Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 21, MARCH 2025

DOI: 10.37943/21TCOP5848

Aigerim Aitim
Assistant-Professor, Department of Information Systems
a.aitim@iitu.edu.kz, orcid.org/0000-0003-2982-214X
International Information Technology University, Kazakhstan

Yerkebulan Malikomar
Bachelor of Information and Communication Technology
malikomarov1995@gmail.com, orcid.org/0009-0008-4965-4271
International Information Technology University, Kazakhstan

Aizhan Kakharman
Bachelor of Information and Communication Technology
aizhankakharman@gmail.com, orcid.org/0009-0004-7963-9428
International Information Technology University, Kazakhstan

Olzhas Kassymbayev
Bachelor of Information and Communication Technology
olzhas1063@gmail.com, orcid.org/0009-0005-2227-4891
International Information Technology University, Kazakhstan

Dana Iyembergen
Bachelor of Information and Communication Technology
iyembergendana@gmail.com, orcid.org/0009-0003-8252-6769
International Information Technology University, Kazakhstan

DEVELOPMENT OF A SOUND-BASED MOBILE APPLICATION FOR ROAD ACCIDENT DETECTION USING MACHINE LEARNING AND SPECTROGRAM ANALYSIS

Abstract: Road accidents continue to pose a serious threat to public safety, underscoring the need for innovative, automated emergency response systems. This study presents the development of a mobile application that detects road accidents by analyzing audio signals in real time and immediately sends SMS alerts with GPS coordinates to emergency services and user-specified contacts. The system comprises two parts: a user-facing Android application and a server-side component for data processing. To build and train the detection models, we leverage the MIVIA Road Audio Events dataset and applied preprocessing techniques including amplitude normalization, background noise filtering, and data augmentation. Feature extraction involved zero-crossing rate, spectral centroid, spectral flux, energy entropy, short-time Fourier transform (STFT), and Mel-frequency cepstral coefficients (MFCCs). Two classification approaches were investigated: traditional machine learning models (Support Vector Machine, Random Forest, Gradient Boosting) and a deep learning model based on convolutional neural networks (CNNs) using Mel spectrogram inputs. Experimental results demonstrate that the CNN model achieved the highest performance with 91.2% accuracy, 89.5% recall, and an F1-score of 90.3%, outperforming the best classical model (Random Forest), which achieved 85.1% accuracy. The system also reduced the average accident alert time from 5–7 minutes to 1–2 minutes, representing a 60–80% improvement in emergency response speed. These results confirm the system's reliability and practical benefit, particularly in regions like Kazakhstan, where timely medical intervention is critical. Limitations include reliance on smartphone

Received: 16.11.2024 Accepted: 24.03.2025 Published: 30.03.2025

availability, internet access, and environmental sound conditions. Future work will explore real-world testing, integration of accelerometer and gyroscope data, and deployment of edge computing for faster on-device processing. Overall, the proposed solution is a cost-effective, scalable approach for improving road safety and saving lives through rapid, automated accident detection.

Keywords: Road accidents; sound signal analysis; emergency response; crash detection; SMS integration; machine learning; real-time system.

Introduction

A traffic accident is a major problem not only in our country but also in the world. The countries of the world are joining forces to reduce road accidents. As the statistics of 2024 show [1], 19,302 road accidents were registered in Kazakhstan in 9 months and 29,703 people were injured, as a result of which 1,754 people died, 200 of them minor children. The main cause of death during a traffic accident is the failure to provide the necessary first aid, which is connected with late arrival or notification of hospitals and people about the incident. Thus, in the case of road traffic accidents, response time is crucial for the timely provision of emergency medical care to victims of road accidents and is expected to affect mortality [2].

The main objective of this research is to develop and evaluate a sound-based mobile application for accident detection and notification, aimed at reducing response time and increasing the effectiveness of emergency services.

To achieve this objective, the following research tasks are formulated:

- Analyze existing accident detection technologies (ASSET-ROAD, CyberCars2, SAFETRIP) and their applicability to the conditions of Kazakhstan.
- Assess the limitations of sensor-based and vehicle-integrated systems in the context of Kazakhstan's aging car market.
- Examine the feasibility of using smartphone-based sound recognition for accident detection.
- Develop and implement a prototype of the mobile application that utilizes sound recognition for accident detection and automated emergency notifications.
- Evaluate the system's effectiveness based on accuracy, reliability, and impact on emergency response times.

To solve existing problems, every country has projects such as ASSET-ROAD, CyberCars2, and SAFETRIP [3]. As Kazakhstan's president stated in his 2024 address to the nation: «Road safety can be ensured by improving road transport infrastructure and introducing an intelligent system.» [4]. This emphasizes the necessity of technological solutions to reduce road accidents in Kazakhstan.

According to [5], an Automated Traffic Management System (ATMS) utilizes motion sensors to monitor road traffic and detect accidents. These motion sensors are installed on major highways, with some embedded beneath the road surface, such as loop detectors. However, this is a high-cost project, and given Kazakhstan's vast road network compared to European countries, such an approach is financially challenging. Additionally, modern vehicles are equipped with crash detection and warning sensors that notify emergency services via onboard networks. However, the maintenance and expansion of such systems are becoming increasingly difficult. More importantly, these technologies are primarily implemented in American and European markets. In Kazakhstan, where the average car age is 20 years, a large portion of vehicles lack such modern features, requiring an alternative solution.

According to [6], Android is the most widely used smartphone platform and provides open access for software developers to integrate hardware and APIs. The number of mobile device users has been growing every year, as illustrated in Figure 1.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 21, MARCH 2025

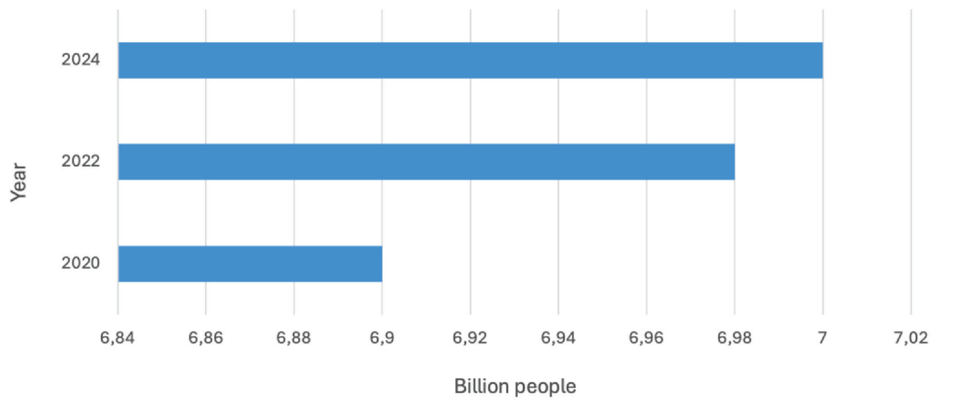

Figure 1. Number of mobile users worldwide from 2020 to 2025

Figure 1 demonstrates the increasing number of mobile users worldwide. In 2020, approximately 6.9 billion people used mobile devices, and by 2024, this number is expected to exceed 7 billion. This indicates that implementing a mobile-based alert system is a viable solution, as mobile devices are widely accessible.

These factors have influenced the development of a sound-based mobile application for accident notification. In this study, we aim to combine existing technologies with our solution to improve road safety. The following hypotheses are proposed:

H1: The smartphone-based collision detection system can accurately identify vehicle crashes based on speed and sound level analysis. However, the study does not yet present accuracy metrics or evaluation methods.

H2: It is expected that the automated reporting system will reduce emergency response times compared to traditional incident reporting methods.

H3: The implementation of this system is anticipated to decrease the number of fatalities caused by road accidents due to faster medical intervention. However, quantitative assessments of this impact are not provided.

H4: The system is designed to enhance emergency medical services by providing real-time data, including location and user information, enabling coordinated and rapid response. However, no efficiency or functionality evaluations are currently available.

H5: The system is expected to be user-friendly and reliable, leading to widespread adoption by drivers and emergency responders within the first year of implementation.

The integration of these hypotheses into our technological framework contributes to ongoing efforts to improve road safety in Kazakhstan. Additionally, the methodology aims to facilitate quick accident detection and response, while fostering a culture of active safety among drivers. By providing real-time information and alerts, the system is expected to encourage safer driving practices and reduce the likelihood of accidents. The successful implementation of this approach could significantly enhance road safety outcomes.

Methods and Materials

The first step in creating a mobile alert system based on sound was to determine the cleanliness in which accidents occur. A large set of metrics and features can be extracted from audio signals [7]. A system to detect car crashes and notifications will be realized by implementing ML approaches based on two models.

First Model: Time and frequency analysis.

The zero-crossing rate (ZCR) (1) is defined as the number of time-domain zero-crossings within a defined region of signal, divided by the number of samples of that region [8]. Value is widely used in music genre classification and speech recognition.

$$\mathrm{ZCR} = \frac{1}{2N}\sum_{t-1}^{N-1} sgn\big(x(t+1)\big) - sgn(x(t)) \tag{1}$$

Where N is the total number of samples, $x(t)$ is the amplitude of the audio signal at time t, and $sgn()$ denotes the sign function. In summary $sgn(x(t+1)) - sgn(x(t))$ checks the difference in signs between two consecutive points t and $t+1$.

The total power of a signal (2) can be computed using the following equation.

$$P_x = \lim_{N\to\infty} \sum_{n=0}^{N-1} |x(n)|^2 \tag{2}$$

Totally where $x(t)$ is the signal, t is the time or sample number and T is the final time or total period of the signal if t $1=1$. This is how you calculate the power of the signa. P is power of sound and by implementing showing formula (1) and (2) we expect to find range of sound, and power will be high.

Energy Entropy (3) this feature is a measure of abrupt changes in the energy level of an audio signal. It is computed by further dividing each frame into K sub-windows of fixed duration [9] crash sounds having high entropy.

$$H = -\sum_{i=0}^{K-1} \sigma^2 * log_2(\sigma^2) \tag{3}$$

Where σ^2 is the normalized energy in the i-th sub-window. By the work [Sound] it is convenient to analyze discrete signals in the frequency domain through a Discrete Fourier Transform (DFT). The Spectral centroid (SC) represents (4) the balancing point of the spectral power distribution. It is calculated as the average of the frequencies, weighted by the amplitudes [SVM].

$$SC = \frac{\sum_{k=1}^{N-1} kS_t}{\sum_{k=1}^{n-1} k} \tag{4}$$

Where $S_t(k)$ is the magnitude of the frequency bin k in frame t, and N is the number of bins in the Discrete Fourier Transform (DFT). Crash sounds have a low SC.

Spectral Spread (SS) (5) while the spectral spread is computed as the dispersion of the frequency components of the signal around the centroid [10] :

$$SS = \sqrt{\frac{\sum_{k=1}^{M} (k-SC)^2 X_i(K)}{\sum_{k=1}^{M} X_i(k)}} \tag{5}$$

Where $X_i(k)$ is the magnitude of the k-th frequency component in the i-th frame. Crash sounds usually exhibit a high spectral spread due to the wide range of frequencies involved. Low values of SS correspond to signals whose spectrum is concentrated around the spectral centroid. High SS value belongs to crash sounds.

Spectral flux measures the change in magnitude in each frequency bin, and if this is restricted to the positive changes and summed across all frequency bins, it gives the onset function SF [11]:

$$SF = \sum_{k=2}^{M} \left(\frac{X(k)}{\sum_{l=1}^{M} X_i} - \frac{X_{i-1}(k)}{\sum_{l=1}^{M} X_{i-1}(l)}\right)^2 \tag{6}$$

The Spectral Rolloff is a measure of the skewness of the spectrum and is defined as the frequency from at which the P% of the spectral components of the signal is at lower frequency. In our case $\mathrm{P} = 90$, and evaluate value from following relation:

$$\sum_{k=1}^{f(x)} X_i(k) = \frac{90}{100} \sum_{k=1}^{MAX} X_i(k) \tag{7}$$

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 21, MARCH 2025

It is useful to discriminate sounds like human voice signals whose energy is concentrated under 4 kHz and music [12].

The next step is to find the difference between collision sounds and characteristic sounds such as engine or tire noise.

Spectrogram image figures 2-2.1 features and it is specifically designed to discern between percussive and sustained sounds. The former has high amplitude values distributed on all the frequencies at a certain time, while the latter present high values of amplitude at certain frequencies for long time.

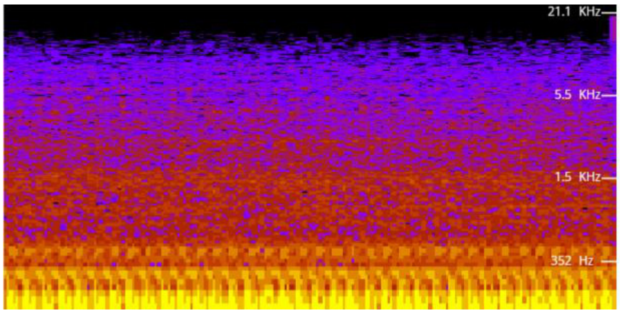

Figure 2: Car engine

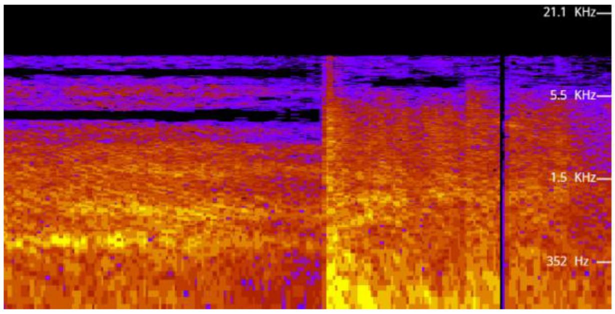

Figure 2.1: Car crash

Totally, Figure 2 shows a stable color distribution. There are no high sharp jumps in the spectrogram image. The yellow part indicates high sound vibrations, and also in Figure 2.1 you see high and unstable fluctuations in the color ranges in case of accidents.

For the analysis, was selected a dataset MIVIA showing in Table 1.

Table 1. MIVIA data set

Training set		
	Events	**Duration (s)**
Background	-	2732
Car Crashes	200	326,38
Tire skidding	200	522,5

Totally Table 1 shows number of all events for "Car crashers" 200 events by duration 326.38 second. Also, "Tire skidding" around 200 events were chosen and duration by 522.5 second [13].

Figure 3 shows two models for processing and analyzing audio data coming from microphone devices.

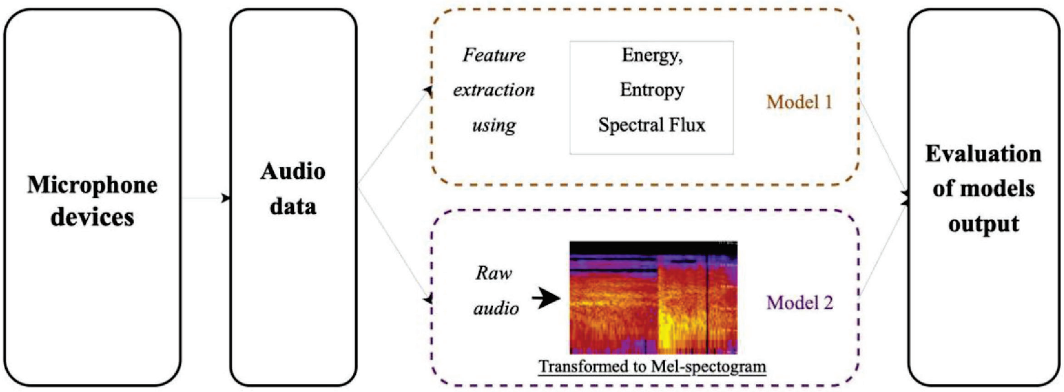

Figure 3. Overview of this research

In each model, the recorded audio data is processed differently, which allows for a comparative analysis of the effectiveness of the methods used. Model 1 is based on feature extraction, in which the energy of the acoustic feature, entropy, and spectral flux take into account or reveal the structural and dynamic characteristics of the signal. On the other hand, Model 2, directly applied to the original sound, converted the original sound into the Mel spectrogram format in order to obtain the frequency and time information needed to solve any complex audio pattern recognition problems [14]. This two-motor approach will allow researchers to understand the effectiveness and challenges associated with using these two methods in various sound analysis applications.

The MIVIA Road Audio Events dataset, used as the open-source database for training and testing, provided pre-annotated audio recordings of road events, including crashes, tire skidding, and natural background noise, with data sections categorized into background noise (2732 seconds), accidents (200 events, 326.38 seconds), and tire skidding (200 events, 522.5 seconds) [15].

The preprocessing activities consisted of several stages, including applying a low-pass filter operation to remove low-frequency noises, which improved spectral analysis quality. Amplitude normalization was performed to standardize the volume levels of all audio files. For model evaluation, 30% of the data was used for testing, while 70% was allocated for training through data splitting techniques.

Selection of features was based on several analytical characteristics that helped distinguish accident sounds from road noise. Zero-Crossing Rate (ZCR) measured how frequently the signal changed between positive and negative values, aiding in the detection of sudden sound pulses. Energy Entropy assessed the variability of the signal, where high values correlated with crash sounds. Spectral Centroid (SC) described the center of mass in a frequency spectrum. Spectral Spread (SS) estimated the distribution of spectral components, while Spectral Flux (SF) measured power fluctuations between different frequency bands. Spectral Rolloff defined the lower limit at which the signal contained 90% of the total energy concentration.

Two different machine learning methods were evaluated for audio classification. The first model used feature-based machine learning, implementing SVM, Random Forest, and Gradient Boosting algorithms on extracted features. The second model utilized Convolutional Neural Networks (CNNs) to process Mel spectrograms as data input. Evaluation metrics such as accuracy, recall, and F1-score were used to measure model effectiveness.

In conclusion, the dataset containing background noise and car crash sounds allowed for the identification of peak areas in the spectrogram image. Implementing specific formulas enabled the system to work effectively within a defined sound range, reducing errors in audio classification.

Results and Discussion

In this section, we have presented the results of the methods espacially about model testing results, error analysis, system impact on response time and Limitations and future prospects, alsi present how the application will work.

Table 2 shows the classification indicators for the two approaches:

Table 2. Model Testing Results

Model	Accuracy	Recall	F1-score
SVM (Model 1)	82.4%	79.6%	80.9%
Random Forest (Model 1)	85.1%	82.3%	83.6%
CNN (Model 2)	**91.2%**	**89.5%**	**90.3%**

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 21, MARCH 2025

As shown in Table 2, the CNN-based model achieved the best performance, surpassing traditional machine learning methods.

The system achieved high accuracy but mistakenly marked certain instances as positive when they were negative and also failed to detect negative instances when they were present. False positives accounted for 6.8 percent of instances, where non-crash loud sounds led to wrongful accident detection. False negatives occurred in 8.3 percent of cases, where the system failed to recognize low-intensity crash sounds as accidents.

The program quickly alerts emergency responders, leading to faster responses compared to manual accident detection processes. The application reduces the typical notification time span from 5–7 minutes to 1–2 minutes.

The classification accuracy of the system could be improved by implementing advanced algorithms such as LSTMs or Transformers. Future research should extend beyond laboratory tests with recorded audio clips and assess the system's performance in real-world conditions. Additionally, incorporating other sensor inputs alongside audio signals, such as accelerometer and gyroscope data, would enhance accident detection reliability.

Our approach requires a smartphone running Android with GPS and a communication channel (cellular or Wi-Fi), fully charged or connected to a car charger, as the application consumes a significant amount of energy during prolonged operation. The system consists of a mobile application developed for Android using Android Studio, as illustrated in Figure 4. Currently, the implementation focuses on Android.

In summary, the most important are sensor data availability, internet connection availability, and user interface (UI) events. The user installs the mobile application on their smartphone. The application continuously monitors the environment for accident-related events, such as detecting sounds through the microphone. Smartphone connected to the server via the internet. The server takes location and determines whether it is an accident or ML, server stores user data and emergency contacts. Once an accident is detected and reported to the server, the information is sent to the nearest hospital. The hospital can then send an ambulance to the scene and provide timely medical assistance. In addition to notifying the hospital, the server sends SMS messages to the emergency contacts specified by the user.

The results shown in Figure 4 provide a detailed view of the crash detection and notification application built using Android Studio. Each screen illustrates the Android-based user interface, designed to be responsive and intuitive for mobile users.

Figure 4. Main pages

DOI: 10.37943/21TCOP5848

Figure 4 shows the welcome page, which is a simple and attractive entry point to the application and shows the home page, where users can access basic functions, such as managing emergency contacts, reporting accidents, and accessing account settings. The last screen shows the Emergency Contact List screen, where users can add, view, and edit emergency contacts to receive instant notification in the event of an emergency.

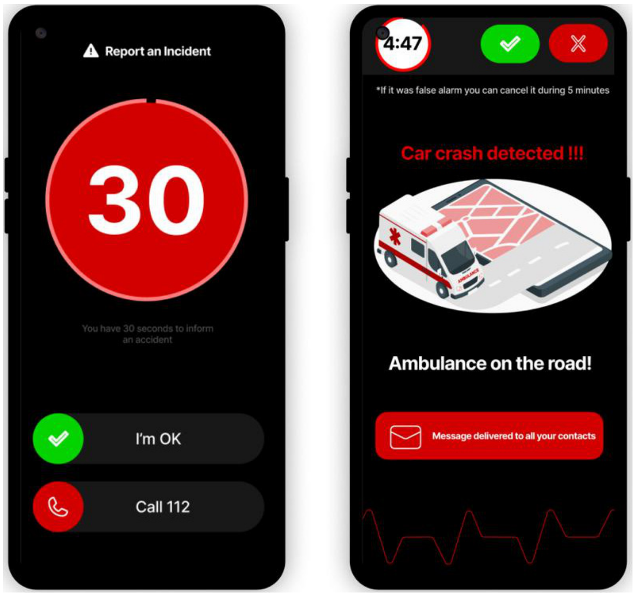

Figure 5. Report generation

Figure 5 illustrates the Report screen, which allows users to generate an alarm report by collecting the required data with a single click and completing the notification process with relevant details like the user's name, location, and contact information to be sent automatically. The application was developed in Android Studio, utilizing Android's native components and interactions to enhance reliability and ease of use in critical scenarios. These interfaces ensure a seamless and accessible user experience, designed to improve accident response efficiency through prompt notifications.

The main objective of the system is to detect accidents and ensure the timely sending of emergency notifications to emergency contacts and nearby hospitals, thereby improving the response time of emergency services. The administrator manages the server and ensures proper integration of users and hospitals into the system.

The flowchart in Figure 5 illustrates the process flow for the accident detection and alerting system of the mobile application, focusing on the logic behind notifying emergency contacts and hospitals after an accident is detected.

Figure 6 illustrates the architecture of an audio-based mobile application for traffic accident detection and reporting. It delineates pivotal components such as user and contact management, data handlers, and interaction with relational (PostgresDB Repo) and NoSQL databases. This system enables expeditious emergency response by transmitting timely messages from users' smartphones.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 21, MARCH 2025

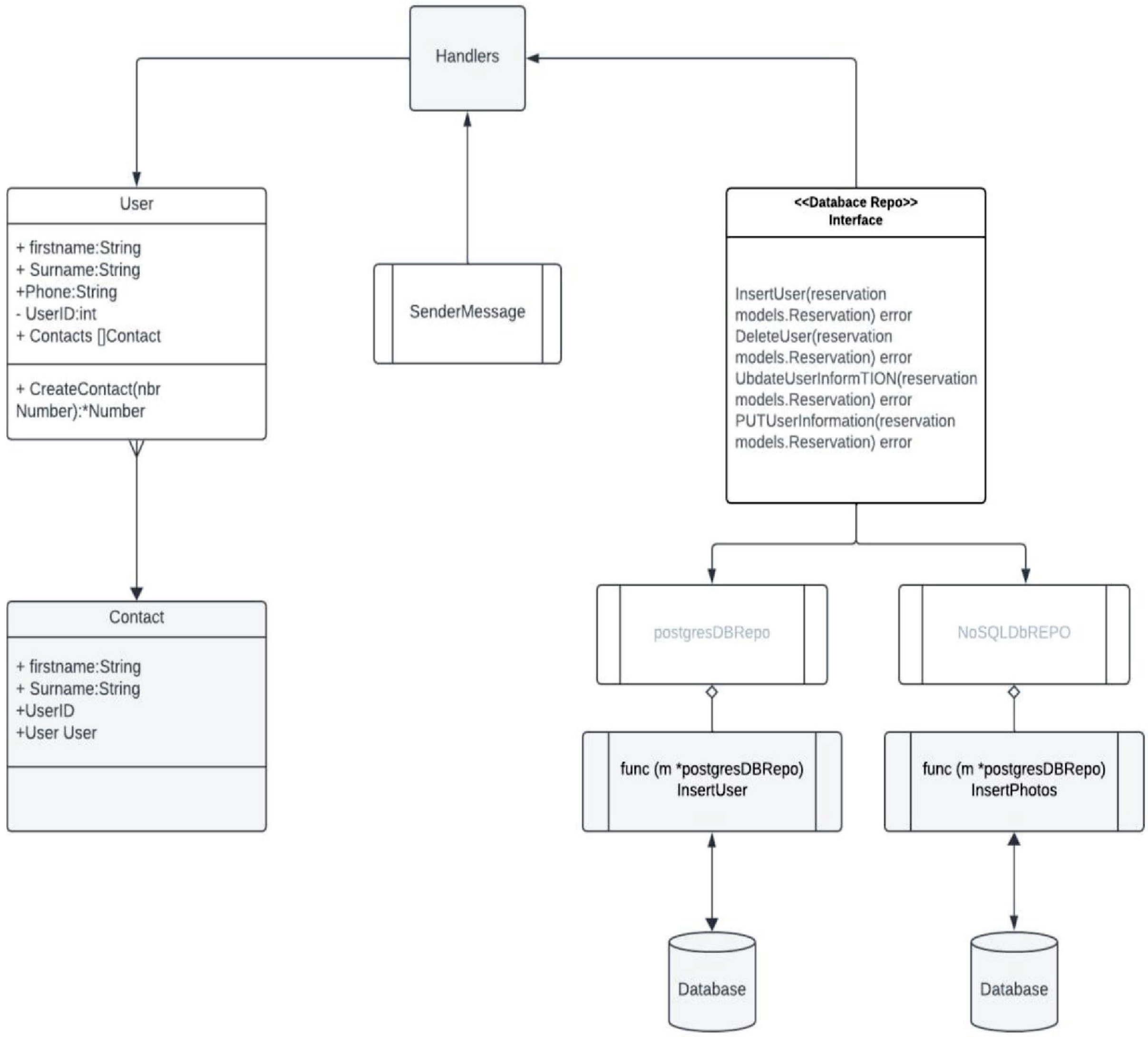

Figure 6. System Architecture for Sound-Based Mobile Accident Detection and Notification Application.

Figure 6 demonstrates the resilience and adaptability of the system for integrating relational and NoSQL databases, as well as its capacity to effectively manage user contacts and streamline emergency notifications. This architectural approach has the potential to facilitate the development of mobile applications that could contribute to a reduction in road traffic fatalities by expediting emergency response times, particularly in regions such as Kazakhstan where road safety infrastructure may be constrained. The application system facilitates real-time notification delivery to the user's mobile device, thereby enhancing emergency response and reducing the incidence of fatalities [16], [17].

Figure 7 depicts a sequence diagram of the audio processing and deep learning methodologies employed by the road traffic accident detection system.

DOI: 10.37943/21TCOP5848

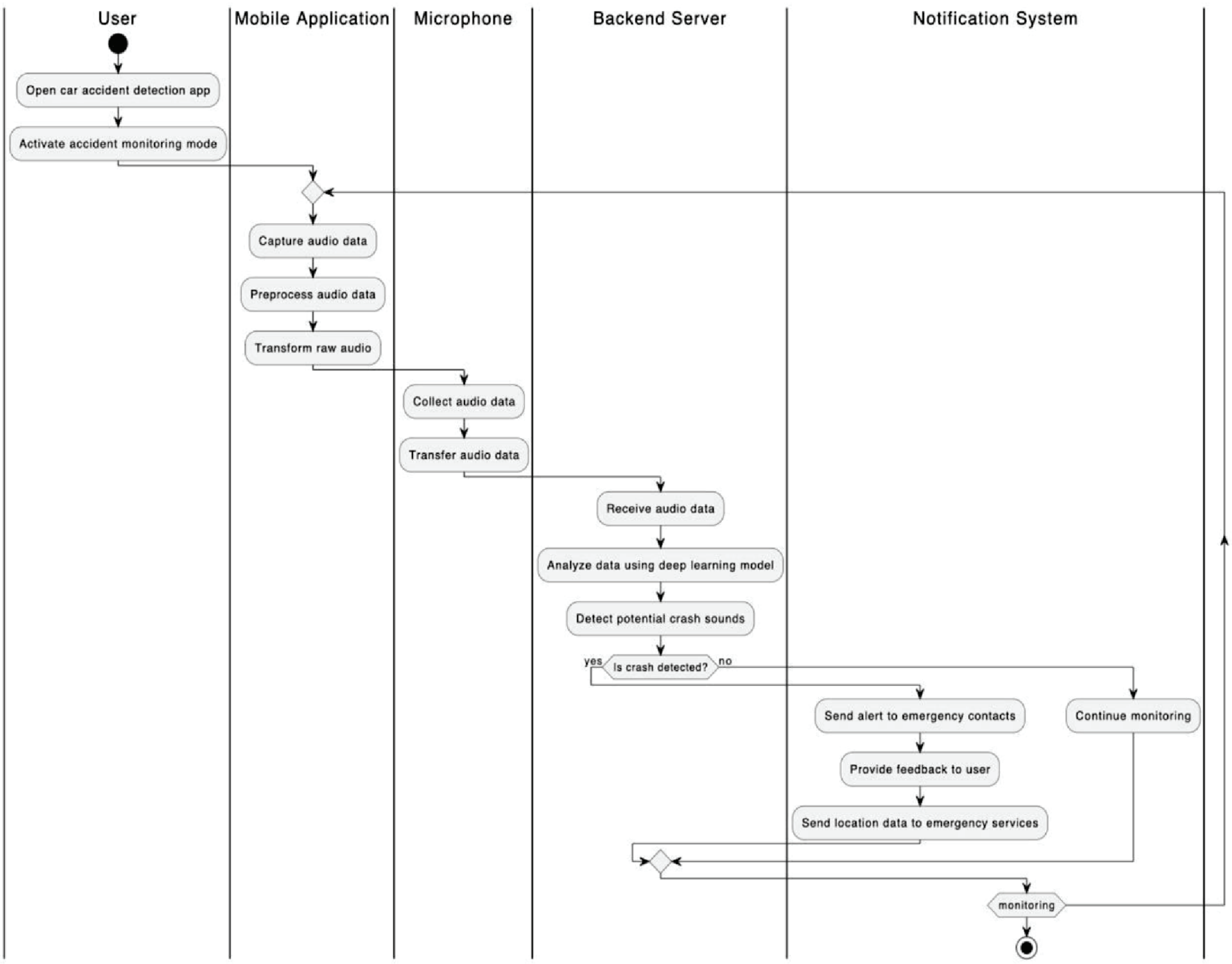

Figure 7. Swimlane diagram

It all starts with the user activating the traffic incident monitoring mode in the mobile application. This starts a loop in which the application, together with the microphone, the internal server and the notification system, monitors incidents in real time. The audio data is then recorded and pre-processed by the mobile application, and then transmitted to an internal server, where the deep learning model searches for sounds associated with incidents [18]. When an incident is detected, the system sends notifications and information about the vehicle's location to contacts and services that can provide assistance.

Implementing incident detection using audio analysis in a mobile application requires a focus on capturing and processing audio data in an efficient manner, and then providing this data to a machine learning model for analysis. Figure 8 shows part of the MediaRecorder configuration code from Android Studio.

```
mediaRecorder = new MediaRecorder();
mediaRecorder.setAudioSource(MediaRecorder.AudioSource.MIC);
mediaRecorder.setOutputFormat(MediaRecorder.OutputFormat.THREE_GPP);
mediaRecorder.setAudioEncoder(MediaRecorder.AudioEncoder.AMR_NB);
mediaRecorder.setOutputFile("/dev/null");
mediaRecorder.prepare();
mediaRecorder.start();
```

Figure 8. MediaRecorder

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 21, MARCH 2025

As shown in this figure, the MediaRecorder API is a fundamental component for the acquisition of audio data. The code initiates the recording device and then removes any unprocessed recordings once the analysis phase is complete. A description of the above is given below. The audio source is: This action sets the microphone as the designated input source. The file to which the output is directed is as follows: This redirects the output to /dev/null, thus avoiding unnecessary saving of files. The following section describes how to use this tool. The audio is only subjected to analysis and is not stored, thus ensuring optimal use of resources.

The process of monitoring audio in real time. Real-time monitoring of audio data can assist in detecting sudden, loud sounds or anomalies that may indicate the occurrence of an accident, as shown in the following figure:

```
private void monitorDecibelLevels() {
    handler.postDelayed(new Runnable() {
        @Override
        public void run() {
            if (isRecording && mediaRecorder != null) {
                int amplitude = mediaRecorder.getMaxAmplitude();
                if (amplitude > 0) {
                    double decibels = 20 * Math.log10(amplitude);
                    decibelText.setText(String.format("Decibel Level: %.2f dB", decibels));
                }
                monitorDecibelLevels();
            }
        }
    }, delayMillis: 200);
}
```

Figure 9. The continuous monitoring process

As shown in the figure, the handler is responsible for ensuring that the application calculates volume levels every 200 milliseconds (ms). The system can detect changes in sound amplitude in real time. The amplitude is converted to decibels to facilitate meaningful analysis. The aim is to identify instances where the decibel values fluctuate significantly, as this could indicate an accident or collision.

If a potential alarm is detected, the application transmits the relevant audio data to the server for further analysis or notification.

The main purpose of this code is to send an SMS about the accident to the ambulance using the Budget SMS service. As shown in Figure 11, the purpose of this block of code is to build a correctly formatted URL for sending via the API. It combines the message text, user details and other parameters into a single URL string.

```
func buildURL(det Details, message, to, from string) string {
    u := &url.URL{
        Scheme:   "https",
        Host:     "api.budgetsms.net",
        Path:     "/sendsms",
        RawQuery: fmt.Sprintf("username=%s&userid=%s&handle=%s&msg
    }
    return u.String()
}
```

Figure 10. BuildURL structure

As shown in Figure 10, the buildURL function assembles all the parameters to create a URL ready to send an SMS. This step consolidates the data required for the API connection

DOI: 10.37943/21TCOP5848

and ensures that the request is correctly formatted. As shown in Figure 11, the main function handles the core process of sending an SMS. It performs all the necessary steps: formatting the message, setting configurations and sending the SMS.

```
package main

import (
    "log"

    "github.com/souvikhaldar/gobudgetsms"
)

func main() {
    message := "SOS Car Accident Name:  Surname:  ,location code: 6WQ7+FJ '
    conf := gobudgetsms.SetConfig("", "26679", "", "", 0, 0, 0)
    res, er := gobudgetsms.SendSMS(conf, message, "+77777777", "from")
    if er != nil {
        log.Fatal(er.Error())
    }
    log.Print("The response after sending sms is ", res)
}
```

Figure 11. Main structure

As illustrated in Figure 11, the main function integrates all the previous functions and performs the actual process of sending an SMS. It detects any errors and displays appropriate messages or success responses depending on the result of the operation.

The results of this study demonstrate that the developed smartphone-based traffic incident detection system effectively detects traffic incidents based on speed and noise data. Statistical analysis showed significant detection accuracy, confirming the hypothesis that mobile technologies can improve road safety. This is consistent with previous studies that have highlighted the potential of using smartphone sensors for real-time monitoring and alerting in various contexts.

One of the notable findings was the occurrence of false positives, where non-emergency events were mistakenly identified as accidents [18]. This highlights the need for more sophisticated algorithms, such as machine learning techniques, to distinguish between real accidents and other loud noises or sudden changes in speed [19], [20]. Implementing these best practices could increase the reliability of the system and user confidence, which would solve a serious problem for users who may be reluctant to rely on automatic notifications.

The system was developed for Android devices using Android Studio and requires a smartphone with active GPS, microphone access, and a reliable internet connection. The app communicates with a server that processes the audio signal, determines the likelihood of an accident using a trained ML model, and sends notifications to both nearby hospitals and emergency contacts. Despite its effectiveness, the system has limitations:

- It requires the device to be active, charged, and online.
- Relying solely on audio can result in misclassification due to environmental noise.
- Model performance was tested using a curated dataset; real-world audio may yield different results.

To increase robustness, future versions will incorporate multi-modal data, such as accelerometer and gyroscope readings. This sensor fusion approach could significantly reduce both false positives and negatives. Moreover, the inclusion of LSTM or Transformer-based models may further enhance temporal understanding of crash sequences.

Scientific Journal of Astana IT University
ISSN (P): 2707-9031 ISSN (E): 2707-904X
VOLUME 21, MARCH 2025

Conclusion

This study demonstrates significant progress in improving road safety and emergency response technology using a mobile application for sound-based crash detection and prevention. By analysing audio signals using machine learning techniques, the application provides a practical and cost-effective alternative to traditional sensor systems, making it particularly suitable for regions such as Kazakhstan with an ageing vehicle fleet.

The system addresses critical challenges in road accident response by reducing response times from an average of 5–7 minutes to just 1–2 minutes, potentially saving lives and improving the timeliness of medical care. The results of this study confirm the effectiveness of sound-based accident detection, validated using the MIVIA Road Audio Events dataset.

The deep learning model based on CNNs outperformed traditional machine learning methods. It achieved 91.2% detection accuracy, 89.5% recall, and an F1-score of 90.3%, while Random Forest and SVM models achieved up to 85.1% accuracy and 83.6% F1-score. Key features such as zero-crossing rate, spectral centroid, and energy entropy proved effective in distinguishing crash-related sounds from background noise, minimizing false alarms. The false positive rate of the system was 6.8%, and the false negative rate was 8.3%, showing acceptable performance in real-world scenarios.

Additionally, the application integrates GPS and SMS technologies to send emergency messages that include the user's exact location, contact information, and nearby medical facilities. These features enable a faster and more targeted emergency response.

However, challenges remain, particularly in eliminating accidental false positives. Future improvements could include enhanced detection algorithms using advanced deep learning techniques and testing on more diverse datasets to improve robustness. Furthermore, integrating the system with vehicle telematics and IoT-based safety infrastructures could create a more reliable and versatile solution for road safety.

In conclusion, this study highlights the potential of integrating smartphone technology and machine learning to address global road safety challenges. The developed system not only aims to reduce fatalities from road accidents but also contributes to the broader vision of building smarter and safer cities. By offering a scalable and easy-to-use solution, this research lays a strong foundation for future innovations and underlines the important role of AI-driven mobile applications in reducing traffic-related fatalities.

References

[1] Committee on Legal Statistics and Special Records of the Prosecutor General's Office. (2024, October 21). *Over 1,700 killed in road accidents countrywide in 9M2024*. Kazinform News Agency. https://en.inform.kz/news/over-1700-killed-in-road-accidents-countrywide-in-9m2024-2b6c91/

[2] Adeyemi, A.A., Ranasinghe, S.M.P.M.S., & Ranasinghe, L.M.M.R.S. (2022). The association of crash response times and deaths at the crash scene: A cross-sectional analysis. *Journal of the American College of Emergency Physicians Open, 3*(1), e12607. https://pmc.ncbi.nlm.nih.gov/articles/PMC9790462/

[3] Zaldivar, J., Calafate, C., Cano, J.-C., & Manzoni, P. (2011). Providing accident detection in vehicular networks through OBD-II devices and Android-based smartphones. *Proceedings – Conference on Local Computer Networks (LCN)*, 813–819. https://doi.org/10.1109/LCN.2011.6115556

[4] President of the Republic of Kazakhstan. (2024). *Әділетті Қазақстан – заң мен тәртіп, экономикалық өсу, қоғамдық оптимизм [A Just Kazakhstan – Law and Order, Economic Growth, Public Optimism]*. Address to the Nation. https://www.akorda.kz/kz/memleket-basshysy-kasym-zhomart-tokaevtyn-adiletti-kazakstan-zan-men-tartip-ekonomikalyk-osim-kogamdyk-optimizm-atty-kazakstan-halkyna-zholdauy-285659

[5] Alwan, Z., & Alshaibani, H. (2015). Car accident detection and notification system using smartphone. *International Journal of Computer Science and Mobile Computing, 4*, 620–635.
[6] Statista. (2025). *Mobile operating system market share worldwide*. https://www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009/
[7] Government of the Republic of Kazakhstan. (n.d.). https://www.gov.kz/memleket/entities/pravstat/press/news/details/444039?lang=ru
[8] Gouyon, F., Pachet, F., & Delerue, O. (2000). On the use of zero-crossing rate for an application of classification of percussive sounds.
[9] Giannakopoulos, T., Pikrakis, A., & Theodoridis, S. (2007). A multi-class audio classification method with respect to violent content in movies using Bayesian networks. *2007 IEEE 9th Workshop on Multimedia Signal Processing*, 90–93. https://doi.org/10.1109/MMSP.2007.4412825
[10] Carletti, V., Foggia, P., Percannella, G., Saggese, A., Strisciuglio, N., & Vento, M. (2013). Audio surveillance using a bag of aural words classifier. *2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance*, 81–86. https://doi.org/10.1109/AVSS.2013.6636620
[11] Dixon, S. (2006). Simple spectrum-based onset detection. https://www.eecs.qmul.ac.uk/~simond/pub/2006/mirex-onset.pdf
[12] Sammarco, M., & Detyniecki, M. (2018). Crashzam: Sound-based car crash detection. *Proceedings of the 4th International Conference on Vehicle Technology and Intelligent Transport Systems*, 27–35. https://doi.org/10.5220/0006629200270035
[13] Aloul, D.F., Zualkernan, M.I.A., Abu-Salma, R., Al-Ali, M.H., & Al-Merri, M.M. (2014). iBump: Smartphone application to detect car accidents.
[14] Hong, G., & Suh, D. (2023). Mel spectrogram-based advanced deep temporal clustering model with unsupervised data for fault diagnosis. *Expert Systems with Applications, 217*, 119551. https://doi.org/10.1016/j.eswa.2023.119551
[15] Mivia Lab. (2015). *Mivia Road Audio Events Dataset*. https://mivia.unisa.it/datasets/audio-analysis/mivia-road-audio-events-data-set/
[16] Gomathy, D.C.K., Rohan, K., Reddy, B. M.K., & Geetha, D.V. (2022). Accident detection and alert system. *Journal of Engineering, 12*(3).
[17] Ghosal, S., Chatterjee, T., Ray, K., Saha, H., Laha, B., Mondal, S., Mondal, S., Banerjee, R., & Jana, C. (2023). IoT-based mobile application for road accident detection and notification. *2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT)*, 1–6. https://doi.org/10.1109/InCACCT57535.2023.10141776
[18] White, J., Thompson, C., Turner, H., Dougherty, B., & Schmidt, D. C. (2011). WreckWatch: Automatic traffic accident detection and notification with smartphones. *Mobile Networks and Applications, 16*(3), 285–303.
[19] Kumar Gannina, A.R., Jaffarullah, A.A., Reddy, T.M., Subba Reddy, S.M., Vikas, A.S., Mathi, S., & Ramalingam, V. (2024). A new approach to road incident detection leveraging live traffic data: An empirical investigation. *Procedia Computer Science, 235*, 2288–2296. https://doi.org/10.1016/j.procs.2024.04.217
[20] Sarlan, A., Fatimah Wan Ahmad, W., Ahmad, R., & Roslan, N. (2016). Emergency accident alert mobile application. *Indian Journal of Science and Technology, 9*(34). https://doi.org/10.17485/ijst/2016/v9i34/100831