International Journal of Trend in Scientific Research and Development (IJTSRD)
Special Issue on Smart Innovations in Computer Science and Applications
Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Urban Transit Vigilance and Defense System

Nishtha Shambharkar
Department of Science and Technology,
G. H. Raisoni Skill Tech University, Nagpur, India

Annotation

Urban transit systems like airports, railway stations, and metro networks are essential public infrastructures that need strong security measures to keep passengers safe. Traditional surveillance mainly relies on video monitoring and manual oversight, which can delay responses in spotting potential threats. This research introduces the Urban Transit Vigilance and Defense System, a technology-based solution aimed at improving security through real-time speech recognition and automated threat detection.

The proposed system captures audio from transit environments and processes it using real-time speech recognition technology. It converts spoken conversations into text. The transcribed data is then analyzed to find specific "alert words" that may signal security threats. When the system detects these keywords, it logs the event with exact timestamps and automatically sends notifications, including email alerts to authorized security staff. This setup allows for quick responses and helps Security Operations Center (SOC) analysts identify and manage potential risks effectively.

The system combines various technologies to ensure it is scalable, efficient, and easy to deploy. The front-end interface is built using React.js, while the back-end services use Flask or Django frameworks with PostgreSQL for managing the database. It also includes features like QR code scanning for user interaction, real-time speech processing, and automated alert systems to improve its functionality. The application is designed with Docker-based containerization and CI/CD pipelines, making deployment on cloud services seamless and maintenance easier.

By blending speech recognition, automated alert generation, and modern web technologies, the proposed system aims to offer a proactive and scalable security solution for urban transit settings. The Urban Transit Vigilance and Defense System provides actionable insights for security analysts, enabling quicker identification and handling of potential threats, ultimately leading to safer public transportation systems.

1. Introduction

Airports, railway stations and metro systems are vital public facilities that rely on the effective and robust security systems to ensure passengers stay safe in transit. The traditional way of surveillance primarily focuses on video monitoring and manual control, and it has a lag in identifying possible threats. The Urban Transit Vigilance and Defense System is a technology-based solution that proposes to increase security by utilizing real-time, speech recognition and automated threat detection.

The system will collect audio signals from the surrounding environment in a transit mode of travel, and analyze them with real-time speech recognition. Transcribes verbal dialogue into writing. The

data is then transcribed and analysed for specific "alert words" that may indicate a security threat. The system automatically records such events with the highly accurate timestamps and send notifications such as e-mail alerts to authorized security staff. With this setup, it's easy to respond fast and Security Operations Center (SOC) analysts are better equipped to identify and handle possible threats.

The system is built using a variety of technologies, with the aim of keeping it scalable, efficient and easy to deploy. The front-end interface is created with React.js and the back-end services are written either with Flask or Django frameworks with PostgreSQL to handle the database. It also has functions such as QR code scanning for interaction with the user, real-time speech processing and automatic alert systems to enhance its use. Application is built with Docker containerization and CI/CD pipelines for easy deployment on cloud services, with easier maintenance.

The proposed system combines speech recognition, automatic alert generation, and contemporary web technologies to provide a proactive and scalable security solution for urban transit environments. The Urban Transit Vigilance and Defense System delivers actionable intelligence to security analysts, helping them to more rapidly detect and manage any potential threats, resulting in more secure transit systems.

2. Related work

Ensuring safety in urban transit environments is an important research area. The threat in public places, such as airports, metro stations, and railway terminals, is increasing. The researchers and developers have investigated several technologies to enhance safety and threat detection, such as surveillance systems, artificial intelligence, speech recognition, and automated alert mechanisms.

There are several studies that are targeting the application of AI and machine learning in security monitoring. The traditional surveillance systems are predominantly based on the use of video surveillance and human surveillance for the detection of suspicious activities. Nonetheless, these systems can have a lag time and human error can arise. Recent advancements have added automated monitoring techniques that analyze data streams in real time, which can speed up threat detection.

Speech recognition technology has also been the subject of research as an aid to security. With NLP, modern real-time speech recognition systems can now record and transcribe a conversation from speech to text. The technology has also been used in various applications, such as emergency response systems, law enforcement monitoring, and public safety analytics. These systems can detect potentially threatening or suspicious speech by checking for specific words or analyzing the meaning of the speech.

Keyword-based threat detection systems are also a research interest. These are systems that keep an eye continuously on certain “alert words” that are connected to possible threats within audio streams or communication channels. If they see these words, they automatically alert the relevant parties. This method has been used in the areas of airport security, emergency call analysis and law enforcement communication monitoring, among others, for the purposes of early identification of threats.

In addition, a number of projects have explored integrated security solutions that would integrate multiple technologies, such as cloud computing, web application frameworks, and database technologies, to provide scalable monitoring solutions. These systems can be implemented using cloud-based architectures, which can handle large amounts of real-time data and provide high availability and scalability. Docker and continuous integration/continuous deployment (CI/CD) pipelines contribute to streamlined deployment and maintenance.

Internet connectivity and automated alerts are also part of the package for advanced transit safety solutions. These interfaces give SOC analysts a way to view logs, investigate alerts and keep track of

security events as they happen. Email alerts and dashboards alert security staff to suspicious activity, enabling them to react quickly.

While these developments have occurred, much of the current systems continue to be based on visual surveillance and/or post event analysis, rather than real time conversational threat detection. The Urban Transit Vigilance and Defense System is designed to fill these gaps, involving the use of real-time speech recognition, keyword detection, automated notification mechanisms, and a scalable cloud deployment. The proposed system combines these technologies to provide enhanced proactive threat monitoring and fast response in urban transit scenarios.

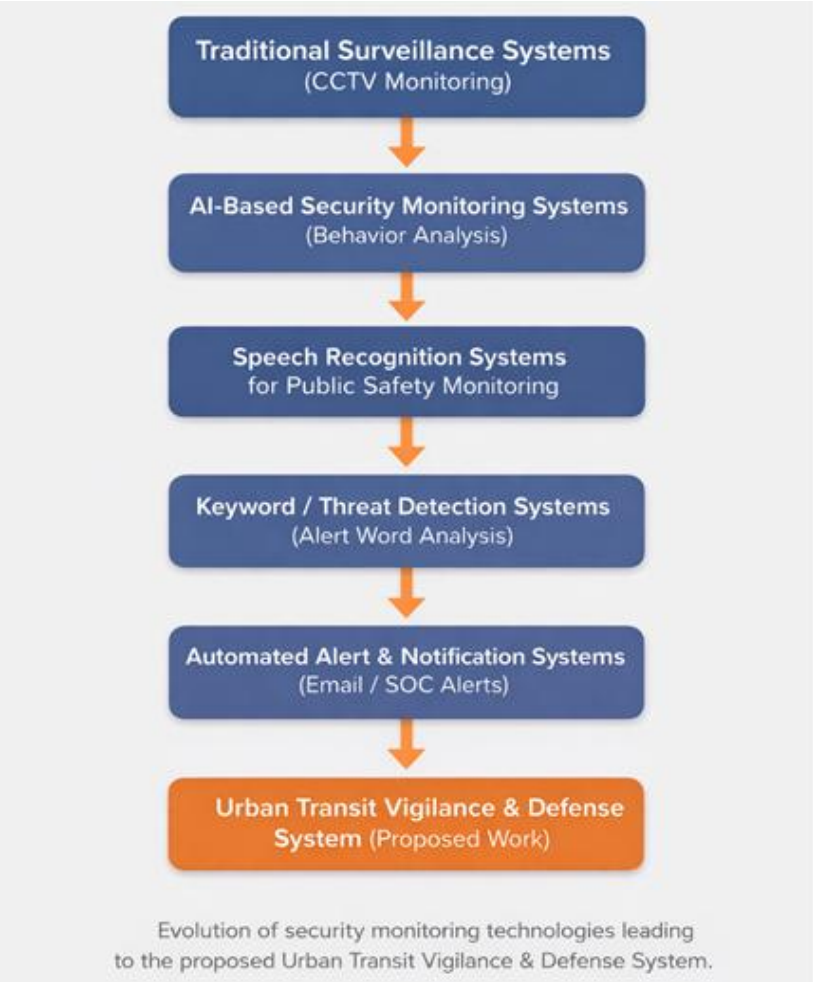

Fig. Evolution of security monitoring system

3. Research Methodology

The research methodology for the Urban Transit Vigilance and Defense System centers on designing and implementing an intelligent monitoring system that uses real-time speech recognition and automated alert mechanisms to boost safety in urban transit environments. The methodology involves several stages, including system design, data processing, threat detection, system implementation, and evaluation.

3.1. Problem Statement

Airports, railway stations and metros have large numbers of customers daily, and are vulnerable to security risks. Traditional security solutions rely primarily on video surveillance, manual surveillance, and only responding to incidents. These methods provide simple security coverage, but are heavily dependent on the human eye and require a lot of time to be able to see a threat initially as a result of operator fatigue and slow analysis. They also don't have the ability to follow more than one data source at a time.

One major problem with current surveillance tools is the lack of tools for analyzing verbal communication in transportation settings. Traditional systems focus more on visual monitoring, while suspicious conversations or threats may go unnoticed due to a lack of auditory information. Potential threats can therefore go unnoticed until after a threat has happened.

Moreover, standard security systems typically do not have inbuilt systems to detect suspicious behavior, record the required information, and alert security staff in real-time. Automated alert mechanisms can help improve the effectiveness of threat mitigation strategies by reducing response time if not present.

This scenario necessitates the need for an intelligent and proactive monitoring system, capable of interpreting vocal communication, detecting possible danger indicators and rendering timely notifications to security authorities. The Urban Transit Vigilance and Defense System (UTVDS) addresses this need by combining real-time speech recognition, keyword-based threat detection, automated alert notifications, and scalable cloud deployment. The overall goal of this system is to enhance situational awareness for Security Operations Center (SOC) analysts and response time to threats in urban transit, to improve overall public safety.

3.2. The design and architecture of systems.System Design and Architecture.

The system is modular and consists of front-end interfaces, back-end services, and database management systems. Front end is built using React.js to give an interactive and usable interface to monitor alerts and system activities. The Back-end services are developed using Flask/Django frameworks which process speech, detect alerts, and communicate among various parts of the system. All transcribed text, system logs, timestamps and alert records are stored in a PostgreSQL database.

3.3. Audio Capture and Speech Recognition

The system records audio information from the travelling environment, through microphones or audio inputs. It uses a real-time speech recognition module which recognizes speech and turns it into text. The transcription process enables the system to efficiently analyze spoken conversations. The recognized text is then passed to the threat detection module for more in-depth analysis.

3.4. Keyword-Based Threat Detection

The system has a list of specific words that can indicate possible security breaches (alert words). The speech data is continually analysed for the presence of these alert words. If a suspicious word or sentence is discovered, an alert event is triggered. All the events detected are logged with the following system information: detected keyword, system time, and the system location.

3.5. Alert Generation and Notification System

A security "alarm" is triggered and an alert is stored in the database when a threat-related keyword is found. The alert is then automatically sent to the people authorized to respond, like via email. These alerts help keep security staff and SOC analysts alert, enabling prompt action and investigation on possible threats.

3.6. User Interaction through QR Code Integration

The system provides QR code scanning capabilities for secure interaction with users and access to the system services. This function enables authorized users to communicate with the monitoring system or utilize certain functions in a secure and efficient manner.

3.7. Deployment and Scalability

The application is containerized with Docker to ensure smooth deployment and scalability. This process streamlines the installation of the system, and enables consistent performance in different environments. Furthermore, CI/CD pipelines are used to automate the testing, updating, and

deployment processes. The system is intended to be easily deployed on cloud platforms, to support large-scale monitoring operations.

3.8. System Testing and Evaluation

The last step of the methodology is the test of the system to evaluate its performance, accuracy and responsiveness. Different test scenarios are used to validate speech recognition, detection of keywords and alerts. The assessment focuses on accuracy of detection, response time and system reliability to guarantee the best performance in real-time transit environments.

3.9. Proposed Algorithm

Several algorithms for speech processing, keyword detection, alert generation and logging mechanisms are proposed to handle effective threat detection and real-time monitoring in the Urban Transit Vigilance and Defense System. These algorithms can collaborate to identify conversations that are questionable and alert security staff quickly.

To capture and pre-process the speech.To capture and pre-process speech.

This algorithm processes audio input from microphone(s) and prepares the input for speech recognition processing.

Steps:

Turn on the system and configure audio input devices.

1. Record audio information from the surroundings, continuously.
2. Use noise reduction and filtering to enhance audio.
3. Divide the audio signal into smaller frames to work on.
4. Send processed audio frames to speech recognition module.

Purpose:

To give precise and uninterrupted speech-to-text conversion.

The algorithm for real time speech recognition is 2.

This algorithm converts captured audio to text using speech recognition technology.

Steps:

1. Get processed audio frames from audio capture module.

Use speech recognition model or API to convert speech to text.

2. Temporarily store the transcribed text for analysis.
3. Send the transcribed text to the keyword detection module.

Purpose:

To accurately transcribe spoken text to text for further threat analysis.

The 3rd step is called Keyword Based Threat Detection Algorithm.

This algorithm detects the suspicious words/phrases that might indicate a threat.

Steps:

Load predefined list of alert keywords.

1. Get text output from speech recognition module.
2. When an alert word is stored, compare the word with the transcribed text.

3. If there is the keyword match:

- Note the conversation as a possible danger.
- Generate an alert event.

If no keyword is detected, then go back to step 4.

Purpose:

To detect threats from conversations with suspect language.

1. Alert Generation and Notification Algorithm

This algorithm will alert authorized personnel when it sees a threat.

Steps:

Get threat detection signal from the keyword detection module.

1. Log the detected keyword and the system's time.
2. Store the alert information in the PostgreSQL database.
3. Create an automatic notification.
4. Notify appropriate security personnel by email.
5. Show alert on the monitoring dashboard.

Purpose:

To provide quick access to threat to security teams.

1. Event Logging Algorithm

This algorithm records all the activities system has been doing and all the alerts it has detected.

Steps:

1. Accept events from system modules.
2. Log event type, time and status of system.
3. Save logs in the PostgreSQL database.
4. Give SOC analysts access to logs via monitoring interface.

Purpose:

To keep a log of system activity for investigation and analysis.

By implementing these proposed algorithms, the Urban Transit Vigilance and Defense System can continuously monitor, detect potential threats in real-time, and alert security authorities swiftly, enhancing safety and response efficiency in transit settings.

4. Research Methodology

The research methodology for Urban Transit Vigilance and Defense System is to create a smart monitoring system, capable of identifying potential threats in urban transportation zones through real-time speech recognition and automated alert systems. The methodology includes several stages such as system design, data acquisition, speech processing, threat detection, system implementation, and performance evaluation. The stages are all contributing to building a secure monitoring system that is reliable and scalable.

4.1. Research Design

This research adopts system development method that designs, implements and evaluates smart monitoring system. The primary aim is to create a system that can record live speech, transcribe it

into text, detect malicious keywords and alert the security staff. The design is designed to be precise in speech recognition and respond immediately to threats. The research fuses technologies of different domains such as Speech recognition, NLP, Web development and cloud computing to develop a comprehensive monitoring system.

4.2. System Architecture Development

The system design enables real-time monitoring and data processing. It is based on a 3-layer architecture, comprised of the presentation layer, application layer and data layer.

- Presentation Layer:

The layer has been created using React.js and enables Security Operations Center (SOC) analysts to view alerts, logs, and control system configurations.

- Application Layer:

This layer is responsible for the primary functions such as speech recognition, keyword detection, alert generation, and system communication. It has been developed with Flask/Django frameworks.

- Data Layer:

This layer contains information on user system logs, speech transcriptions, detected keywords, timestamps, and alert records, and is also stored in PostgreSQL. This allows data to be reliably stored and readily retrieved for analysis.

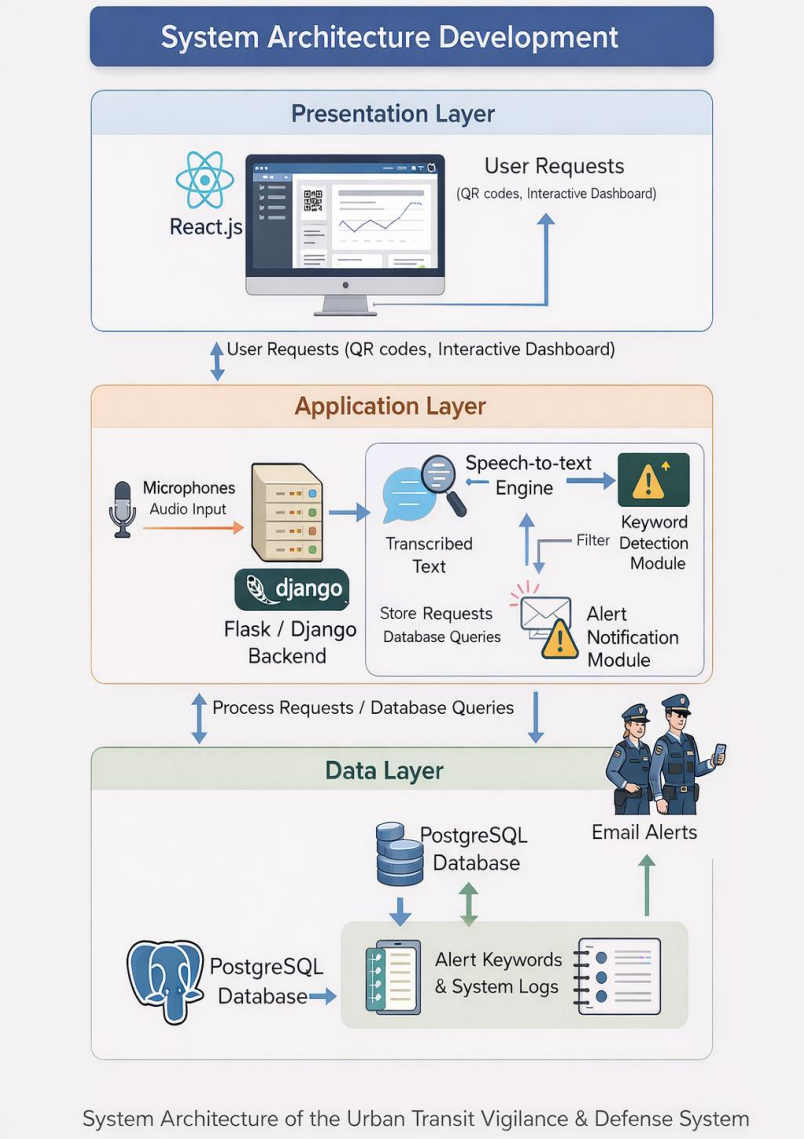

Fig. System Architecture Of Urban Transit Vigilance & Defence System

4.3. Audio Data Acquisition

The system is always collecting audio data from microphones located in monitored transit areas. The most common input devices for collecting speech data from conversations in the immediate vicinity are these microphones. Audio pre-processing techniques could be used to enhance audio quality and to mitigate background noise and include:

- Noise filtering
- Signal normalization
- Voice activity detection

are used prior to the speech recognition process. These are useful to improve accuracy of speech-to-text conversion.

4.4. Speech Recognition Processing

Real-time speech recognition technology processes the audio signals that are captured. The speech recognition module transforms spoken text into a plain text representation, using models or APIs. This process involves:

1. Audio signal processing
2. Feature Extraction from speech signals
3. Speech-to-text conversion
4. Outputting the transcription.

The outputted text information is then passed on to the threat detection module to be analyzed again.

4.5. Threat Detection Mechanism

The threat detection module analyzes the transcription to identify suspicious words or phrases that could indicate possible security threats. The system employs keyword detection algorithm that matches the generated text with a list of alert keywords. Words can be related to:

- Explosive threats
- Violent activities
- Illegal actions
- Security violations

Any of these words in the speech that are transcribed are indicated as a potential threat event.

Fig. Threat Detection Mechanism

4.6. Provide alerts and notifications.Generate and notify alerts.

If it is detected that a threat exists it automatically creates an alert notification. The Alert information contains:

- Detected keyword
- Timestamp of detection
- Source of audio input
- Event identification number

This data is kept in the PostgreSQL database for the use of authorized people and then sent to these users via emails and dashboard alerts. That way, the security team can be alerted to any potential threats quickly.

4.7. Integrate QR codes for user interaction.Enable QR Code user interaction.

It is a system that incorporates QR code scanning technology for safe interactivity with users. QR codes can be used by authorized staff to gain access to features of the system, to validate users or to retrieve particular information from the system. This feature makes the system more efficient and secure for getting access to the modules of the system.

4.8. Cloud Deployment & Containerization

The system is containerized using the Docker technology to ensure scalability and ease of deployment. Having the system running in containers means that it will run with no problems on different environments. The project also includes Continuous Integration and Continuous Deployment (CI/CD) pipelines, which automate the process of updating, testing, and deploying a system. This allows efficient maintenance and to easily scale the system to accommodate large transit networks.

4.9. System testing and performance evaluation

Once it is developed, the system is extensively tested to evaluate its performance and reliability. The evaluation process is based on a number of important metrics such as:

- Speech recognition accuracy
- Keyword detection efficiency
- Alert response time
- Dependable operation and steadiness
- Database logging accuracy

Different scenarios are simulated and tested to determine the effectiveness of the system in detecting possible threats and alerting security personnel. The results will aid in evaluating the efficacy of the system to enhance security monitoring in urban transportation settings.

4.10. Educating students on ethical and privacy concerns.Exploring ethical and privacy issues.

As it is a speech monitoring system, privacy and ethical issues are of significance. The system only looks for certain keywords in speech that are related to threat and does not keep or misuse unnecessary personal conversations. System logs and alerts are only available to authorized security personnel to ensure that the information remains confidential and protected.

The conclusions and future work are included.

The Urban Transit Vigilance and Defense System is a technology-based solution designed to increase the security of urban transportation facilities such as airports, metro stations and railway terminals.

Traditional security systems typically depend on the installation of video surveillance and manual monitoring, which are not always able to detect threats. To overcome this, the proposed system will be based on real-time speech recognition, keyword-based threat detection, and automated alerts, all of which will enhance the situational awareness and response efficiency.

It picks up the sound in the surrounding and translates spoken speech into text by means of speech recognition. The analysis of the transcribed text is done to look for potential threats by using predefined alert keywords. The system triggers alerts when it identifies suspicious words or phrases, logs the events with timestamps and notifies authorized personnel via email alerts and monitoring dashboards. This will aid security teams and Security Operations Center (SOC) analysts in their timely and efficient response to potential threats.

The architecture is a mix of modern web technologies, including React.js for the front-end, Flask/Django for the back-end services, and PostgreSQL for the database. Its enhanced capabilities, such as QR code integration, automated logging, and cloud deployment with Docker and CI/CD pipelines, guarantee scalability, reliability, and easy maintenance. This system demonstrates the potential of automated monitoring and speech recognition in improving threat detection in public transit environments.

In conclusion, the proposed Urban Transit Vigilance and Defense System plays a significant role in ensuring the safety and security of urban transit systems by providing smart security solutions that enable proactive monitoring and swift response.

Future Work

While the proposed system provides an effective framework for real-time threat detection, there are opportunities for further research to enhance its functionality and performance.

An improvement that could be implemented is the use of more sophisticated machine learning and natural language processing algorithms that could process not only the keywords but also the context and sentiment of a conversation. This would enable the system to be alerted to more sophisticated threats, which are not reliant simply on trigger words.

Multilingual speech recognition could also be added to future versions of the system, which would allow threats to be detected in multiple languages that are commonly spoken in urban transit environments. This would make the system more effective in a variety of public scenarios.

A further extension can be the incorporation of video surveillance and facial recognition systems along with speech monitoring. An integrated solution incorporating both audio and visual analysis would give a more holistic security monitoring solution, and boost the accuracy of threat detection.

Furthermore, the system could be enhanced by incorporating an app on mobile devices that alerts security staff in real-time, enabling a more rapid response and better coordination in case of an emergency. Mobile integration would allow authorities to be alerted, check their logs, and track their threats on the device itself.

In conclusion, future research could involve the development of improved cloud infrastructure and distributed processing systems to further enhance scalability and large-scale deployment. This would enable the Urban Transit Vigilance and Defense System to be operational simultaneously in several transit centres.

These enhancements could help transform the system into a smarter, more robust security solution to meet the evolving safety and threat detection needs of today's transit networks.

References

1. F. R. Byers, Speech Analytics for Public Safety, PSCR Stakeholder Meeting, National Institute of Standards and Technology (NIST), 2020. (NIST)

2. Automatic Speech Recognition of Public Safety Radio Communications for Interstate Incident Detection and Notification, C. M. Gartner, V. Vajpayee, J. Desai, D. M. Bullock, vol. 8, no. 5, Smart Cities 2025. (MDPI)
3. K., A. Rao, S. S. Bhat and V. B. Durdi, "A Literature Review on Speech Recognition and Cyber Security Systems," International Journal for Research in Applied Science and Engineering Technology, 2022. (IJRASET)
4. P. Haley, "The Impact of Biometric Surveillance on Reducing Violent Crime: Strategies for Apprehending Criminals While Protecting the Innocent," 2025, vol. 25, no. 10, Sensors. (MDPI)
5. G. Laffitte et al., "Disruptive Situation Detection on Public Transport using Speech Emotion Recognition," Intelligent Systems with Applications, 2024. (ScienceDirect)
6. M. Biju, A. Sajeev, A. Bimal, M. Fais and S. Chandran, "AI Surveillance System for Detecting Criminal Intent in Voice or Chat Communication", International Journal of Engineering Development and Research, 2025. (RJ Wave)
7. J. Villalba, S. Joshi, P. Żelasko, and N. Dehak, "Representation Learning to Classify and Detect Adversarial Attacks against Speaker and Speech Recognition Systems," arXiv preprint arXiv:2107.04448, 2021. (arXiv)
8. K. Rajaratnam, K. Shah, and J. Kalita, "Isolated and Ensemble Audio Preprocessing Methods for Detection of Adversarial Examples against Automatic Speech Recognition," arXiv preprint arXiv:1809.04397, 2018. (arXiv)
9. Y. Xie, C. Shi, Z. Li, J. Liu, Y. Chen, and B. Yuan, "Real-time Universal Adversarial Attacks against Speaker Recognition Systems," arXiv preprint arXiv:2003.02301, 2020. (arXiv)
10. A. Sharma, S. Kumar, "AI-Based Speech Recognition Systems for Security Surveillance and Threat Detection", 2023, International Conference on Artificial Intelligence and Security Systems. (SciTePress)