VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

School of Computer Science and Engineering

"AUDIO INTRUSION DETECTION SYSTEM USING DEEP LEARNING"

A project submitted
in partial fulfillment of the requirements for the degree
of
Bachelor of Technology
In
Computer Science and Engineering

By

Yash Jeena 20BCT0227
Gayathri Reddy Patlolla 20BCI0235
Pandaga Vamshidhar – 20BCT0310

Course Instructor

Professor
Dr. Anil Kumar K

AUDIO INTRUSION DETECTION SYSTEM USING DEEP LEARNING

ABSTRACT

An intrusion detection system (IDS) is a device or software programme that keeps track of system or network activity and looks for signs of hostile activity. Concerns regarding how to safely share and protect digital information are raised by the internet's phenomenal expansion and usage. In the modern world, hackers employ a variety of techniques to obtain valuable information. One of these is audio intrusion, which refers to the unauthorised introduction of sound or music into a secured space, such as a home or business, with the goal to harm or steal. Our objective is to implement an audio intrusion detection system to identify suspicious sounds and analyse audio patterns to identify potential security breaches. Deep learning technique is used to detect the intrusion in our project. This technology is advantageous over conventional methods as it can pick up sounds that cameras or other sensors might miss, including those coming from closed doors or windows.

Keywords: Intrusion detection system, Audio intrusion, Security Breaches, Deep Learning

INTRODUCTION

Internet security has become an issue for enterprises in the modern world. Since a long time ago, Web Firewalls, encryption, authentication, and Virtual Private Networks (VPN) have been used to secure network infrastructure and internet communication in the process of protecting data. A very recent addition to the category of security technology is audio intrusion detection.

Entering a secure audio environment without authorization is known as audio intrusion. Without the intended recipient(s)' knowledge or consent, it entails the act of intercepting, watching, or modifying audio signals. A variety of activities, including espionage, eavesdropping, sabotage, and the theft of private information, can be carried out through audio intrusion. Particularly in sensitive sectors like government, banking, healthcare, and technology, it poses a severe danger to privacy and security.

Host-Based infiltration is the most prevalent type of infiltration. The main method used to steal sensitive data from a network end device is host-based intrusion. A host-based intrusion is when a hacker or malicious software that has acquired access to the system gains unauthorised access to a computer or other device. The use of audio listening devices is one method for committing this kind of intrusion. End devices, such as computers or smartphones, can be equipped with audio listening devices, such as microphones or other audio sensors, to eavesdrop on crucial talks. The hacker or assailant can then use the audio data that was collected to get private information or to carry out other attacks. It can be challenging to identify the use of audio

listening devices for host-based intrusion, which poses a major security risk. Security officers may have difficulty locating these devices because they can be small and discrete. They can also be installed remotely, giving attackers access to a system without having to physically enter it.

The goal of an audio intrusion detection system is to identify suspicious sounds and analyse audio patterns to identify potential security breaches. To detect the sound of breaking glass, forced entry, or other suspicious noises, the system uses audio sensors or microphones positioned strategically throughout the monitored area. An algorithm that analyses the audio data gathered by the sensors and distinguishes between normal and abnormal sounds is often part of the audio intrusion detection system. For reliable identification of various sound patterns, this system is trained on a sizable amount of labelled audio data. The technology can set off an alarm or notify security personnel when it detects an irregular sound pattern. To give security workers more in-depth information, the system can also distinguish between various sounds, such as the sound of a door opening versus the sound of a window shattering.

The usage of audio intrusion detection systems can improve security and guard against theft, vandalism, and other security risks in a number of contexts, including residential and commercial buildings. To offer a complete security solution, they can also be coupled with other security systems like cameras and motion detectors.

An improved version of the conventional audio intrusion detection system is a deep learning-based system. To analyse audio patterns and spot potential security breaches in a specific area, it employs deep learning algorithms. With the help of a sizable dataset of labelled audio recordings, the deep learning algorithm utilised in this system is trained to recognise various sound patterns and tell apart normal from aberrant noises. The system has the ability to distinguish between different sound kinds, such as the difference between the sound of a door opening and the sound of a window breaking, and it can even pinpoint the location of the sound's source in a particular space.

The technology can set off an alarm or notify security personnel when it detects an irregular sound pattern. By removing background noise and concentrating on sound patterns, the system can also be configured to decrease false alerts. This technology can also distinguish between various sound sources, giving security professionals access to more specific information and minimising false alerts.

Overall, a deep learning-based audio intrusion detection system is a potent tool for boosting security in several contexts. It uses sophisticated algorithms and audio sensors to identify potential security breaches and warn security staff in real-time, assisting in the prevention of theft, vandalism, and other security issues.

RELATED WORK

Real-Time Intruder Detection System Using Sound Localization and Background Subtraction

In this paper, an efficient and real-time intruder detection system, based on the concepts of sound localization and background subtraction, is proposed. The system consists of a camera capable of 180° rotation and a microphone array. The system detects the presence of possible intruders in real-time, using background subtraction, and alerts the security centre on detection. One of the challenges is that the camera turns towards the direction of the source of disturbance by sound localization. Therefore, there is a delay in turning of the camera towards the source of disturbance.

Intrusion Detection System

The research of intrusion detection systems (IDS) is presented in the paper, which also suggests a brand-new IDS architecture based on the Snort IDS engine. The suggested IDS is made to recognize many forms of attacks, such as buffer overflow, port scanning, and DoS attacks. The many types of intrusion detection systems, including host-based and network-based IDS, as well as their benefits and drawbacks, are also covered in the article. It emphasises the significance of using several IDS to offer thorough defence against attacks. The study also explains how the suggested IDS operates, which makes use of a signature database, protocol analysis, and anomaly detection methods. Network traffic is compared against known attack patterns using the signature database, and unusual traffic patterns are found using protocol analysis. Traffic that differs noticeably from typical traffic patterns is identified via anomaly detection. The paper offers a general overview of intrusion detection systems and suggests a novel architecture for an IDS that can recognise various sorts of attacks.

Understanding the effectiveness of ultrasonic microphone jammer

The usage of ultrasonic microphone jammers to safeguard user privacy by preventing unauthorised audio recording is discussed in the study. The ability of an ultrasonic jammer to degrade the audio quality of recorded speech was tested by the authors. Several real-world locations, including a coffee shop, a meeting space, and an office environment, were used for the studies. The findings demonstrated that, in the majority of instances, the ultrasonic jammer was successful in degrading the audio quality of recorded speech, however the effectiveness varied depending on the particular situation and the type of microphone being utilised. The limitations of ultrasonic microphone jammers are also discussed in the article, including the fact that they do not stop the use of other kinds of recording tools, such covert cameras. The authors also pointed out that if the recording device was placed very close to the speaker or if the jammer wasn't placed properly, ultrasonic jamming's efficiency might be diminished.

Eavesdropping attack in wireless ad hoc networks under shadow fading environment

This study examines how shadow fading affects eavesdropping attempts in wireless ad hoc networks. Shadow fading is the term used to describe the phenomena where the signal intensity changes as a result of environmental conditions or the presence of obstructions. The authors examine several shadow fading situations and eavesdropping attack scenarios to determine how secure a wireless ad hoc network is. They also suggest a fresh method for detecting eavesdropping attempts based on assessments of signal strength. The efficiency of the suggested approach in identifying eavesdropping attacks, even in the presence of shadow fading, has been tested through simulation. The research sheds light on how shadow fading affects wireless ad hoc network security and emphasises the significance of creating strong security protocols that can withstand environmental changes. In practical applications, the suggested method for identifying eavesdropping attempts may be helpful for boosting the security of wireless ad hoc networks.

Internet Eavesdropping information security challenge in cyberspace

The difficulties with information security in cyberspace brought on by internet eavesdropping are the main topic of this article by Rudy Gultom. The author emphasises how easier it has become for hackers to collect and eavesdrop on sensitive information carried via the internet due to advancements in technology and communication technologies. The article examines numerous eavesdropping methods and devices, including trojans, man-in-the-middle assaults, packet sniffers, and other tools. Also covered by the author are the possible effects of eavesdropping assaults, such as identity theft, financial fraud, and reputational harm. In order to defend against eavesdropping assaults, the essay emphasises the necessity of strong security measures like encryption and authentication. To avoid listening in, the author advises using virtual private networks (VPNs) and secure communication protocols like Secure Sockets Layer (SSL) and Transport Layer Security (TLS).

METHODOLOGY

The use of deep learning algorithms in audio intrusion detection systems has become an increasingly popular and effective approach in recent years. This methodology involves training a deep neural network to detect specific audio patterns and classify them as either normal or anomalous.

Data Collection: Collect audio data that includes both normal and intrusion audio signals. This data is used to train and validate the system. The normal samples can be recorded from a quiet environment, while the anomalous samples can be simulated by adding various types of noise or other sounds to the normal samples.

Data Pre-processing: Pre-process the collected audio data to remove noise and enhance the features of the audio signals that are important for intrusion detection.

Feature Extraction: Extract relevant features from the pre-processed audio data using techniques such as Mel-frequency cepstral coefficients (MFCCs), which can capture the frequency content of the audio signals.

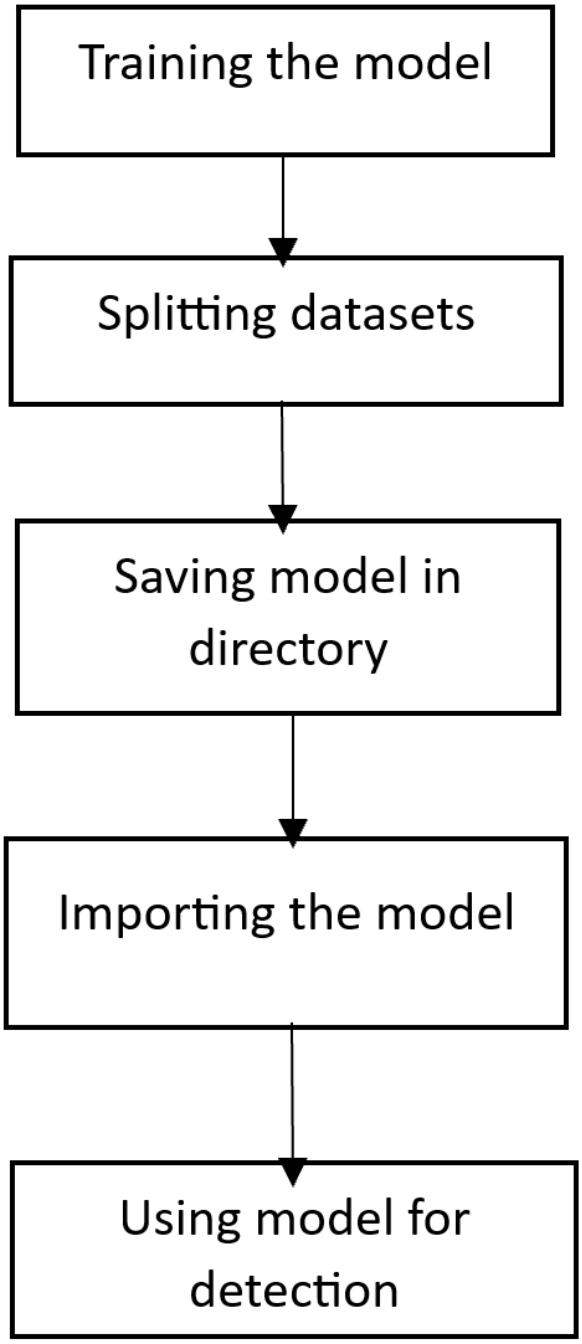
Model Selection: Choose an appropriate deep learning or machine learning model for the intrusion detection task. Commonly used models include Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Support Vector Machines (SVMs).

Model Training: The deep neural network must first be trained after completing the pre-processing stages. This entails training the network to categorise audio samples as either normal or anomalous using the pre-processed audio samples.

Validation: Validating the model's performance on a different dataset is crucial after training. Both typical and unusual audio samples that the model has not encountered before should be included in this validation dataset. Metrics including accuracy, precision, recall, and F1 score are used to assess the model's performance.

Testing and Deployment: Once the model is trained and validated, it can be used to detect audio intrusions in real-time. The system can be deployed as a standalone application or integrated into existing security systems.

SYSTEM FLOW


Training the model
Splitting datasets
Saving model in directory
Importing the model
Using model for detection

CODES

For Training the Model:

```
# Import required libraries
import os
import numpy as np
from scipy.io import wavfile
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from keras.models import Sequential
from keras.layers import Dense, Dropout, Flatten, Conv1D,
MaxPooling1D
from keras.utils import np_utils

# Define paths to the dataset and output directory
data_dir = "C:/Users/yashj/OneDrive/Desktop/dataset"
output_dir = "./models"

# Load dataset
def load_dataset():
    X = []
    y = []
    for subdir, dirs, files in os.walk(data_dir):
        for file in files:
            file_path = subdir + os.path.sep + file
            label = subdir.split("/")[-1]
            rate, data = wavfile.read(file_path)
            X.append(data)
            y.append(label)
    encoder = LabelEncoder()
    encoder.fit(y)
    y = encoder.transform(y)
    y = np_utils.to_categorical(y)
    return np.array(X), y

# Split dataset into training and testing sets
X, y = load_dataset()
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=42)

# Define model architecture
model = Sequential()
model.add(Conv1D(filters=32, kernel_size=3, activation='relu',
input_shape=(X_train.shape[1], 1)))
model.add(Conv1D(filters=64, kernel_size=3, activation='relu'))
model.add(Dropout(0.5))
model.add(MaxPooling1D(pool_size=2))
model.add(Flatten())
```

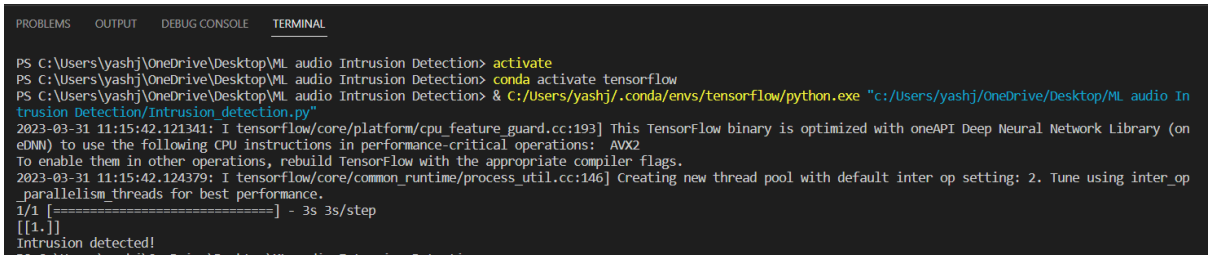
```
model.add(Dense(100, activation='relu'))
model.add(Dense(y.shape[1], activation='softmax'))

# Compile model
model.compile(loss='categorical_crossentropy', optimizer='adam',
metrics=['accuracy'])

# Train model
model.fit(X_train.reshape(X_train.shape[0], X_train.shape[1], 1),
y_train, validation_data=(X_test.reshape(X_test.shape[0],
X_test.shape[1], 1), y_test), epochs=50, batch_size=128)

# Evaluate model
scores = model.evaluate(X_test.reshape(X_test.shape[0],
X_test.shape[1], 1), y_test, verbose=0)
print("Accuracy: %.2f%%" % (scores[1]*100))

# Save model
model.save(output_dir + "/audio_intrusion_detection_model.h5")
```

For importing the model and using it for detection:

```
# Detection of Intrusion
from keras.models import load_model
from scipy.io import wavfile


# Load saved model
model = load_model("C:/Users/yashj/OneDrive/Desktop/ML audio
Intrusion Detection/models" + "/audio_intrusion_detection_model.h5")

# Load new audio data
rate, data =
wavfile.read("C:/Users/yashj/OneDrive/Desktop/testing_audio1.wav")

# Reshape audio data to match model input shape
data = data.reshape(1, data.shape[0], 1)

# Make prediction
prediction = model.predict(data)
print(prediction)
```

```python
# Decode prediction
intrusion_detected = prediction[0] > 0.5
if intrusion_detected:
    print("Intrusion detected!")
else:
    print("No intrusion detected.")
```

RESULTS



CONCLUSION

In conclusion, deep learning-based audio intrusion detection systems have demonstrated promising results in precisely identifying and categorizing various forms of audio intrusions. To accurately classify audio samples, convolutional neural networks and recurrent neural networks have been proven to be useful for capturing their temporal and spectral features. Additionally, even with scant training data, model performance has been demonstrated to be improved by the application of transfer learning and data augmentation techniques.

However, there are still some challenges to be addressed in the development and implementation of audio intrusion detection systems. These include the need for large amounts of labeled training data, robustness to environmental noise and variability, and potential privacy concerns. Future research can focus on addressing these challenges and further improving the accuracy and effectiveness of audio intrusion detection systems using deep learning methodologies.

REFERENCES

1. Amaresh, H.S. & Rao Y G, Anil & Hallikar, Rohini. (2014). Real-Time Intruder Detection System Using Sound Localization and Background Subtraction. 131-137. 10.1109/TIIEC.2014.030.
2. Tiwari, Mohit & Kumar, Raj & Bharti, Akash & Kishan, Jai. (2017). INTRUSION DETECTION SYSTEM. International Journal of Technical Research and Applications. 5. 2320-8163.
3. Chen, Y., Li, H., Nagels, S., Li, Z., Lopes, P., Zhao, B.Y., & Zheng, H. (2019). Understanding the Effectiveness of Ultrasonic Microphone Jammer. *ArXiv*, *abs/1904.08490*.
4. Li, Xuran & Dai, Hong-Ning & Zhao, Qinglin & Wang, Qiu. (2014). Eavesdropping Attacks in Wireless Ad Hoc Networks under a Shadow Fading Environment.
5. Gultom, Rudy. (2017). Internet Eavesdropping: Information Security Challenge in the Cyberspace. Jurnal Pertahanan. 3. 243. 10.33172/jp.v3i3.236.