 Model testing

The model was trained in the virtual environment called Google Colab which provides a higher-performing TPU(Tensor Processing Unit) which gives efficient and faster hardware for the training process. The trained model was then inference with testing images which are shown in *Fig. 3* and live videos which accurately detected the ambulance in the images and video with bounding boxes and confident scores [11]. The output of the tested images is displayed below.

Figure 3: *Tested images*

C.) *SIREN SOUND DETECTION:*

A Model is built to classify the incoming sound at the signals. After detecting the ambulance, the audio is captured and sent to the server. At the server, the model is built using ANN architecture in addition to tensorflow. The Librosa library is used in analyzing the audio file. Using the MFCC (Mel-frequency cepstrum) technique, the input audio file is transformed into a spectrum for further classifications. The model actually classifies the noisy traffic sound and the ambulance siren sound. The model is trained with various attributes and it is trained for many datasets for many circumstances. The noisy sound and the siren sound with various noises like other vehicle horns, bird chirping, etc., is given as input data set and those are filtered and decoded to identify the particular ambulance siren sound [12]. A sound recognizing sensor is placed over a traffic signal, which receives the sound and runs the model to detect the siren sound of the ambulances.

D.) *CONTROL AND COMMUNICATION:*

i. *Integrate the features in the Raspberry pi:*

As referred from the review paper [13] of Raspberry pi's uses and characteristics, it is used as the control unit of this entire process. The various processes and controls are integrated and controlled by a single device which is the microprocessor Raspberry pi model. The Raspberry pi is programmed in such a way to control the overall action in step-by-step procedure in the raspberry pi environment. The main work of the raspberry pi is to control the traffic signal system based on the communication with the server. It also transmits the data to the server. All the devices such as the hardware devices are connected to raspberry pi to ensure the working of the product.

ii. Building communication between the server and the product:

A server is set up nearby the signal through a LPWAN which sends the live data from the signal hardware to the signal. This server has all the Deep learning models running [14]. Based on the data from the hardwares (camera and microphone), the raspberry pi is set to feed that data into the server, the model first searches for the ambulance in the video footage. If the ambulance is detected then the mic is triggered and audio is recorded from the signal and fed into the model to acknowledge the emergency of the ambulance. Then, the signal is cleared through the controller which was programmed to control the traffic signal to clear the ambulance on that road.

E.) *DEVELOP INTO PRODUCT:*

For building this project as a commercial product in the market, we need to fulfill some steps. Conducting some deep research and development on this project and setting up hardware infrastructure [15]. To commercialize this product, we need to prepare a business model for the product. The model considers the following aspects: scalability, flexibility, robustness, user-friendly, cost-effectiveness, reliability, marketing and so on.

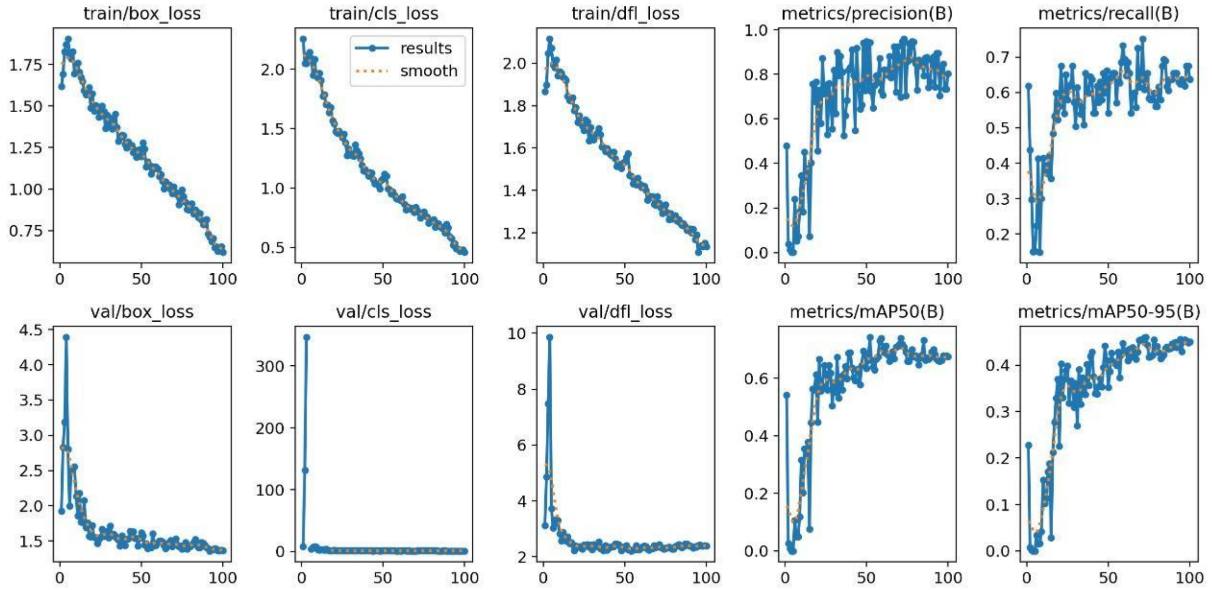

Figure 4: *Results of performance of graphs*

V. EXPERIMENTAL RESULTS

The model for YOLO is trained with a dataset containing 500 images that are split in a proportion of 80% as training images and 20% as validation image sets. The training process is extended to 70 epochs with 8 batches per epoch which gave an accuracy of 92.8% in mAP-50(mean average precision) and 65% in mAP50-95 for detecting the ambulance objects and an overall accuracy of the model is 70% in mAP50 which can be depicted from the plot in *Fig. 4* from which all the metrics such as accuracy and loss of the model while training and validation process are displayed. The loss functions in the plot can be seen as a decreasing graph over the number of epochs thus decreasing the loss by increasing the performance. Further, training with more datasets and several epochs can improve the model. In *Fig. 4*, the first six plots show the individual box and class losses of the training and testing phases.

The image *Fig. 5* shows the confusion matrix of the model and how accurately it classifies the given images for different classes. It compares the predicted and true values through the predicted vs true confusion matrix.

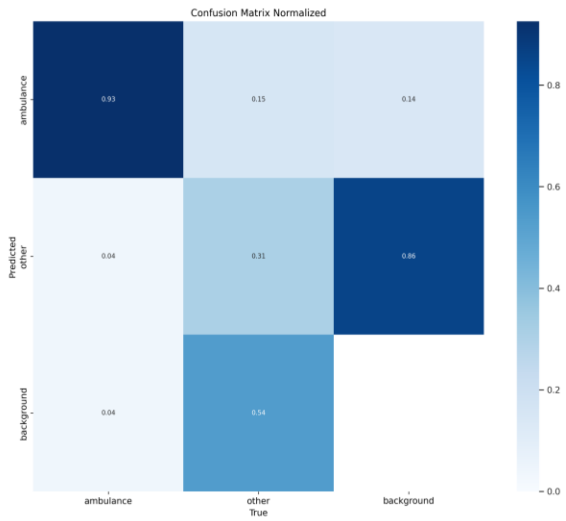

Figure 5: *confusion matrix*

In our state-of-the-art audio classification model, we utilized TensorFlow and Librosa to excel in siren detection. Using normalized MFCC features, a single vector streamlines training and predictions. Our model achieves a remarkable 84% accuracy. Configured with 50 epochs and a batch size of 20, it sets a new standard in efficient and accurate siren sound identification. To make it so elegant an interface is built using an open-source library in python for detecting the audio. The Fig. 6 shows the interface given for adding the audio track and detecting the class whether it is a siren or not.

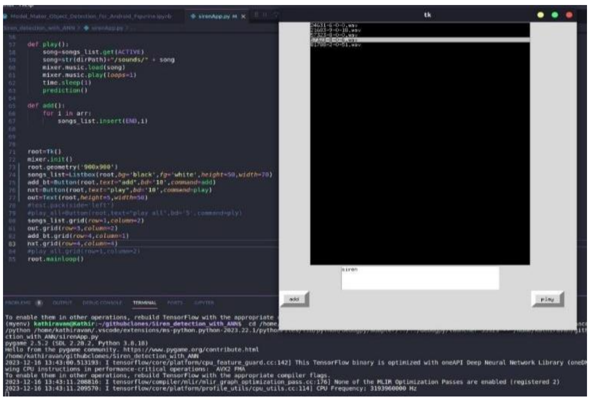

Figure 6: Screenshot of siren detection

VI. CONCLUSION

Since traffic accidents occur frequently in busy cities, especially at important intersections, many people are especially affected, emergency services are trapped in traffic accidents, and many people lose their lives due to the slowness of the ambulance. Our plan to overcome this problem is to ease the traffic and ensure that the ambulance reaches the hospitals on time. How the automatic signal generator works and how it is used is explained in detail in the paper. Using deep learning and embedded systems, the project can identify ambulances and replace the best detection currently available. Deep learning is used to train the model to identify the sound of the ambulance, using a microphone (ETS omnidirectional microphone) to identify it as one of the inaccessible ones, and CCTV cameras to capture the video input. In conclusion, the project can further be developed to seek partnership with the government to implement such automated systems. We believe that by implementing this project we can save many lives caused by ambulance delays, improve transportation and support the development of sustainable cities.

VII. REFERENCES

[1] Usaid, Muhammad, Muhammad Asif, Tabarka Rajab, Munaf Rashid, and Syeda Iqra Hassan. "Ambulance siren detection using artificial intelligence in urban scenarios." *Sir Syed University Research Journal of Engineering & Technology* 12, no. 1 (2022): 92-97.

[2] Reddy, R. Gangadhar, G. S. Kumar, M. P. Ananth, S. Anil, and M. Darshan. "GREEN WAY FOR AMBULANCE BY AUTOMATIC CONTROL OF TRAFFIC LIGHTS."

[3] Hashim, Norlezah, Fakrulradzi Idris, Ahmad Fauzan Kadmin, and Siti Suhaila Jaapar Sidek. "Automatic traffic light controller for emergency vehicle using peripheral interface controller." *International Journal of Electrical and Computer Engineering* 9, no. 3 (2019): 1788.

[4] Saradha, B. Janani, G. Vijayshri, and T. Subha. "Intelligent traffic signal control system for ambulance using RFID and cloud." In *2017 2nd International Conference on Computing and Communications Technologies (ICCCT)*, pp. 90-96. IEEE, 2017.

[5] Sivaraj, S. N., K. Vigneshwaran, S. Vigneshwaran, and M. Vishnu Priyan. "Iot Ambulance With Automatic Traffic Light Control." *SSRG International Journal of Industrial Engineering (ICRTECITA)* (2017): 12-18.

[6] Kadmin, A. F., A. Z. Jidin, Abu Bakar, KA A. Aziz, and W. N. Abd Rashid. "Wireless voice-based wheelchair controller system." *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* 8, no. 7 (2016): 117-122.

[7] Geethapriya, S., N. Duraimurugan, and S. P. Chokkalingam. "Real-time object detection with Yolo." *International Journal of Engineering and Advanced Technology (IJEAT)* 8, no. 3S (2019): 578-581.

[8] Oliveira, Luiz Fernando Pinto de, Leandro Tiago Manera, and P. D. G. Luz. "Smart traffic light controller system." In *2019 Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*, pp. 155-160. IEEE, 2019.

[9] Munasinghe, K. D. S. A., T. D. Waththegedara, I. R. Wickramasinghe, H. M. O. K. Herath, and V. Logeeshan. "Smart Traffic Light Control System Based on Traffic Density and Emergency Vehicle Detection." In *2022 Moratuwa Engineering Research Conference (MERCon)*, pp. 1-6. IEEE, 2022..

[10] Balu, Shibin, and C. Priyadharsini. "Smart traffic congestion control system." In *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 689-692. IEEE, 2019.

[11] Miyim, Abubakar M., and Mansur A. Muhammed. "Smart traffic management system." In *2019 15th International Conference on Electronics, Computer and Computation (ICECCO)*, pp. 1-6. IEEE, 2019.

[12] "Emergency vehicle detection on heavy traffic road from cctv footage using deep convolutional neural network." In *2019 international conference on electrical, computer and communication engineering (ECCE)*, pp. 1-6. IEEE, 2019.

[13] Dhatrak, A. S., and S. T. Gandhe. "Automatic traffic signals in smart cities for speedy clearance of emergency vehicles." In *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, pp. 1-6. IEEE, 2018.

[14] Armstrong, Anu, Anwar Sahil CA, Finla NA, and Godwin Simon. "Emergency Ambulance Service with Mapping." *Grenze International Journal of Engineering & Technology (GIJET)* 8, no. 2 (2022).

[15]. Nayyar, Anand, and Vikram Puri. "Raspberry Pi-a small, powerful, cost effective and efficient form factor computer: a review." *International Journal of Advanced Research in Computer Science and Software Engineering* 5, no. 12 (2015): 720-737.