DOI: https://doi.org/10.4038/engineer.v57i1.7613

ENGINEER - Vol. LVII, No. 01, pp. [85-94], 2024

Audio-Based Vehicle Detection System: Enhancing Safety for Bicycle Riders

R.A.L.B. Wimalasena and D.D.M. Ranasinghe

Abstract: To enhance the safety of bicycle riders an android mobile application was developed that utilizes real-time audio analysis to detect approaching vehicles and alert the riders when there is a safety issue. The application employs a CNN lite model to detect the presence of the vehicle. The CNN model was trained on a dataset of 0.5-second audio clips of vehicle and non-vehicle sounds. The root mean square value of the audio clip is calculated to figure out the approaching vehicles. Based on the model prediction and the root mean square value of amplitude, the application issues alerts to the user. Three alert levels are defined, ranging from level one for low amplitude to level three for high amplitude. The alert types are audio, visual and vibration. Users can customize and adjust the alert types and threshold values within the predetermined range according to their preferences. The evaluation of the model revealed favourable results, as indicated by low loss and high recall and precision metrics, affirming the efficacy of the model for accurately detecting vehicles. The ability of the model to proactively detect and alert bicycle riders of incoming vehicles, enable the users to take timely actions to prevent potential hazards and positions the application as a crucial and effective life-saving tool.

Keywords: Bicycle accident, CNN, Mel-spectrogram, Collision detection

1. Introduction

Bicycle riding, as an eco-friendly and sustainable mode of transportation, plays a critical role in promoting a greener future and contributes to the achievement of sustainable development goals. Embracing this eco-friendly transportation alternative actively contributes to environmental preservation endeavours and cultivates a society that is healthy, sustainable, and resilient. Yet, cyclists pose a great danger due to road accidents, and ensuring the safety of cyclists is of utmost importance if cycling is to promote as a sustainable and a safe mode of transportation. According to [14], annually in the United States of America, nearly 1,000 bicyclists die and over 130,000 are injured in crashes that occur on roads, which account for about 2% of total deaths out of all the people who die in a crash involving a motor vehicle. In Sri Lanka, according to [16], in the year 2022, a total number of 226 bicyclists have died which accounts for about 9% of out of total deaths due to road accidents. Hence it is clearly evident that the safety of bicycle riders is of most importance, considering that the most common cause of bicycle accidents is collisions with vehicles and near misses [1]. Near misses often occur due to concerns associated with cycling in traffic, such as encountering inattentive or aggressive drivers and being cut off by turning vehicles. To ensure the safety of bicycle riders, it is crucial to implement vehicle detection systems that can effectively prevent collisions and mitigate potential risks, thereby further improve bicycle riding as a safe mode of transportation.

Several solutions are currently available to address this issue, including the utilization of computer vision techniques to detect vehicles, as well as the incorporation of sensors such as LIDAR, ultrasonic, and laser technology. However, these existing systems often pose challenges for the public due to the requirement of attaching an extra device to the bicycle such as camera or sensor devices, which limit their accessibility to a wider audience incurring an additional cost for the riders.

The solution proposed in this research for the above-mentioned problem is to develop a mobile application that uses a deep learning

Mr. ***R.A.L.B. Wimalasena****, Student Member of IESL, B.Sc. Eng. (Hons) (OUSL), Department of Electrical and Computer Engineering, Faculty of Engineering Technology, The Open University of Sri Lanka.*
Email:ymrlbw@gmail.com
https://orcid.org/0009-0006-6473-8233
Dr.D.D.M. Ranasinghe, *B.Sc.(Statistics)(Hons) (Colombo), PhD (OUSL), Senior Lecturer, Department of Electrical and Computer Engineering, Faculty of Engineering Technology, The Open University of Sri Lanka.*
Email:ddran@ou.ac.lk
https://orcid.org/0000-0002-3890-3278

model to detect and classify approaching vehicles using ambient sound. The app will use the built-in microphone [2] of the smartphone of the user to capture audio data, which will then be processed into spectrograms. The deep learning model is trained on a large dataset of vehicle and non-vehicle sounds to accurately identify and classify the presence of nearby vehicles. To ensure real-time processing and minimal delay, the app will use on-device processing rather than sending the data to a remote server for analysis. The user will receive instant visual and auditory alerts, as well as vibration feedback, if a vehicle is approaching. When an approaching vehicle is detected, the increasing value of the root mean square of amplitude indicates the closeness of the approaching vehicle. The system will provide a simple and effective tool for improving bicycle rider safety by providing real-time information and incident reporting. The system aims to enhance the safety of bicycle riders by providing real-time alerts about nearby vehicles. The system will offer three different levels of alerts based on the root mean square values of the amplitude, which will help riders to take appropriate actions to avoid collisions.

The rest of the paper is organized as follows: Section 2 describes the related concepts and theories, along with a review of relevant literature. A detailed explanation of the methodology, including dataset preparation, preprocessing, and the deep learning model is given in Section 3 and Section 4 presents the results obtained. Section 5 discusses the findings while analysing the limitations and conclusions are given in Section 6.

2. Literature Review

The primary focus of the literature review revolves around analysing research work on detecting vehicles utilizing smartphone technology. Numerous studies have investigated various aspects of this topic, ranging from the use of sensors and microphones in smartphones to the application of deep learning techniques for vehicle identification. These works collectively contribute to a comprehensive understanding of the challenges and opportunities in the field of vehicle detection using smartphones. In this section, we delve into the specific studies and approaches that have paved the way for this research, shedding light on their methodologies, findings, and the unique contributions they offer to the realm of smartphone-based vehicle detection. The literature review initially looks at the behaviours of bicycle riders and move on to analysing similar systems with physical sensors, then progress to reviewing similar systems with deep learning methods.

E-bike risky riding behaviours are reviewed in [7]. The authors have analysed the issue from three perspectives, namely, accident characteristics, user behaviour, and prevention/intervention. Research methods include surveys and models, revealing risky behaviours like lane violations, speeding, running red lights, and illegal manoeuvres. Variations in behaviour occur based on factors like gender, age, abilities, responsiveness, and effective prevention measures involving license registration, quasi-drive systems, improved environments, and safety training.

Bicycle riders face a higher risk of head trauma compared to motorists [15], and bicycle-related accidents occur due to crashes, rider errors, and environmental hazards. To address this, bicycle accident detection system in [3] utilizes hardware modules with a MARG sensor-based system to measure 24 features related to riding status and falls, achieving a 95.2% accuracy in detecting accidents due to falls during cycling.

A magnetic sensor-based system which is inexpensive, non-intrusive, and energy-efficient in identifying vehicles and measuring their speed, was proposed in [6]. The measuring algorithm, which consists of an adaptive auto-offset algorithm, a power-based detector, and a correlation-based speed estimator, is used in this study for the analysis. Real-world measurement examples are used to demonstrate the performance accuracy of the suggested approach.

The system in [8] labelled as a "Cyber-Physical Bike" uses a video camera attached to the saddle of the bicycle in backward direction and an embedded microcomputer to detect arriving vehicles and alert the bicycle user. The study allows the evaluation of both the probability of a collision as well as the detection of cars. However, since a camera or a sensor device is required for this system, it is challenging for regular people due to the extra cost and the burden of attaching additional hardware to the bike. The objective of the research described in this paper is to develop a vehicle detection system utilizing smartphone as a standalone device, eliminating the need for external hardware components or sensors.

The research paper by Guofeng Ma et al. [5] presents a framework for video object detection using the Single Shot MultiBox Detector (SSD). The method improves detection by fusing feature maps from different layers and enhancing the detection of small cars. It achieves comparable performance to state-of-the-art video object detectors on the VID dataset, increasing the consistency of video vehicle detection results.

ImageNet-pretrained deep CNN models were used for audio classification in [9] and [10] effectively, despite the differences between audio spectrograms and ImageNet images. Pretrained weights outperform randomly initialized weights, with visualization of gradients revealing what the CNNs learn. The study achieves state-of-the-art performance on the ESC-50 and UrbanSound8K datasets, obtaining validation accuracies of 92.89% and 87.42%, respectively, using an ensemble of ImageNet pretrained DenseNet models.

The study By Tabei et al. [4] aimed to identify vehicles through acoustic detection using a general microphone and Matlab for data analysis. However, setting a threshold for noise avoidance and false alarms limited the success rate to approximately 70%. In contrast, the research presented in this paper employs the built-in microphone of a smartphone and deep learning techniques for vehicle identification, deviating from the reliance on external microphone and Matlab analysis.

Next, from the literature, the methods used for classifying the sound waves using deep learning techniques will be analysed. The paper by Nikhil and Morris [11] addresses the need for efficient diagnosis and monitoring of respiratory diseases by proposing a hybrid neural model that utilizes deep learning techniques. The model combines a convolutional neural network (CNN) for feature extraction from lung sound spectrograms with a long short-term memory (LSTM) network for temporal dependencies and classification.
The proposed model achieves state-of-the-art results on the ICBHI 2017 Respiratory Sound Database, demonstrating its effectiveness in classifying four types of sounds of the lung with varying data splitting strategies.

The work in Acoustic Scene Classification (ASC) addresses and introduces SubSpectralNet, a novel model that incorporates frequency band-level differences to capture discriminative features in soundscapes [12]. The proposed approach utilizes mel-spectrograms and trains a convolutional neural network (CNN) on band-wise crops of the input representations. Evaluation on the DCASE 2018 Challenge dataset demonstrates that the SubSpectralNet model achieves a significant +14% improvement in classification accuracy compared to the baseline system.

Study by Ming Liu [2] emphasizes the importance of using sensors in smartphones, distinguishing them from traditional computing devices. The research suggests that the microphone, camera, accelerometer, and gyroscope, known for their accuracy, are widely utilized in various applications, and their usage generates substantial data. Hence this research leverage smartphone sensors that enables the collection of highly accurate data that will potentially yield to obtain more precise results.

In the context of this research, the proposed approach entails the construction of a vehicle detection system by analysing ambient audio. Recognizing the research gap in audio-based vehicle detection systems for cyclists, this work proposes a novel approach leveraging readily available smartphone sensors (microphones). Traditional sensor-based solutions often burden users with additional hardware, whereas image-based systems struggle with poor lighting. Hence this research is inspired by the promising results in CNN-based audio analysis and the reliability of smartphone sensors (microphones) for developing a solution for identifying vehicles having a potential threat for the bicycle riders.

3. Methodology

The proposed system comprises several modules, namely, deep learning model, mel-spectrogram, and approaching vehicle detection algorithm. The overall system design is given in Figure 1. The core function of the deep learning model is to identify vehicles within ambient audio clips. As a fundamental preprocessing step, the ambient audio is transformed into a mel-spectrogram before feeding into the deep learning model. In a nutshell, the overall system functionality is as follows.

The proposed system is a mobile application that uses the microphone of the mobile to analyse the ambient sound using a CNN model. The system captures audio data using the built-in microphone, converts it into Mel-spectrograms, and then uses the CNN model to identify the presence of vehicles. The root mean square (RMS) value of the amplitude is considered as an indicator to determine whether a vehicle is approaching or not. When an approaching vehicle is detected, the user will be alerted through visual and auditory signals, as well as vibration. The app uses a two-way approach to detect approaching vehicles. It analyses 0.5-second audio clips via both a CNN model and RMS value calculation. The CNN identifies vehicle presence within the audio, while the RMS tracks intensity changes. Only when both the CNN detects a vehicle and the RMS shows a sustained increase (with a threshold value of 500 than the previous value), the app trigger vibratory alerts with varying intensity.

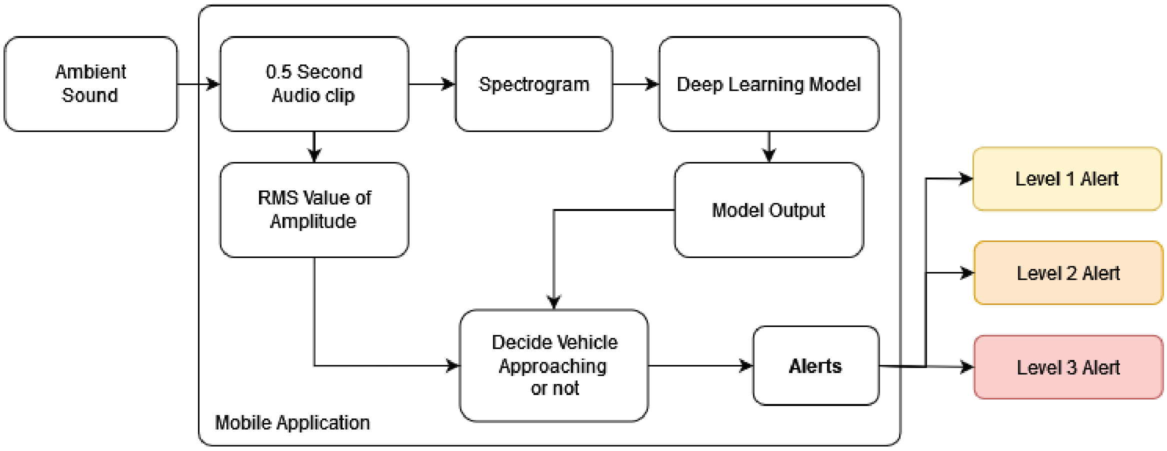

Figure 1 - Overall System Design

3.1 The Dataset

The audio data was collected from both rural and urban areas to ensure a diverse and comprehensive dataset. To ensure comprehensiveness, data collection spanned a total of 90 minutes, with recordings intentionally capturing a variety of vehicle sounds, including cars, motorcycles, and trucks. To ensure consistency and to minimize confounding factors, recordings were exclusively made during clear weather conditions. The recorded audio clips were trimmed to 0.5 second clips and sorted as vehicle and non-vehicle audios. The dataset includes 2000 vehicle audio clips and 2000 non-vehicle audio clips. To amplify the robustness of the dataset, data augmentation techniques were applied.

The data collection process involved the use of a Huawei Honor X6 smartphone, also known as Huawei Nova 7i, which served as the primary device for recording audio samples. The smartphone was equipped with a high-quality built-in microphone capable of capturing audio with sufficient clarity and fidelity.

3.2 Data Preprocessing

The audio data was collected using a mono channel configuration at a sampling rate of 44.1 kHz. Down sampling the collected audio from 44100 Hz to 16000 Hz during the preprocessing stage was made to reduce the computational complexity and memory requirements while preserving the relevant frequency information for vehicle detection. Additionally, the lower sampling rate allows efficient processing and analysis of the audio data without sacrificing the essential characteristics that are necessary for accurate detection of vehicle-related features. The waveform is then processed to have a fixed length of 7900 samples. If the waveform is shorter than 7900 samples, it is padded with zeros at the beginning to match the desired length. This step standardizes the lengths of the audio clips for further analysis. The pre-processed audio clips are used to obtain Mel-spectrograms. Spectrograms are visual representations of the frequencies present in an audio signal over time, which is a 2D plot where the x-axis represents the time, and the y-axis represents the frequency. The intensity or colour of each point in the plot represents the magnitude or power of the corresponding frequency component. Figures 2 and 3 depict the audio spectrograms obtained for vehicle and non-vehicle scenarios, respectively.

First, the waveform data is resampled from the original sample rate to 16000 Hz (between 8000Hz and 16000Hz, 16000Hz gives the best results). Then, the resampled waveform is transformed into a spectrogram using STFT (Short-time Fourier Transform), which computes the magnitude spectrum using a window size of 512 and a stride of 256. Finally, the spectrogram is converted into Mel spectrograms using Mel-scale, where the frequency axis is transformed into Mel scale with 128 Mel frequency bins. Additionally, it ensures that the Mel spectrogram captures the perceptually relevant frequency components from the audio waveform directly, aligning with the research objective of detecting vehicle-related features efficiently and accurately.

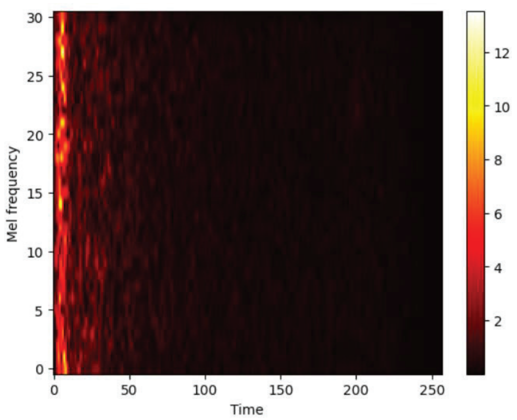

Figure 2 - Vehicle Audio Spectrogram

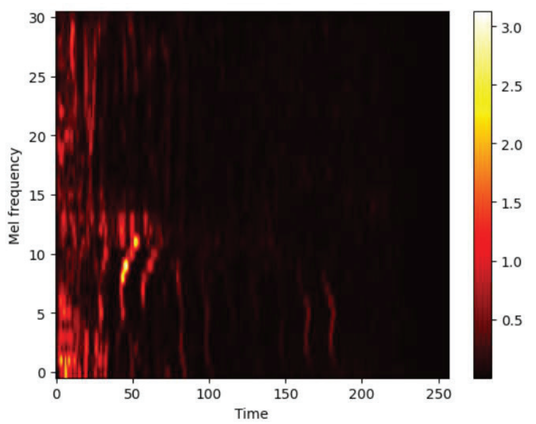

Figure 3 - Non-vehicle Audio Spectrogram

3.3 Deep Learning Model

Recurrent Neural Networks (RNNs) excel at capturing temporal dependencies in sequential data, making it a natural choice for tasks like music genre classification where the order of sounds matters. However, the performance can be impacted by long audio sequences and training is computationally expensive. In this scenario, focusing on capturing spatial features within short clips seemed more relevant, making CNNs preferable. Traditional machine learning method of Support Vector Machines (SVMs) model offers strong interpretability and handle class imbalance well. However, its reliance on handcrafted features is a limiting factor compared to the automatic feature extraction capabilities of deep learning models. The ability of CNNs to learn optimal features directly from spectrograms is a significant advantage for this scenario. The labelled sound data was divided into 0.5 seconds intervals, and the frequency features of the sound spectrum were extracted using Short-time Fourier Transform (STFT) [13]. Spectrograms were obtained from the STFT data. STFT parameters were set as follows: a window size of 512, and a hop size of 256. The window size determines the length of the window used for analysing the audio signal at a given time, while the hop size determines the amount of overlap between consecutive windows. The choice of these parameters is crucial as it balances the time-frequency resolution trade-off. A larger window size provides better frequency resolution but sacrifices time resolution, while a smaller window size allows for better time resolution but may result in lower frequency resolution. Similarly, the hop size determines the amount of information overlap between adjacent windows, influencing the temporal smoothness of the resulting spectrograms. For the deep learning model, a Convolutional Neural Network (CNN) was chosen due to promising results in CNN-based audio analysis.

The CNN model architecture was defined using the TensorFlow Keras library, and consisted of several layers including Lambda, Conv2D, MaxPooling2D, Flatten, and Dense. Input layer readily accepts audio spectrograms of dimensions 31x257. Lambda layer, deftly expands the dimensionality of the spectrogram to accommodate subsequent convolutional layers, effectively preparing the stage for feature extraction. First convolutional layer unleashes 64 filters, each 3x3 in size, to extract foundational features from the spectrograms, indicative of vehicle presence. First max pooling layer, down samples the feature maps by a factor of 2, retaining essential information while reducing computational demands, which improves the efficiency. Second convolutional layer activates 128 filters, each 3x3 in size, to delve deeper into the spectrograms to extract more intricate features for refined vehicle detection. Second max pooling layer down samples by a factor of 2, and further streamline computations while preserving the essence of learned features. The third convolution layer contains 256 filters, each 3x3 in size, to uncover the most profound patterns within the spectrograms using the acoustic signatures of vehicles. Third max pooling layer does the final down sampling, condensing the feature maps while retaining the most informative elements. The flatten layer transforms the multidimensional feature maps into a single, streamlined vector, preparing the extracted knowledge for further analysis. The dropout layer uses only 50% of neurons during the training to make the model resilience against overfitting and ensure the adaptability of the model to unseen data.

During model evaluation, 5-fold cross-validation confirms precise detection of approaching vehicles at closest approach with a F-value of 93.5% and various other metrics such as loss, precision, recall were utilized to assess the performance of the trained CNN model.

indicating the ability of comprehensive classification capability of the model.

3.4 Developing Mobile Application

The Android application was developed using Kotlin language and Android Studio IDE. The app is designed to listen to ambient sounds in real-time and detect approaching vehicles. The app will issue three levels of alerts depending on the proximity of the approaching vehicle. The sound analysis is performed using a CNN lite model that is optimized for mobile devices. The TensorFlow interpreter is used in the app to preprocess and feed the audio clips to the CNN model and fetch the model output. The root mean square value of the audio-clip is also calculated in real-time from worker thread in android application. The app will start by recording ambient sound for 0.5 seconds and preprocessing the recorded audio-clip. The pre-processed audio-clip is then fed to the CNN model, for prediction. The app has three alert levels and three threshold values for root mean square values of amplitude.

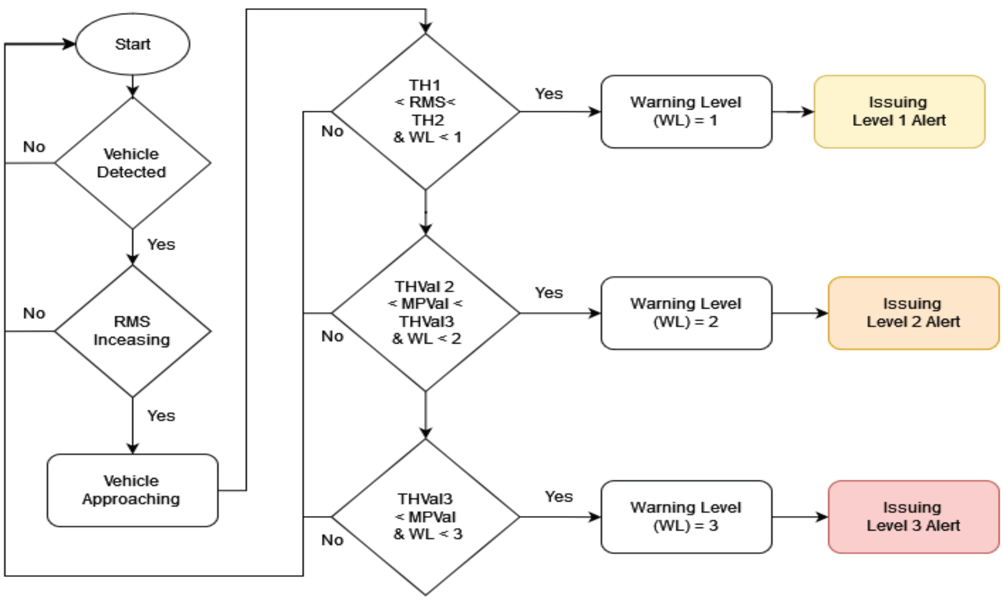

Figure 4 - Approaching Vehicle Detection Method

Loss function was employed as a metric to quantify the discrepancy between predicted and actual labels, providing insight into the model's training process. Higher precision and the recall confirm the ability of the model to correctly identify positive samples thus

Alerts are designed to vibrate the phone with one vibration for level one, two vibrations for level two, and three vibrations for level three. The entire process of detecting approaching vehicles and issuing vibrations is given in Figure 4. Users can choose which alerts they want to receive by disabling the unwanted alert types. Advanced options allow users to edit the threshold values as per their preference. The app will have a user-friendly interface with easy navigation to allow users to quickly access various features of the app.

The Android application developed in the research incorporates various UI/UX design considerations to ensure a seamless and user-friendly experience. The UI design features a single activity, facilitating straightforward navigation and minimizing complexity.

The interface includes two switches for enabling or disabling audio and vibration alerts, offering control for the users over their notification preferences. Additionally, three sliders are provided to adjust threshold values, allowing users to customize the sensitivity of the vehicle detection system. The UI layout dedicates 40 percent of the screen to visual alerts, emphasizing their significance as shown in Figure 5 and 6.

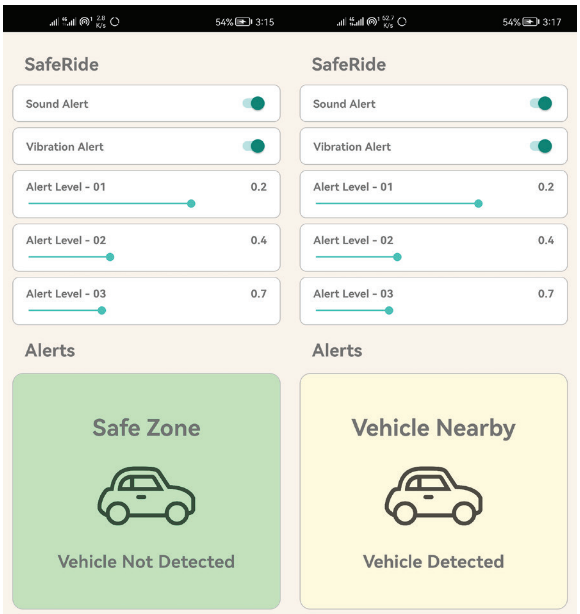

Figure 5 - Frontend of the Mobile Application (Idle and Level 1 alert)

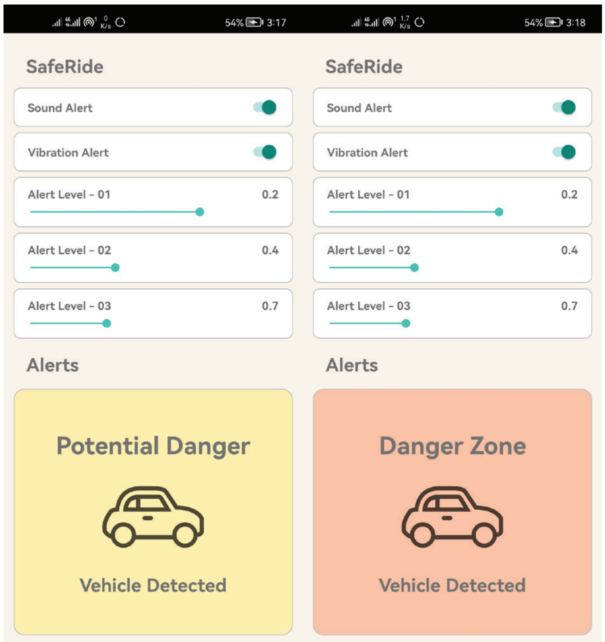

Figure 6 - Frontend of the Mobile Application (Level 2 and 3 alerts)

The design employs light colours to enhance visibility and readability under direct sunlight, optimizing usability in outdoor environments. Using a simple and intuitive design approach, the application prioritizes ease of use and reduces cognitive load for users. The combination of switches, sliders, and a visually focused interface ensures flexibility, personalization, and effective communication of information to users, resulting in an enhanced UI /UX experience.

4. Results

The mobile app developed utilizes the power of audio rather than relying on visual cues and overcomes the inherent limitations of image-based methods. This novel approach enables robust vehicle detection capabilities regardless of ambient light conditions, ensuring effectiveness even in nighttime scenarios. Several testing and validation procedures were carried out to evaluate the performance of the trained CNN model. The model was tested on a separate dataset containing sound data of vehicles that were not included in the training dataset. 5-fold cross-validation confirms precise detection of approaching vehicles at closest approach with a F-value of 93.5%.

The trained deep learning model was integrated into a mobile application for bicycle riders. The application captures audio data above 100Hz in real time and alerts the user about the potential threat from the approaching vehicle. The system has three levels of alerts. The first level of alert is issued when the vehicle was detected for the first time with a level one alert with one beep tone with one vibration effect and a visual effect with a colour change in the app as shown in Figure 5. If the distance to the detected vehicle approximately becomes less than 5 meters, then the system issues a second level alert with two beep tones and two vibrations, again with a colour change in the app as indicated in Figure 6. When the prediction value is above the third level threshold, it includes three beep tones and three vibrations with a colour change of the UI indicating the rider is in the danger zone.

The efficacy of the developed mobile application was evaluated through a questionnaire with human bicycle riders, a sample size of 20, representing a diverse range of cycling habits. The cyclists mainly used handlebar-mounted smartphones in the evaluation of the mobile application and each cyclist was asked to use the mobile application at least five times before answering the questionnaire to ensure the familiarity with the app. The questionnaire included 10 questions asking about the usability and the applicability of the usage of the developed mobile application alone with some general information related to bicycle riding. The selected user group rode the bicycles at least four times a week. Among the participants, 50% reported daily bicycle rides, 30% engaged in cycling a few days per week, and 20% cycled less frequently, predominantly during morning and evening hours on some days of the week. Notably, 65% of respondents identified vehicles attempting to overtake in proximity as the most significant safety concern during their bicycle rides, while 35% cited that vehicle approaching closely from behind as a key threat.

Table 1 gives the approximate distances of generating alarms when vehicles are overtaking from right as well as when vehicles approaching from behind. The application garnered high praise, with 72% of users expressing strong satisfaction with its usability and the effectiveness of its alarming system in enhancing their safety and overall riding experience and 80% of users agreed that they received alerts accurately.

Table 1 – Result Summary

	Vehicle overtaking from right	Vehicle approaching from behind
Initial Detection Approximate Distance (m)	8 - 10	8 - 10
Level 2 Alert Approximate Distance (m)	5 - 6	5 - 6
Level 3 Alert Approximate Distance (m)	2 - 3	1 - 3
Detection Accuracy (%)	88	90
False Positive Rate (%)	10	7

These findings emphasize the valuable contribution of this research to the safety enhancement of the bicycle riders and the positive reception of the application by the users.

5. Discussion

The aim of this research was to develop a mobile application that uses deep learning techniques to detect approaching vehicles using ambient sound to improve the safety of bicycle riders. The objectives of the study include collecting a large dataset of sound data, labelling the data with frequency characteristics corresponding to approaching vehicles and ambient noise, extracting features of the frequency spectrum when the vehicle approaches, developing a machine learning model using the labelled sound data as training data, and evaluating the performance of the model through testing and validation procedures to improve its accuracy. The findings of the study suggest that the proposed mobile application has the potential to effectively detect and classify nearby approaching vehicles using ambient sound to provide real-time alerts and incident reporting for bicycle riders. Using the built-in microphone of the smartphone of the user to capture audio data, the proposed solution provides a simple and cost-effective tool for improving bicycle rider safety. The system offers three different levels of alerts which will help users to take appropriate actions to avoid collisions.

This system does not measure the exact distance to the approaching vehicle nor measure the speed of the vehicle. However, the system can detect approaching vehicles by analysing ambient sounds using a deep learning model and by measuring the root mean square of the amplitude.

The study has limitations that need to be acknowledged. Although the model accuracy can be significantly improved with a large and diverse dataset, the amount of data that can be collected and used for training is practically limited. Moreover, the sound data collected in different environments may vary due to factors such as weather, traffic, and other background noise, which may impact the accuracy of the model in detecting vehicles. The model is limited in its ability to detect the speed of the vehicle.

The findings of this study have implications for the field of bicycle safety. The research is designed to enhance the safety of bicycle riders by detecting approaching vehicles and alerting them in real-time. The research uses on-device processing to analyse audio data in real-time, providing instant alerts to the user while incorporating advanced features that allow the user to adjust the sensitivity of the vehicle detection system according to their preference.

Future research could be conducted to evaluate the performance of the model in different environments, such as in highways, city roads, and rural roads and increase the size of the data set as well as the diversity and thereby the accuracy of the model can be improved. It is noted that the usage of this application does not need to be confined only for bicycle riders, yet it can be expanded to other types of vehicle riders as well with appropriate adjustments in the model as well as in the data set.

6. Conclusion

The CNN lite model trained on vehicle and non-vehicle audio clips was able to detect approaching vehicles and the three-level alert system implemented in the app can provide timely warnings to riders. The output alert levels are based on the root mean square value of the amplitude of the deep learning model. However, by improving the dataset the model accuracy can be further improved to detect the type of the vehicle as well. While the research has limitations in terms of the amount and diversity of data available for training and the variability of sound data collected in different environments, potential future works could address these limitations and further improve the accuracy and performance of the application. For future works, the system can be improved by incorporating features to detect vehicle types, speed of the vehicle and distance between user and the vehicle. In conclusion, the mobile application developed with deep learning techniques to detect approaching vehicles in real-time has the potential to greatly improve the safety of bicycle riders.

References

1. Huang, L., Wu, J., You, F., Lv, Z. and Song, H., 2016. Cyclist Social Force Model at Unsignalized Intersections with Heterogeneous Traffic. *IEEE Transactions on Industrial Informatics,* 13(2), pp.782-792.

2. Liu, M., 2013. A Study of Mobile Sensing using Smartphones. *International Journal of Distributed Sensor Networks*, *9*(3), p.272916.

3. Tabei, F., Askarian, B. and Chong, J.W., 2020. Accident Detection System for Bicycle Riders. *IEEE Sensors Journal*, *21*(2), pp.878-885.

4. Manori, G., Singh, G., Dushad, M., Vinod, M., Choudekar, P. and Tayal, V.K., 2018, April. Acoustic Detection of Rear Approaching Vehicles for Cyclists. In *2018 International Conference on Power Energy, Environment and Intelligent Control (PEEIC)* (pp. 354-357). IEEE.

5. Ma, G., Wu, M., Wu, Z. and Yang, W., 2021. Single-Shot Multibox Detector-and Building Information Modeling-Based Quality Inspection Model for Construction Projects. *Journal of Building Engineering*, *38*, p.102216.

6. Koszteczky, B. and Simon, G., 2013, May. Magnetic-Based Vehicle Detection with Sensor Networks. In *2013 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)* (pp. 265-270). IEEE.

7. Ma, C., Yang, D., Zhou, J., Feng, Z. and Yuan, Q., 2019. Risk Riding Behaviors of Urban E-Bikes: A Literature Review. *International Journal of Environmental Research and Public Health*, *16*(13), p.2308.

8. Smaldone, S., Tonde, C., Ananthanarayanan, V.K., Elgammal, A. and Iftode, L., 2011, March. The Cyber-Physical Bike: A Step Towards Safer Green Transportation. In *Proceedings of the 12th Workshop on Mobile Computing Systems and Applications* (pp. 56-61).

9. Palanisamy, K., Singhania, D. and Yao, A., 2020. Rethinking CNN Models for Audio Classification. *arXiv preprint arXiv:2007.11154*.

10. Nikhil, N. and Tran Morris, B., 2018. Convolutional Neural Network for Trajectory Prediction. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops* (pp. 0-0).

11. Petmezas, G., Cheimariotis, G.A., Stefanopoulos, L., Rocha, B., Paiva, R.P., Katsaggelos, A.K. and Maglaveras, N., 2022. Automated Lung Sound Classification using a Hybrid CNN-LSTM Network and Focal Loss Function. *Sensors*, *22*(3), p.1232.

12. Phaye, S.S.R., Benetos, E. and Wang, Y., 2019, May. Subspectralnet–using Sub-Spectrogram Based Convolutional Neural Networks for Acoustic Scene Classification. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 825-829).

13. Cheuk, K.W., Anderson, H., Agres, K. and Herremans, D., 2020. nnaudio: An on-the-Fly GPU Audio to Spectrogram Conversion Toolbox using 1d Convolutional Neural Networks. *IEEE Access*, *8*, pp.161981-162003.

14. "Bicycle Safety," Centers for Disease Control and Prevention. https://www.cdc.gov/transportationsafety/bicycle/ accessed on, 2024/01/29.

15. Spörri, E., Halvachizadeh, S., Gamble, J. G., Berk, T., Allemann, F., Pape, H. C., & Rauer, T. (2021). Comparison of Injury Patterns between Electric Bicycle, Bicycle and Motorcycle Accidents. *Journal of Clinical Medicine*, *10*(15), 3359.

16. "Road Accidents", National Council for Road Safety, Ministry of Transport and Highways, https://www.transport.gov.lk/web/index.php?option=com_content&view=article&id=29&Itemid=149&lang=en#type-of-vehicles-involved-in-accidents, accessed on 2024/01/29.