FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

SOUND BASED LOCATION DETECTION USING MACHINE LEARNING METHODS

NURA ABDULLAHI

Master's Thesis

DEPARTMENT OFDIGITAL FORENSIC ENGINEERING

FEBRUARY 2022

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Digital Forensic Engineering

Master's Thesis

SOUND BASED LOCATION DETECTION USING MACHINE LEARNING METHODS

Author
NURA ABDULLAHI

Supervisor
Assoc. Prof. Dr. Erhan AKBAL

FEBRUARY 2022
ELAZIG

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Digital Forensic Engineering

Master's Thesis

Title:	Sound Based Location Detection Using Machine Learning Methods
Author:	NURA ABDULLAHI
Submission Date:	05 January 2022
Defense Date:	07 February 2022

THESIS APPROVAL

This thesis, which was prepared according to the thesis writing rules of the Graduate School of Natural and Applied Sciences, Fırat University, was evaluated by the committee members who have signed the following signatures and wasunanimouslyapproved after the defense exam made open to the academic audience.

	Signature	
Supervisor:	Assoc. Prof. Dr. Erhan AKBAL Firat University, Faculty of Technology	Approved
Chair:	Assoc. Prof. Dr. Türker TUNCER Firat University, Faculty of Technology	Approved
Member:	Assist. Prof. Dr. Fahrettin Burak DEMİR Bandırma Onyedi Eylül University, Faculty of Engineering and Natural Sciences	Approved

This thesis was approved by the Administrative Board of the Graduate School on

..... / ..... / 20 .....

Signature

Prof. Dr. Kürşat Esat ALYAMAÇ
Director of the Graduate School

DECLARATION

I hereby declare that I wrote this Master's Thesistitled “Sound Based Location Detection Using Machine Learning Methods” in consistent with the thesis writing guide of the Graduate School of Natural and Applied Sciences, Firat University. I also declare that all information in it is correct, that I acted according to scientific ethics in producing and presenting the findings, cited all the references I used, express all institutions or organizations or persons who supported the thesis financially. I have never used the data and information I provide here in order to get a degree in any way.

07 February 2022

NURA ABDULLAHI

PREFACE

My countless thanks to Almighty Allah for bestowing open me the capacity to carry out this research work. I sincerely acknowledge my supervisor's intellectual guidance, encouragement, and supervision, Assoc. Dr. Erhan AKBAL, who tirelessly through thick and thin painstakingly ensured that this work was balanced and completed in good time. Sir, I am most grateful. May Allah guide and preserve all that relates to you and your families in your race through life.

My heartfelt appreciation goes to my family, in the person of Alhaji Abdullahi Muhammad and mother Hajiya Saudat Salisu, for putting in me the discipline that brought me this far. And my lovely wife Rukayyah Tanimu and unforgettable children Aishat, Muhammad, and Nura (junior) for their patience and support throughout my studies. I also like to thank the National Information and Technology Development Agency (NITDA) Nigeria for their sponsorship until the end of this program.

Many thanks to all of you for your encouragement!

NURA ABDULLAHI
ELAZIG, 2022

TABLE OF CONTENTS

Page

ABSTRACT

Sound Based Location Detection Using Machine Learning Methods

NURA ABDULLAHI

Master's Thesis

FIRAT UNIVERSITY
Graduate School of Natural and Applied Sciences

Department of Digital Forensic Engineering

February 2022, Page: xi + 44

Today, sound plays a significant role in every aspect of human life. The number of crimes committed in every environment, from personal security to critical inspections in our environment, is increasing. With this increase, the identification of criminals and the clarification of events are of great importance. Audio data can be used as descriptive information to track the criminal.

Activities carried out in our homes produce different sound signals. Various activities are carried out in other locations in the house, and the acoustic properties of the sound in each location differ. As a result of the detection of the produced sound signals, the person's position can be estimated. For this purpose, a new sound dataset consisting of household sound was collected. Videos related to activities to be performed according to home locations on YouTube were used in the generated dataset.In this thesis, the popular machine learning classifiers that have been used by recent research have been applied, such as KNN, SVM, random forest (REF), etc Finally, our model, the Cubic-SVM, determined the position of the classified sound and achieved a good result with an accuracy of 96.85 for identifying the location of the classified sound, compared with existing results obtained by the researchers.

Keywords: Audio forensics, signal processing, artificial intelligence

ÖZET

Makine Öğrenmesi Yöntemleri Kullanarak Ses Tabanlı Konum Tespiti

NURA ABDULLAHI

Yüksek Lisans Tezi

FIRAT ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

Adli Bilişim Mühendisliği Anabilim Dalı

Şubat 2022, Sayfa: xi + 44

Günümüzde insan yaşamının her alanında ses çok önemli bir rol oynamaktadır. Kişisel güvenlikten başlayarak çevremizde kritik denetimlere kadar her ortamda işlenen suçların sayısı artmaktadır. Bu artışla birlikte suçluların tespiti ve olayların aydınlatılması büyük önem arz etmektedir. Suçlunu takibi için ses verileri tanımlayıcı bilgi olarak kullanılabilmektedir.

Evlerimizde gerçekleştirilen faaliyetler farklı ses sinyalleri üretmektedir. Evin içerisindeki farklı konumlarda farklı faaliyetler gerçekleştirilmekte ve her konumdaki sesin akustik özellikleri farklılık göstermektedir. Üretilen ses sinyallerinin tespit edilmesi sonucunda kişinin hangi konumda bulunduğu tahmin edilebilir. Bu amaçla ev içi seslerden oluşan yeni bir ses veriseti toplanmıştır. YouTubedaki ev konumlarına göre yapılacak faaliyetlerle ilişkili videolar kullanılmıştır. Oluşturulan verisetine KNN, SVM, rastgele orman (REF) vb. gibi son araştırmalar tarafından kullanılan popüler makine öğrenmesi sınıflandırıcıları uygulandı. Derin öğrenme, daha büyük miktarda veri ile daha iyi sonuç verse de, derin öğrenme ile kombinasyon uygulanabilir. Son olarak, modelimiz, araştırmalarla elde edilen mevcut sonuçlarla karşılaştırıldığında, sınıflandırılmış ses konumunun konumunu dikkate değer bir performansla belirlemiştir.

Anahtar Kelimeler: Ses adli bilişimi, sinyal işleme, yapay zeka

LIST OF FIGURES

Page

LIST OF TABLES

Page

SYMBOLS AND ABBREVIATIONS

Abbreviations

AI: Artificial Intelligence
KNN: K- Nearest Neighbor
SVM: Support Vector Machine
DNN: Deep Neural Networks
MFCC : Mel Frequency Cepstral Coefficient
LBP : Linear Binary Pattern
1D-LBP : One Dimensional Linear Binary Pattern
DL : Deep Learning
NN : Neural Network
ML : Machine Learning

1. INTRODUCTION

Recently, the use of machine learning applications in sound classification has been used in automated audio supervision systems[1], instrument classifications [2][3], robot navigation[4], healthcare[5], alerts to clients and buyers widely accepted[6], crime analysis systems [7], voice detection[8], sound-based disaster detection[9], environmental monitoring system[10], and many more. The act of incorporating sound classification into most applications proves its importance to the world. Audio classification is used to identify the audio class of an audio clip or recording in which the received information obtained from the audio signal has been analyzed in detail. Understanding the surrounding noise situation is essential and taking immediate action to mitigate the risk. Sound classification has become a subject matter due to its wide range of uses.

Research analysis on sound by making devices to understand its environment is one of the works related to Computational Audio Scene Analysis (CASA)[5].Machine learning systems are part of a wide range of research topics that perform similar processing tasks on human auditory systems and combine areas like machine learning, robotics, and artificial intelligence.

Detection of sound events is valuable. In automatic tagging when indexing audio, automatically analyzing audio for audio segmentation or classifying audio context. The presence of distinctive audio events characterizes audio context or scene. The sound in television news, voice recording, or video includes voices from different sources that prove the existence of an individual, like laughing, speaking, dropping off an object, animal crying, or natural causes.

Therefore, regardless of the categories, we can manage a multi-category description of our audio or video files to identify the types of audio events that occur in the file. For one to say, this event recording from 'playground,' playing with 'children,' before the unexpected appears. These are annotations on a different level. While seashore can be indirect as a context from audio events like waves and splashing of stream, the audio events of 'children' such audio events can occur in other contexts and must be explicitly recognized.

1.1. Problem statement

Sound plays a vital role in every phase of life. Sound has now been a feature in developing computerized systems in these areas for personal safety-critical surveillance. Very few systems are already on the market, but their efficiency is essential for implementation in real-world scenarios. You can use the learning capabilities of machine learning architectures to develop solid classification systems to overcome the efficiency issues of traditional techniques.

1.2. Purpose of the Thesis

Today, the number of crimes committed in every environment is increasing. With this increase, it is crucial to identify the criminals and clarify the events. There are specific Forensic vectors where information is required to be collected from the crime scene. This thesis aims to collect data by recording the activities in different environments. We have proposed a model to detect and identify activity based on sound detection and examine its performance.

1.3. Thesis Structure

This thesis study started with pieces of literature. We examined similar work done by researchers in the literature separately. We applied methods and procedures found as a result of the examinations in the thesis study to achieve better results.

The thesis mainly focuses on machine learning methods. We reviewed five different sound classification methods and environmental sound detection categories. The data sets applied for the classification were obtained from a different location at home and added to the thesis study for later use.

However, after collecting information and data, codes were written using the MATLAB programming language to improve the performance and a better success rate of machine learning and get accurate results.

In the current thesis, Chapter 1 covers the purpose of the idea; Chapter 2 includes sound, environmental sound, sound forensic processes. Chapter 3 provides machine learning algorithm methods, and Chapter 4 covers the application of machine learning techniques on sounds application. Chapter 5 is the method, results, and conclusion.

2. BACKGROUND/THEORY

2.1. Sound

Sound is an essential instrument among every living entity. Yet, the everyday existence of individuals added more kinds of audios to the regular ambient, which might be helpful or may not be. So distinguishing and understanding the audios from the normal informative one is a high time need of the day. The sound makes vibration, and sometimes with solid wave noise, these sound waves are created by vibrating that occurs on the entity. When sound waves reach our ears, they travel through our air, water, and solid objects as vibrations. These waves vibrate the delicate skin of the eardrums, and the brain recognizes these vibrations as the sounds event sources[11].

Sound is said to be one of the most important senses we receive to perceive the environment. Every action or event in our environment has an imbalanced sound. Sounds have three main attributes that can be used to distinguish between the two sounds.[12].

I. Amplitude –audio loudness
II. Frequency –audio pitch.
III. Timbre - Quality and the identity of the audio.

2.2. Sound Event

Sound events are audio clips generated from actions. Actions include speaking, buzzing, snapping fingers, running, and balling. Since childbirth, we have been trained to recognize actions and strange events as humans[13].People often learn all the sound actions they hear very efficiently and are conscious of the sound events. Only speech recognition is used to understand the podcast when listening to a podcast. From time to time, sound detection devices and sensory are used to perceive the surroundings.

2.3. Environmental Sound

Environmental sound is described as the undesirable or dangerous outside sound created in human activities, together with noise emitted via transport, street location visitors, bar spot visitors, tone site visitors, and business activities.

2.4. Sound Based Classification

A sound classification alludes to the most common setting and divides an audio recording. The sequence is one of the fundamental pieces of current AI innovation, including programmed discourse acknowledgment, remote helpers, and text-to-discourse applications. Likewise,

intelligent home security frameworks[14] and multimedia indexing and retrieval in proactive upkeep.

The human annotators can decide the substance of a sound recording and arrange it into a succession of foreordained classes or types. A proper arrangement is utilized in various regular language handling applications like Chabot's, programmed discourse acknowledgment, text to discourse, and that's only the tip of the iceberg[15].

The characterization of sounds achieved from the most familiar domain but research brings concerns in three disciplines and focus on the latest sound event Recognition (SER), The Music Information Retrieval (MIR)[14]Automatic Speech Recognition (ASR)[15], and the Environmental Sound Classification (ESC). However, considering the fundamentals of the domains above, the ESC/SER task is much more complicated. The reason is that music and speech signals have suitable patterns and organization[16], The audio signals in Environmental sound classification possess a small Signal to Noise Ratio (SNR) are generally unstructured.

To compete with Environmental Sound Classification, call out, different element extraction procedures and many AI models have been examined[13]. The area of ESC is vast and involves a large number of distinct datasets[17]. The work considers the known and familiar ESC datasets, ESC-10, ESC-50[17], and the Urban- sound8k (US8K) [18]. The famous models have been carefully applied to different classification methods. SVM, Random Forest Ensemble (RF), and KNN were used with accuracy on the ESC-10 and ESC-50 datasets.

Piczak et al. [18]. Has adopted a convolution neural network (CNN) and a remarkable improvement in the outcome, that proved raise with 7.8%. The approach also achieves a 73.7% accuracy for the US8K dataset. In previous, CNN has been considered in classifying images, and with these results, it is also an excellent tool for organizing ambient noise. Chen et al. We also used extended convolution for the ambient noise classification task on the US8K dataset[19].

2.5. Sound Classification Types

Environmental Sound Classification; this is a classification of sounds established in different environments. It could be due to the detection of city sound samples, like Sirens, voices, horns, and road structures. It can use for various securities system purposes, such as detecting the sound of glass breaking and predictive maintenance by detecting abnormal noise in factory machinery. Wildlife observation and protection are also required to distinguish animal requirements.

Music classification: This is the process that classifies music based on factors such as the category and apparatus you play. Classifications are used to categorize audio libraries by type, improve suggested algorithms, and ascertain trends and listener preference by analysis.

Natural Language Classification; Is there a characterization of standard language accounts dependent on communication in language, tongues, semantics, or other phonetic qualities? This sort of discourse grouping is generally usual in Chabot and menial helpers since it is a human language order; however, it is likewise usual in machine understanding and text-to-discourse applications.

Acoustic Data classification; This is comparable to acoustic occasion location, which recognizes where recorded the sound sign. You can realize conditions like cafés, schools, homes, workplaces, and roads at the end of the day. It assumes an essential part in checking and involving the environment in building and keeping up with sound libraries for sound media.

2.6. Sound Forensics

Sound Forensics is a science that deals with the collection, evaluation, and analysis of sound recordings that can be presented as acceptable evidence at the end of the day in court or other official locations. Evidence obtained is often deliberately taken from audible recording systems such as ring voice recorders, automatic call center recordings, and surveillance tapes obtained as part of law enforcement criminal investigations. Is it part of an official investigation into an accident, fraud, defamation proceeding, or other civil schedules? Evidence may be erroneously collected in some cases, such as soundtracks extracted from electronic news collection systems. In any case, audio evidence should be evaluated for its credibility, content enhancement and interpretability, and relevance to research objectives[20].

2.7. Sound Based Forensics Procedures

Like different types of measurable proof, have the sound and video accounts given an ongoing, onlooker record of wrongdoing that will provide the agents with what unfolded. For instance, a reconnaissance camera can catch underway in a bank, similar to burglary activity, or a secret camera records a covert sting activity[21].

It has been long; the numbers of recorded acoustic and video sources are multiplying for research. Surveillance systems and recorders from any source can be used in a location, including traffic confluences, parking places, ATMs, and mobile phones.

Howsoever, premium audio recordings are often unavailable. Forensic audio and video expertise can help. The experts can use many techniques to enhance recordings and make audio recordings more audible, emphasizing details and providing more transparent images of the event. It will then help you perform a forensic examination of the recording.

Today, audio forensics applications use digital signal processing techniques such as discrete Fourier transform and adaptive filtering.

2.7.1. Digital Acquisition

Acquisition of digital audio recordings has always followed the most acceptable protocols in the scientific community. SWGDE is the organization that established the protocol where audio recordings acquired digitally are divided into three different categories as listed below; examine the chain original request, and the tapes must be retrieved using the acceptable procedures[21].

Chain of custody examination; The main thing is to scrutinize the chain of authority. How to find out what equipment is used for evidence? How the evidence stored from its creation until the expert was handed it over to the court? Are there valid guardianship archives and reports that make up the care chain? From time to time, official protocols based on legal requirements are incorporated to enhance the legitimacy of noise protection. Checking the chain of authority reveals irregularities. It, in principle, does not need to be valid or severe.

Establishment; Forensic scientists can draw evidence from the first source of information. They are usually created and establish an accurate management process when done correctly. This recovery interaction should be archived through video recordings or images to provide an accurate record of the professionals' actions during the recovery cycle. In the unlikely event that it is foolish to expect documents to be restored, the measurable master should carefully examine the entire archive and report revealed at this point, along with evidence. In any case, if the chain of authority cannot be established, the Chain of Custody is where the scientist will have to rely on various strategies and his ability to judge its credibility.

Request for original; "When in doubt, a legal sound research facility should demand the first chronicle or the soonest age accessible. A unique account is the main appearance of sound in a recoverable put-away arrangement. If the first chronicle is on simple media, playback and duplication depend on actual cycles that present commotion and debase the sign, regardless of whether marginally. A duplicate of a simple chronicle can never be a definite copy. A unique advanced chronicle is a piece stream that can create an acoustic sound sign. And make precise of that piece stream. With advanced proof, each phase of replicating can be careful with no deficiency of value between ages. The precision can be tried and affirmed using a hash work. Accordingly, a piece stream copy of a recorded document is comparable to the first."

Retrieval method; "Methods for making sure about the recorded proof should be assessed dependent on their impact on the recorded sign, and the accessible technique for move safeguarding the evidence in a condition as near the first as conceivable ought to be picked. Utilize various methods for assortment on the off chance that it isn't obvious which accessible methods will deliver the highest quality."

The retrieval methods are as follows:

- Original recorder or recording system
- The original storage medium of the image
- Forensic copy obtained from the original file
- Transferring data file from original storage media
- Finally, digital signal transfer, copy.

2.7.2. Audio Analysis

Sound Enhancement; Forensic audio examinations frequently contain recordings that have been made furtively or beneath occasions that didn't allow best microphone placement or optimized signal/noise. Therefore, usual audio may be compromised through additive noise, distortion, lousy equalization or immoderate reverberation. The most common enhancement duties contain noise discount of recorded speech to beautify intelligibility, so a frequently prepared written transcript.

Enhancement measurable expert can recuperate the evidence from the principal source; generally speaking, that will make and set up a genuine chain of authority, IF DONE PROPERLY. Recover connection ought to be chronicled through video or pictures to give an accurate record of what the master did during the recuperation cycle. If it is crazy to hope to recuperate the recording(s), the quantifiable expert ought to mindfully encounter the whole of the files and reports displayed with the evidence. Regardless, on the off chance that can't set up the chain of power. Logical experts ought to rely upon various procedures similarly to their ability to choose the realness of the care chain.

The spectrogram shows both the occurrence components of the recording and the levels of these frequencies over time. It's the most useful tool for audio forensics professionals because it visually displays everything that happens during audio in one window. This allows professionals to identify and deal with individual harmful noise during recording. During the processing of audio, artifacts can easily be introduced into the recording. These artifacts are not needed noise generated by various processing and firmness techniques. If you think that the purpose of audio enhancement is to eliminate extraneous noise, the introduction of artifacts is the exact opposite of what you would expect when working with recordings. Many things can lead to artifacts, but the easiest way to explain the cause is to overdo it.

Authentication; Authentication is the manner so that it will with medical reality the authenticity of the activities which might be represented, in addition to the integrity of the sound recording and the end whether or not or now no longer the audio recording in query has been tampered with. In this period of virtual audio, edits may be made and included very quickly. The audio modifying software program like boldness is to be had online. It might modify the activities or verbal exchange that at the beginning befell in virtual audio recordings. Step one is to set up a

chain of custody when authenticating audio recording. While it miles steps one, the chain of custody does now no longer, in and of itself, set up a recording as being authentic.

The processes of authentication are as follows;

The listening critically; This ought to be the initial step to get comfortable with the soundproof. On the off chance that an alteration is found during the interaction, they are generally as sudden changes. Distinguishing these progressions is difficult and accompanies insight. Scientific masters put themselves in a peaceful, secluded room during this cycle to keep away from any external unsettling influences. Top caliber, proficient grade checking earphones and excellent studio screens (speakers) are best for basic listening investigation of computerized sound accounts. Proficient quality earphones and speakers have the flattest recurrence reaction. At the end of the day, it delivers a nonpartisan and normal sound.

Electronic measurement; Figure 2.1 below is how after performing the listening process, forensic professionals need to examine audio evidence by electronic measurements by determining the specific frequency and background noise of audio or other sources. In this way, you can also measure the level of recording and various frequencies. Once the frequency range of the audio grows or shrinks, or if the frequency range shifts, the edit could have been there. Unexpectedly, mysterious changes in the noise floor, and the presence of different background noise, can be a sign of processing occurs. The spectrogram and frequency analysis are shown in Figure 2.1 shown the details of the signal can be seen using these methods.

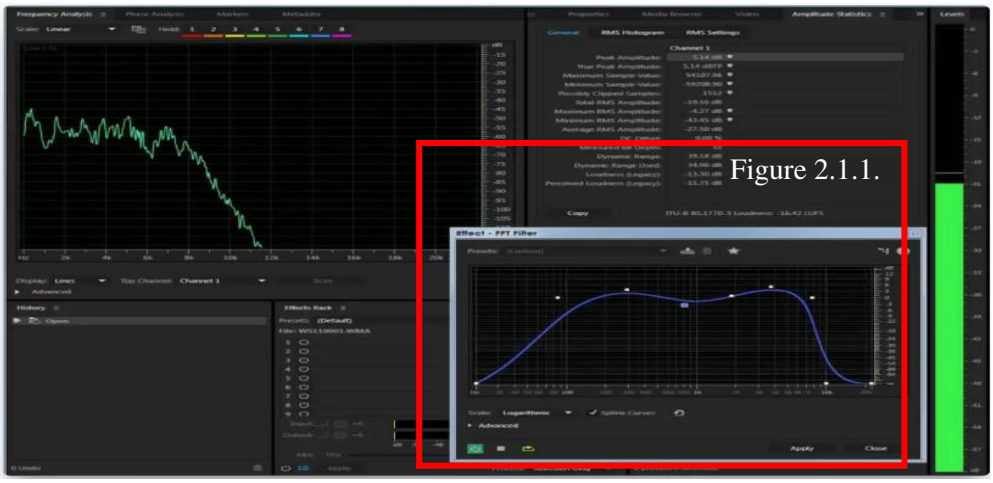

Figure 2.1. Electronic Measurement[22]

Visual Inspection; The next is to visually inspect the waveform and the nature of the spectrogram of the audio as can be seen in Figure 2.2 below. Once there are unexpected breaks in the waveform of a recording, there may be a sign of editing. The audio recording waveform will be seen inverted when zoomed visually. Figure 2.2 below shows how frequency will represent strange colors of the frequency spectrum in the spectrogram. And the background noise can be seen very clearly in this

view, which helps identify breaks in the sound. Every recording has certain background noise, even if it is barely audible. Looking at the spectrogram, noise floor interruptions may indicate processing, and even changes that may occur to the noise floor volume may indicate processing.

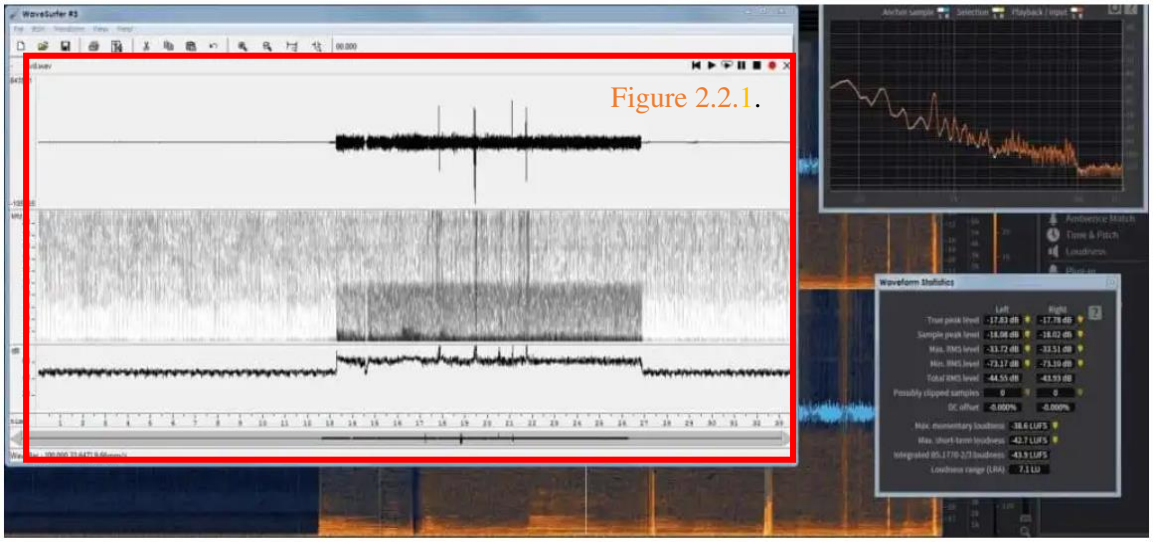

Figure 2.2. Visual Inspection [21]

Digital audio evidence; Different information can be obtained while analyzing the voice recognition process. Hence, computerized sound accounts contain metadata that contains data concerning how the recording was made and which gadget made the recording. At the point when you load a recording into an editable encoding program, the recording frequently has an impression called HEX data that demonstrates the product used to record the sound. In any case, to contrast the metadata and the first metadata, you want to make an example record. Likewise, contrasting the metadata and the HEX data in the two records can assist with distinguishing information irregularities. This is an indication of treating.

2.8. Sound Event in Everyday Environment

Anticipate sitting on the spot of a town, you can hear and perceive a progression of sounds: vehicles cruising by, individuals talking, their strides as they pass, and the steady falling precipitation [23].It's only natural to recognize all these noises and understand the perceptual scene as a sound event on an urban street. But this is the result of many years of "training", the contact and learning of all kinds of sounds, their origins, and the connections between their names in everyday life.

Our environment has always been many sources that make mixing audio signals difficult. Human hearing is specialized in classifying sound sources and drawing attention to the sound

sources of interest. This event is called the mixture Party Effect and is similar to focusing on conversations in a noisy room. Perception groups the spectral time information of an acoustic signal into an audio object, making it possible to understand a group of noise or sound as a whole[24]. A complex set of sounds is perceived as instances of a sound event, such as how a dog barks or walks. Automatic Sound Event Detection (SED) technology is primarily used to detect what is happening and when it is happening in an audio signal. In reality, you need to know which temporary instance of the various tones of the audio signal is active.

The sound detection process is different for each application. However, for a typical sound detection system, sounds like birds, cars, and footsteps passing by the target sound are ambient. In the literature, these are sometimes referred to as non-verbal and non-musical sounds.[18]A natural sound examination from more experienced discourse or music investigation errands. Sound incident identification assignments and ordinary discourse or music examination undertakings additionally have various purposes, because the impression of discourse, music and ecological sounds is likewise unique: music listening centers around the stylish nature of sound, while voice discernment centers around language or paralingual data, Daily listening plans to recognize the sources of the sound [25].

2.9. Challenges of Environmental Sound Detection

Creating automatic sound detection systems is hampered by several challenges, including those related to the nature of the sounds detected and how they occur in the natural environment and those related to data acquisition and annotation techniques. These determine the challenges that machine learning techniques must overcome in the learning process. Several challenges hinder the detection of sound location for automating the system. Some are nature-related, the sounds to be recognized and how they appear in their natural form environments, and some are related to data collection and annotation events. These, in turn, determine the challenges overcome by the techniques of machine learning in the learning process.

Acoustic characteristics are extensive in sound events [17]. Some sounds are concise and seem temporary, while others are events such as "footsteps" and "Animal crying"- a sequence of basic components. In a typical application for sound event detection, the target sound event is far away; the microphone greatly affects the event depending on the acoustic transfer function. Also because of the distance, the sound weight level of the target sound in the microphone events when received by a microphone can often be low than any other sounds that occur in the environment to the complication of detection.

Polyphonic is a various sound that can be active at the same time in the natural environment. Polyphony also exists in music, but often the fundamental frequency forms a small integer ratio because of the constrained way that overlapping sound events are in harmony with each other [26].

Sound signals have characteristic acoustic properties. There is no specific standard rule defining the audio signal produced. Statistical methods such as manual analysis of the signal band and counting of the repetitions are used for modeling the combined sounds. The number of event classes that can occur is unlimited, as there can be any presence and environmental sound in an audio signal [27]. The detection of environmental sounds is quite different from speech classification. Speech recognition has a certain set of features in any language. Recognition can be achieved by using certain features of the language.

The complicated situation is more by the lack of an established ontology for the general description of sound categories. The category of sound is usually defined according to some common characteristics [28], but this leads to a high degree of imprecision in the definition of the category. Figure 2.4 shows how an audio signal changes concerning temporal patterns.

In reality, each sound event detection application targets different sound classes and is used in different environments. Due to this variety, there is no universally applicable dataset or model for detecting sound events, and application requires data collection and system development to meet specific requirements. This is a very different approach from speech recognition, and it is hoped that a system that can process all sound combinations will be developed. In Figure 2.3, an example of temporal information of auditory scene and sound events is shown.

Figure 2.3 Sound event detection: auditory scene and temporal information of sound events [23].

2.10. A general approach in Machine learning for sound event detection

The main approach to tackling the sound event detection task is based on supervised learning[23]. It can be seen in Figure 2.5 below. It trains an acoustic model using a training set of voice recordings and reference annotations for class activities. The annotation contains binary information of each target sound class of temporal activity. It describes, by the hour, whether the class is active.

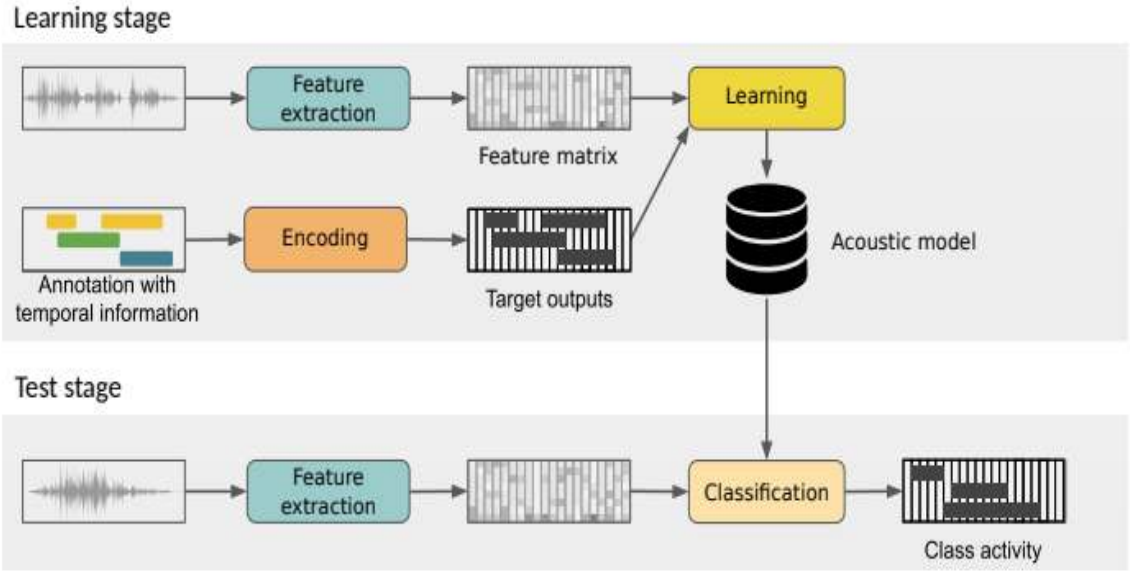

Figure 2.4. Overview of a sound event detection system, where activities of classes in short consecutive segments of audio are estimated by multi-label classification [23]

We can detect events by sound classes when the system must make available the output of each category under consideration to indicate whether the sound is active in that class. During the learning phase, the system learns the communication between the features extracted from the audio signal and the actions of each category**.** Annotations are represented as a binary matrix in which each element represents a class of active (1) or inactive (0) in a short period. The system receives the features extracted from the test audio recording and provides the output presented according to the same system. This is a matrix showing the binary activity of each sound class in consecutive time segments. A single sound event detection system is used to predict the activity of multiple classes of sounds that may be active at the same time, with multi-class, multi-label classification on each segment. It is no longer referred to as detection but rather as tagging.

Sound event detection was proposed for the temporal activity of noise in audio recordings. The label of the temporal activity pattern is estimated in it to find individual sound instances. A multi-class multi-label classification into consecutive time segments of the same audio example will produce output as shown in Figure 2.5. The output resolution is determined by the length of these segments of the target and can be as short (usually 20,100 ms) as the systematic audio frame. Examining frames define the length of calculated audio features, but the system can make predictions at the frame level or different resolutions. Suppose the system provides a single prediction for each class throughout the audio file and does not output a separate activity pattern for each sound. In that case, the task is called tagging rather than detection.

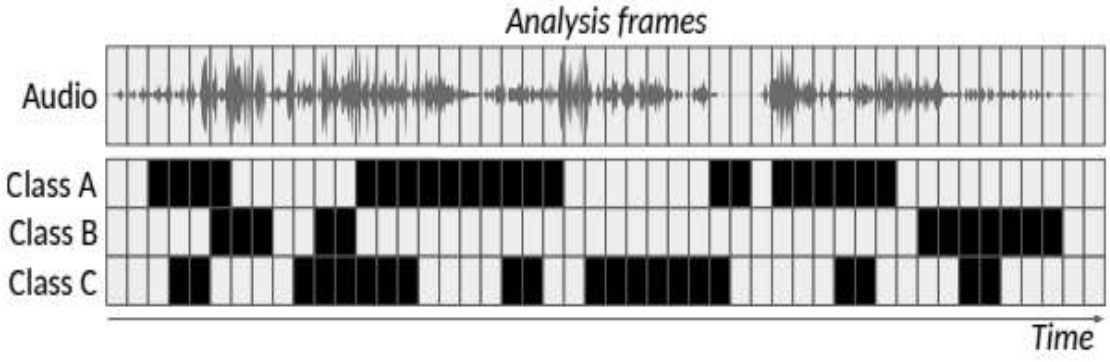

Figure 2.5 Sound event detection as multi-label classification of short consecutive audio segments [22]

3. ENVIRONMENTAL SOUND CLASSIFICATION USING MACHINE LEARNING TECHNIQUES

Environmental sounds are quite easy to detect by humans. However, the detection and automatic classification of sounds by computer-aided systems is a challenging task for researchers [29]. Automatic sound detection and classification systems try to detect whether there is a target sound by using the acoustic properties of sound signals recorded from an environment. To fulfill this task, all the characteristics of the sound must be determined. It defines the machine learning (ML) processes that the machines can make decisions by learning different signal characteristics. ML is a broad field that will enable complex problems to be solved by machines more quickly and effectively.

3.1. Machine Learning Approaches to Activity Detection

Certain processes are carried out so that machine learning methods can be used for activity detection. First of all, it is necessary to extract the features of the collected signals and select the specific features. Obtained features are given as input to machine learning algorithms. Then the features are classified using different classification algorithms. Different methods and algorithms have been used in the literature to develop activity detection models. In the thesis study, machine learning algorithms are emphasized[30]. Obtaining meaningful information from raw audio signals depends on the type of input data. Machine learning algorithms used in the literature are shown in Figure 3.1 [31].

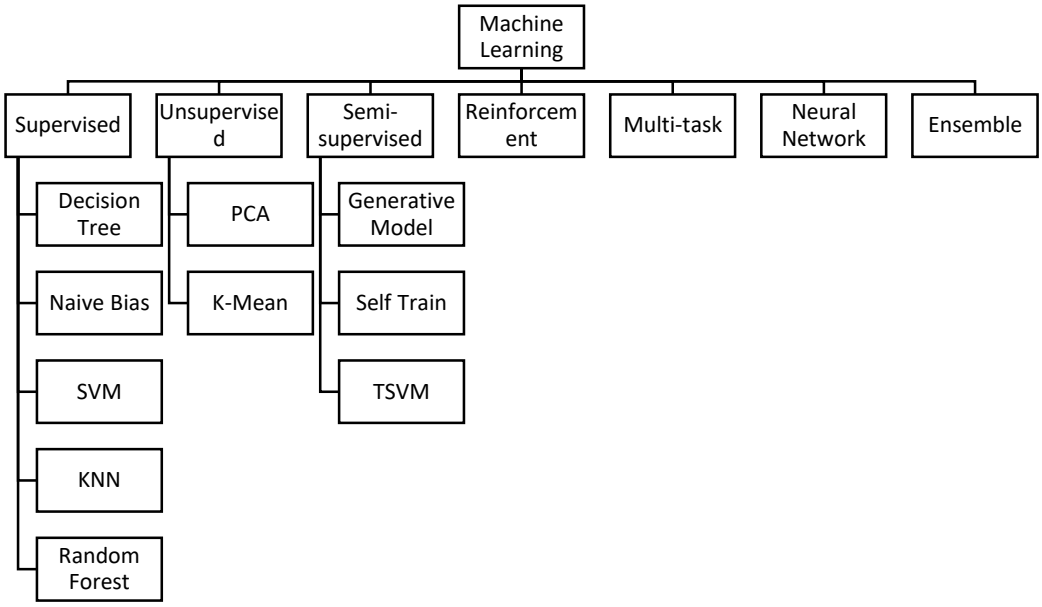

Figure 3.1 Types of well-known machine learning

3.2. Machine Learning Techniques

ML techniques include a large class of algorithms that start with metric methods such as solution trees, genetic algorithms, KNNs, SVMs, statistical methods, Bayesian networks, and end with artificial neural networks [32][33]. This direction is to solve the core problem of intelligent systems, all other actions that predict the evaluation of the current object (situation).

One of the practical applications in which machine learning technology has been widely used since the 1970s is the extraction of commercially available minerals. Artificial neural networks are used, for example, in petrography for the analysis of log data, lithology for the assessment of mineral resources, and seismological exploration[34].

The research paper [33] is dedicated to the application of neural networks and for solving the practical problem of interpreting log data from oil production. Research results[33]describe some of the consequences of applying a neural feed-forward network for the interpretation of geophysical borehole data during uranium exploration. Still, the use of machine learning is much more extensive. These include medicine [33], biology, robotics, community facilities and industry [35], services sector, ecology, innovative communication systems, astronomy.

Different machine learning algorithms are used to solve data problems. Researchers pointed out that no single algorithm best solves the problem in data science. The algorithm used depends on the type of problem being solved, the variables, and the type of optimal model.

3.2.1. Supervised Learning

The supervised learning model works based on the input-output example. There are matches in the input parameter to the appropriate output parameter. It also derives a signed function consisting of training sequences. Supervised learning algorithms expect external input. The entered datasets are divided into two training and testing. Training data has productivity factors that need to be characterized. Generally, algorithms learn from training data. It then uses the test data for classification prediction [36]. The diagram of the supervised learning model is shown in Figure 3.2.

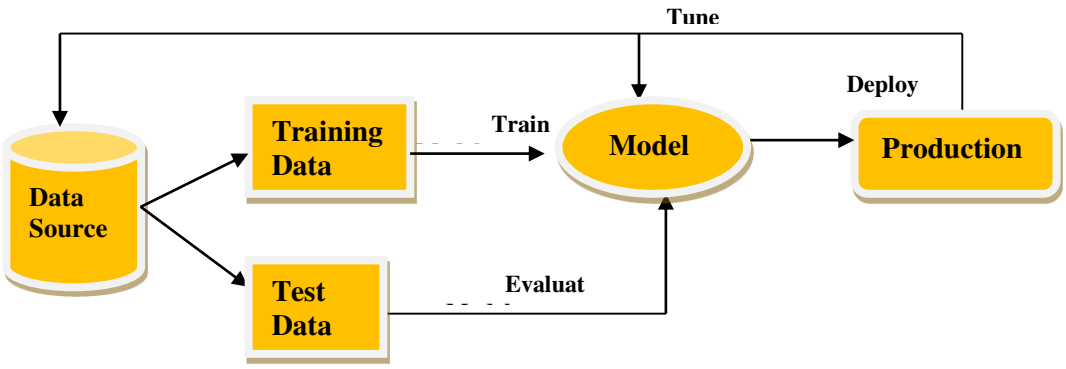

Figure 3.2. Supervised learning workflow

Table 3.1 below shows five examples of records data according to a Label and Unlabeled. The next column is titled “Decision for labeling” talks about the possible condition for each data example we have. And the third one is the possible decision of the implemented labels. The last column expresses which supervisor can take action. It is very clear that in all the first four cases described in Table 3.1 machines can be used, however, the concern is how the accuracy? The recognition in image, sentiment analysis, and speech detection technologies has been in use over time, but before we compare with human performance upgrades are required. For the case of tumor detection, it would be difficult to label the X-ray data, since specialists are required in this area.

Table 3.1.Example of labeling issues

Data Example Unlabeled	**Decision for Labeling**	**Possible Tags**	**Supervised**
Tweet	Feeling	Positive/Negative	Human / Machine
Photograph	Includes house and car	Yes/ No	Human / Machine
Sound recording	Spoken word football issues	Yes/ No	Human / Machine
Video	Use of guns in the video?	Violent/Non-Violent	Human / Machine
X-ray	Tumor on x-ray found	Yes / No	Human / Machine

3.2.1.1. Random Forest Classification

Random forest is one of the supervised learning techniques that are popular as a machine learning algorithm that has been used for solving problems in both regression and classification. The random forest algorithm combines different classifiers for solving complex problems. The emerging model advocates the ensemble learning method to increase its performance.

To improve the prediction accuracy random forest involves a chain of decision trees with a variety of subsets of the given datasets. Random forests take estimates from each tree and estimate the final output based on a majority vote of the estimates, rather than relying on a single decision tree. The more trees you have, the more accurate you are and the more you can avoid the problem of overfitting. The diagram in Figure 3.3 below explains the operation of the Random Forest algorithm [26].

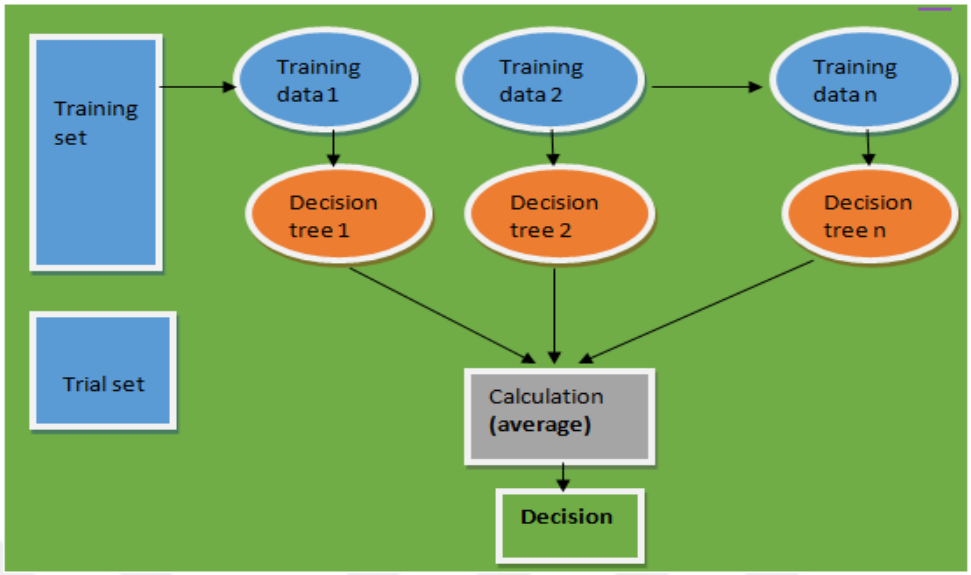

Figure 3.3. Random forest example

3.2.1.2. Decision Tree

The best performance in speed and frequency are among the preeminent qualities that made decision tree to be the most widely used machine learning technique by the researchers in classification. Decision trees usually operate in two stages, the tree creation, and classification stages. It was during classification the rules for classification are applied from the tree. The process for deciding root to leaf and followed the branches.

The diagram in Figure 3.4 below displayed the decisions and results in the tree format of a decision tree. The events represent the nodes while decision rules represent the edges of the graph respectively. Each tree consists of a node and a branch where each node is a single attribute within the classification group and the node expect a value from each branch [35].

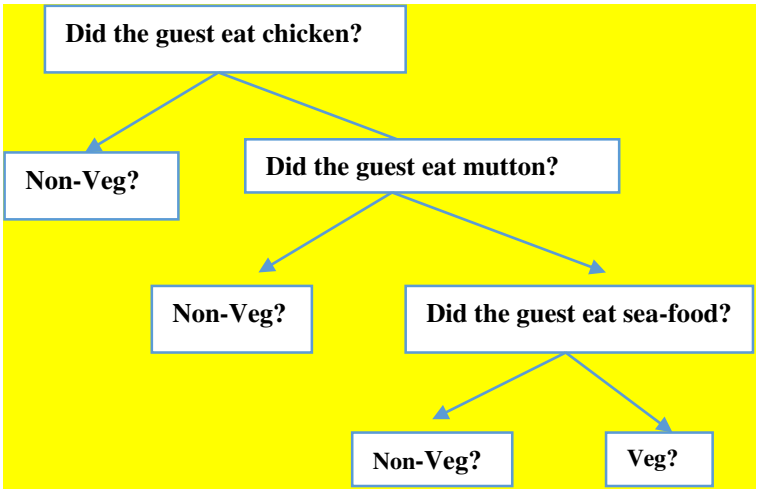

Figure 3.4. Decision tree

3.2.1.3. Naive Bayes

This is a method that defends on Bayes theorem with predictors' independence assumptions for classification. Naive Bayes is primarily for text Classification manufacturing. The purpose of the classification depends on the conditional probability of what's happening and the clustering.

The classification technique here mainly defends Bayes Theorem and is used for high inputs dimensionality [37]. The mathematical expression for Bayes' Theorem is as follows: X is a data sample with an unknown class label and a hypothesis H so that the data sample X can belong to a specified class C. This expression, P(C|X), is used to calculate the probability of P(C|X) after P(C), P(X), and P(X|C). P(C|X) is the posterior probability of the target class in Bayes' theorem; therefore, P(C) is the prior probability of the class, while P(X|C) is the probability of the estimator, given class, P(X) is the prior probability of the estimator of the class. Below is formula 1.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \quad (3.1)$$

$$P(c|X) = P(x_j|c) \times P(x_2|c) \times \ldots \times P(x_n|c) \times P(c)$$

Advantages

• Short computation time for training.

• Good performance

• Classification improvement by removing irrelevant features.

Disadvantages

• The accuracy of some datasets is low compared to other classifier.

• Improving the performance of the results depends on the recordings.

3.2.1.4. Support Vector Machine(SVM)

Another most widely used and modern technique is the Support Vector Machine (SVM). Support vector machine is a supervised learning model with relevant learning algorithms that are very popular and is used by researchers for classification and regression analysis for a long period.

The support vector machine (SVM) performs linear classifications as well as non-linear classification efficiently using kernel tricks by mapping inputs implicitly to higher dimensional feature spaces anywhere [24]. A margin among the classes was drawn. And edges are also considered to maximize the distance between the edge and the class, minimizing classification errors.

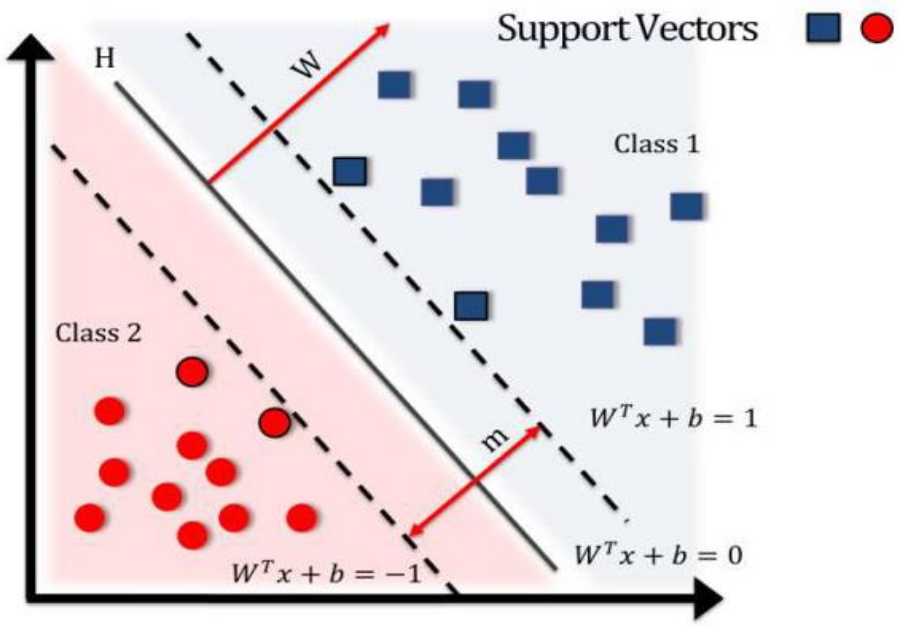

Figure 3.5.Support vector machine[35]

Support Vector machine constructs hyper-plane in multidimensional space that separates different class boundaries and the number of dimensions called the feature vector of the dataset as can be seen in Figure 3.5 such as. Table 3.2 shown SVM pseudo code.

Table 3.2 SVM Pseudo code

Support Vector Machine Pseudo code

Initialize yi = yi for i∈I

repeat

compute svm solution uu, b for data set with imputed labels

compute outputs ii = (uu, xi) + b for all xi in positive bags

set yi = sgn(fi) for every i e i, yi = 1|

for (every positive bag bi) end

if (liei (l + yi)/2 == 0)

compute i = argmaxiei ii*

set yi = 1*

end

while (imputed labels have changed)

output (uu, b)

3.2.1.5. K-Nearest Neighbors Algorithm

The KNN algorithm is a supervised machine learning algorithm that aims to solve classification and regression problems. Its advantage is that it is easy to apply. However, the increase in data size negatively affects the performance.

K-nearest neighbor is defended on the rule that instantaneous environment contains similar samples. Based learning methods like nearest neighbor, where lazy learners like Instance-based classifiers store all training examples, and until a fresh, unlabeled example needs to be classified, a classifier will not be created. Unlike eager learning algorithms, Lazy learning algorithms involve less computation time during the training phase especially, neural networks, decision trees, and Bayesian networks, although during the classification process more computation time.

Classification in KNN defends learning by correspondence, comparing a given test sample and training examples similar to it. A sample X to be classified, X is assigned the class label to which the majority of its neighbors belong its nearest neighbors are then searched.

If the K value is too small and the selection of K also affects the performance of the nearest neighbor algorithm due to the noise of the training dataset, the ANN classifier can be prone to overfitting. However, if k is too large, the nearest neighbor list may contain data points far from that neighborhood, and the nearest neighbor classifier may misclassify the test sample

K-nearest neighbor operates data knowing linked in a feature space. Therefore, all on locating the distance between data points are considered sequentially. Euclidean distance is used according to the data type of the data classes used. With the given value of K, which is used to find the total number of nearest neighbors determining the class label for the sample? If K=1, it is called the nearest neighbor classification.

Working Structure of K-NN

When using the K-NN classification, below are the steps applied.

• K value is reset,

• Determining the differences between input and training data,

• Sorted the distance,

• The nearest neighbors of the top K are taken.

• Apply the simple majority.

• Labels classes with more neighbors for the estimated input sample.

Figure 3.6 shows X, Y, and Z as three example classes. To find the class label for the data instance P. the value of K=5 and the calculated Euclidean distance for each. The four nearest neighbor samples fell into the X class label as was found that the pair of samples, while the only bunch belonged to the Z class label. Consequently, The main class for this instance P is given as the X class [5].

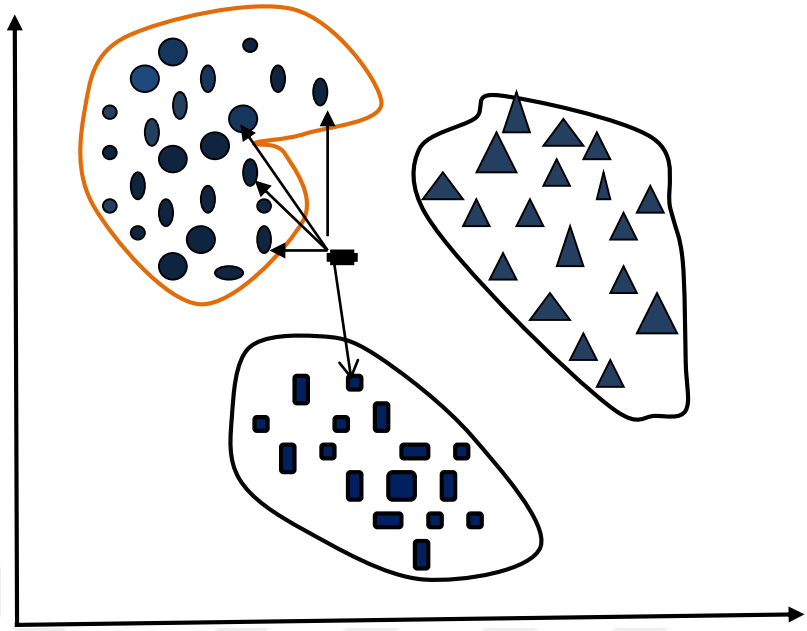

Figure 3.6 An Example of KNN Classifier

KNN classifier Advantages

• Very fast in training.

• Understanding easy to apply.

• Noisy data resistance.

• Performs well in applications where a sample can be used.

KNN classifier Disadvantages

• Memory limitations.

• Supervised learner works slowly.

• Memory limitation.

• Sensitive to the local nature of the data.

3.2.2. Unsupervised Learning

Unlike supervised learning above, this is known as unsupervised learning because there is no correct answer or trainer. The algorithm represents interesting structures in the data and its device to discover. The algorithm learns some features from the data. And the new data uses previously learned characteristics to identify the class of data when introduced. Unsupervised learning is primarily used for feature reduction clustering.

3.2.2.1. Principal component analysis

Principal component analysis -Is an orthogonal transformation that uses statistical techniques to transform a set of experimental values of potentially correlated variables into a set of values of

linearly uncorrelated variables. As a dimensionality reduction technique, the dimensions of the data are reduced to make calculations faster and easier. It is used to describe the variance-covariance structure of a set of variables in terms of linear combinations as can be seen in Figure 3.7 such as.

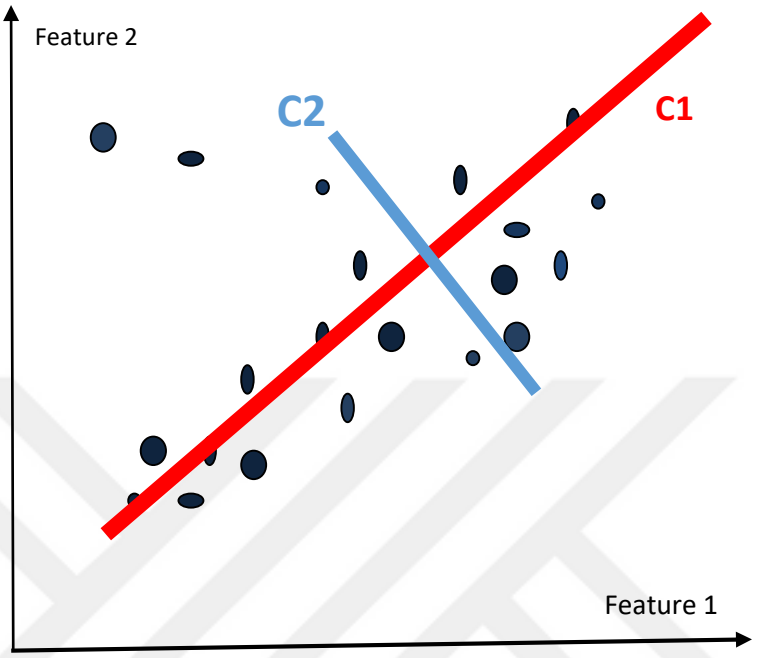

Figure 3.7 Principal Component Analysis

3.2.2.2. K-Means Clustering

The algorithm that solves the well-known clustering problem is K-means, which is the easiest to use in unsupervised learning [38]. The simple and easy method with a procedure as the following: Classify dataset by a certain number of clusters. And the main ideas are to define k-centers, one for each collection. The center should be placed in a nifty way as different locations will have different results, to get a better result. The Figure 3.8 explains the working of the k means clustering algorithm.

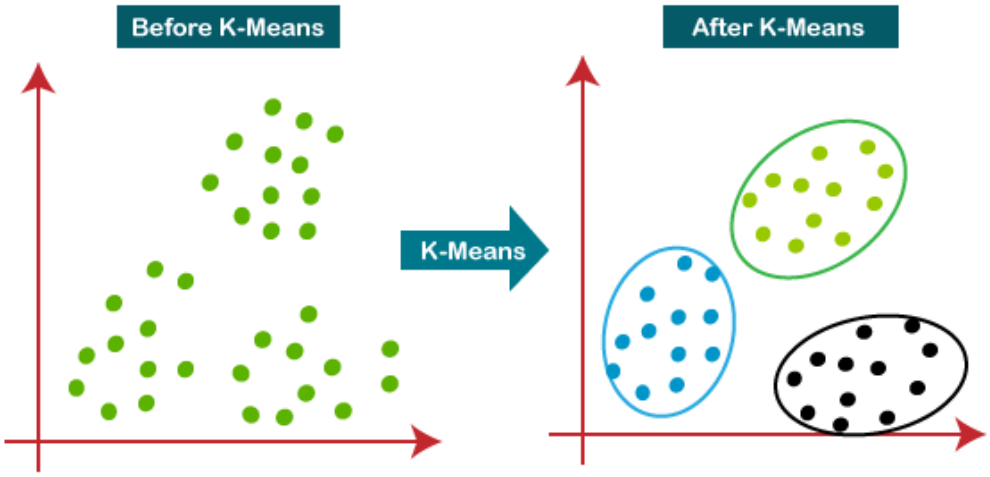

Figure 3.8 K-means clustering

The next steps are to take all the points associated with a particular data from the given dataset then assign it to the nearest center. If it doesn't make pending, first step is completed, initial group age has been completed. We need to recalculate k new centurions of gravity. As the focus of the cluster that emerged from the previous result. Table 3.3 shown k-means algorithm pseudo code.

Table 3.3 K-means algorithm pseudo code

K-means Pseudo Code

Input:
D= {t1, t2.. Tn} // Set of elements
K // Number of desired clusters
Output:
K Set of clusters
K-Means algorithm:
Assign initial values for ml, m2, ...mk
Repeat
Assign each item t1 to the clusters which has
The closest means; calculate new means for each
Cluster:
Until convergence criteria is met;

3.2.3. Semi Supervised Learning

A semi-supervised learning model is a combination of supervised and unsupervised models. These areas of machine learning can also be fruitful with data mining, where unmarked data that is already retrieving data tagged is a process. The standard methods of supervised machine learning algorithms are for the labeled datasets, and each record containing result information [39]. Some of the semi-supervised learning algorithms are described below.

3.2.3.1. Support Vector Machines (TSVMs)

Semi-supervised learning is frequently used (TSVMs) as a means of processing partially tagged data. Mystery around it makes the basics of generalization incomprehensible. The idea used to label unlabeled data to maximize the space linking labeled data to unlabeled data.

3.2.3.2. Generative Models

The process that generates data easily and equally models features and classes for data completely. The probability distribution is used to generate the data points; therefore, the entire algorithm that models P (x, y) is generative. Labeling the example for each ingredient is enough to confirm the mixture distribution.

3.2.3.3. Self-Training

The self-training becomes a classifier for labeled data. Then the data that is not marked on the classifier is supplied. The training set, unlabeled point, and predicted label are added together. Subsequently, the process is repeated further away. The name is due to the fact the classifier learns itself from Self-training.

3.2.4. Reinforcement Learning

Machine learning deals with the way software agents should perform in an environment to get ideas is reinforcement learning. Three learning steps are used to maximize the performance of this method. These are monitoring, an essential machine learning paradigm, and then unsupervised learning. Figure 3.9 shows an example of reinforcement learning such as.

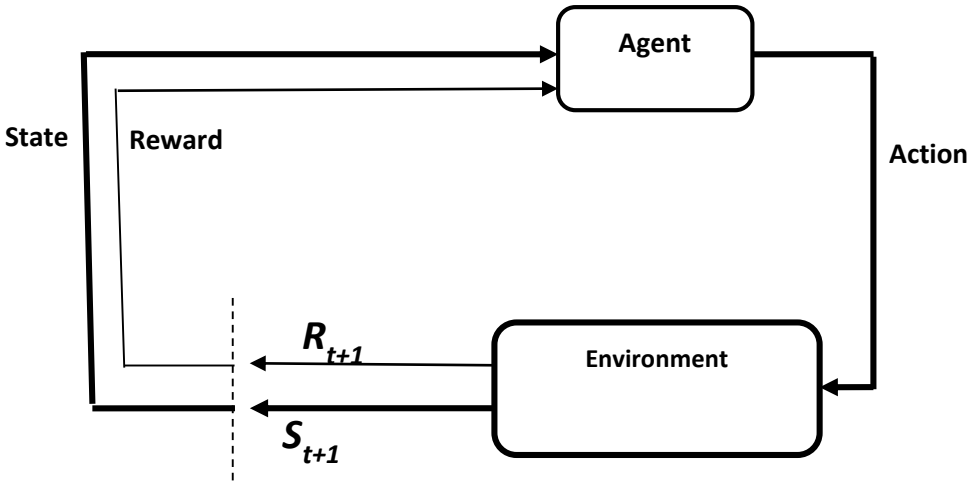

Figure 3.9 Reinforcement Learning

3.2.5. Multitask Learning

This is a part of a machine learning task that aimed to solve several tasks and take advantage of the similarities among different tasks. It adjusts and improves efficiency in learning. Formally, if you have number of tasks to execute (the conventional deep learning approach targets only one task Specific model), these number of tasks or subsets thereof are related, but not the same. Multitask learning is a specific model by the knowledge contained in all of the numbers of tasks.

3.2.6. Ensemble Learning

This is a model that helps to handle any computational intelligent problems, where multiple models, like experts or classifier, are generated and combined to solve the problem. Ensemble learning is mainly used to reduce the probability of an unfortunate selection of poor features or improve the performance of a model. Further applications of this learning comprise assigning confidence to the decision making, optimal feature selecting, data fusion, incremental learning, and so on.

3.2.7. Neural Networks (NN)

A neural network uses processes that try to understand and learned how the human brain work using some set of algorithms and to discover the fundamental relationship of datasets. In this sense, the organic or artificial neurons are the essentials of a neural network system.

Neural networks, one of the artificial intelligence techniques, have made a great contribution to the development of commercial systems. NN can settle into changing inputs on how the network produces the best possible results without redesigning the output process. Figure 3.10 shows how artificial neural networks work. It works at three levels. These layers are responsible for accepting and processing the input and computations in the output layer.

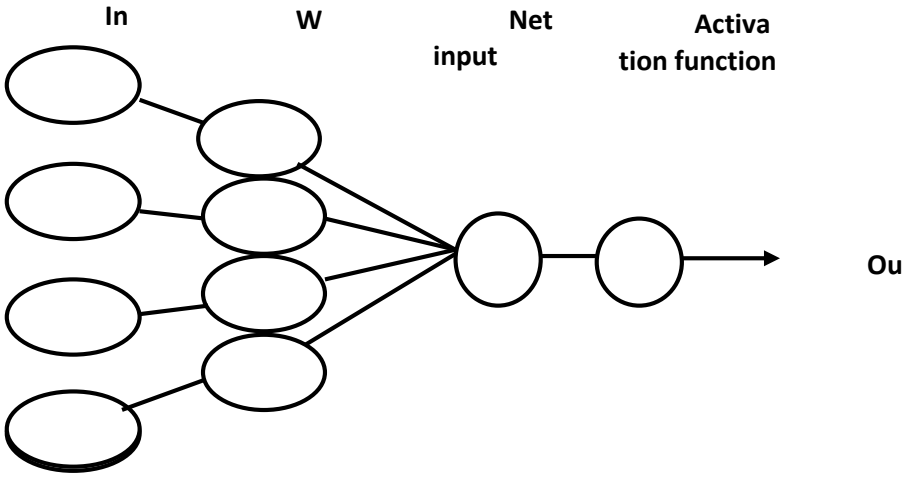

Figure 3.10 Neural networks

4. LITERATURE REVIEW

4.1. RELATED WORK

Environmental Sound Classification usually consists of manually intended feature extracts processed by traditional classifiers such as forests and k-nearest neighbors. (KNN) [10], [39], [13]. An excellent investigation work separates the methods into fixed STE, MFCC, MPEG-7, ZCR, STE, or and Gaussian mixture models (GMM) and wavelet-based, which are non-stationary approaches such as continuous wavelet transform (CWT), fast wavelet transform (FWT),

Kobayashi et al[40] Was the first author to work on a sound classification using the image descriptor LBP. The work improved the discriminating power of the LBP functional features with L2-Hellinger normalization and used linear SVMs in the sound event dataset RWCP to obtain good classification results. Their study was in 2017 by Ren et al. Tracked[41]. He applied multi-channel LBP to a gamma tone spectrogram and robot hearing and showed excellent performance on two sound event datasets, the RWCP and NTU-SEC.

With the success of deep learning in the field of image classification, many works have replaced traditional classifiers with CNNs that can better learn time-frequency features using weight sharing and clustering. However, compare HUTIFY [42]Short-Range Fourier Transform (STFT), QCT, and CQT on CNNs.

Sharma et al. [43] CNN has implemented a deep multichannel network consisting of MFCC, GFCC, CQT, and Chromagram. However, through extensive training, CNNs may be applied directly to a signal or its spectrogram without initial feature extraction.[44].

The research of SoundNet[45]They the trained network from visual recognition networks to audio networks by transferring discriminatory knowledge. Also of note is the temporal attention mechanism, Restricted Boltzmann Machine (RBM)[46]There are discussions of deep learning techniques that have been successful in ESC, like extreme deep networks and interclass techniques. Even though, the traditional classifiers were outperforming by the CNN method, because there were problems in lack of data and dataset diversity. Data enhancements are techniques generally solves, like time stretching, pitch shift, dynamic range compression, and background sound. Although CNN methods achieve good accuracy, they are much time-consuming and resource-intensive. This is because; the combination of hand-crafted features and classic machine learning methods represents a good trade-off among accuracy and speed with a reasonably small amount of data.

The human ear can distinguish different sound signals based on frequency. In addition, by giving meaning to an audio signal consisting of sound waves of a specific frequency, humans can derive the information that a specific sound means.[39]. Classification of sound is simply a method

in which an algorithm learns sound patterns and their meanings then extracts information from sound signal data, similar to the human ear[2]. Based on the sound to be detected, sound classifications techniques can be categorized accordingly. Speech recognition One of the most frequent methods that recognize natural languages in the category [15], Mental and physical condition[12], Of the human voice, music recognition is a widely used tone recognition technology that recognizes song titles and their genres, as well as the instruments played in the music. However, there are two voice detection methods, voice event detection, and acoustic scene classification, as methods for sensing the environment and situation, and they have become more popular recently than the above two methods [30].

Sound event detection is a method of detecting and classifying event types using specific sounds produced by relatively short-term events (such as broken glass beakers or falling metal objects), and acoustic scene classification is How to classify location types. Situations and environments by learning several sound events in a particular room [17].

For instance, collections of sound in a classroom, like a student's talk, computer keyboards sound, writing motion. Sounds from different locations, such as car engine sound, noise in the motor park, and that of the passengers, all can be made. Lee and Ellis [47] collected sound records from a different environment, where people acted. And classify the event using acoustic scene classification. The method of recognizing sound events is based on the assumption that different events produce different noises that can be distinguished, while the method of classifying acoustic scenes is that the sounds of different scenes, situations, and environments are different, distinguished from that of the place.

Akbal[48] has proposed novel method that came with both textural and statistical features. With the multileveled feature extraction method, he generated 9580 features. 125 most distinctive of these features were selected using NCA. The selected features were forwarded to SVM, and 90.25% accuracy was achieved during classification. ESC methods and comparatively results demonstrated that this method is successful for ESC when compared to others.

The research of Aucouturier et al. [14] Studies the differences between urban environments, called urban soundscapes, and polyphonic music. He investigated the type of environmental sound. In this system, the similarities between urban and musical sounds were proposed to determine. They studied the temporal and statistical homogeneity of each of these classes and demonstrated differences in the temporal and statistical structure in the environment.

Minkyu Lim et al. [49] have created a system that uses an audio event classifier based on Convolution Neural Networks. This system uses a feature extractor where the audio signal (as image) gets transformed into a PCM format. The Location classifier uses three different layers: convolution, pooling, and fully connected. The convolution layer identifies have occurred in the

preceding layer [50]. The pooling layer helps to reduce the dimension and combines the features that are alike into one. The fully connected layer identifies the sound of the image input.

Mendoza et al. [8] experimented with sounds classification with three different input forms. They are the Constant -Q Transform (CQT) [13], the spectrogram function, and the spectrogram image. Of these three, CQT is the most appropriate functional technique. It is based on the 3 seconds of audio given as input using probability voting, and the output delay is based on the number of windows processed by the system. The accuracy of the classification method is determined by testing dropouts [18] and batch normalization in the CNN architecture.

Additionally, Mesaros et al.[51]Worked on Event detection in real-life recordings was used based on hidden Markov models. Here HMMs were trained for 61 sound event classes using Mel-frequency cepstral coefficients MFCC) extracted from the acoustic signal mixture. The event detection used the Viterbi algorithm to decode the optimal path through the HMM state and connect 61 model HMMs to the network. Because the system was evaluated for annotations that mark duplicate events, the ability to decode only one event at a time while the annotation contains concurrent events limits recognition accuracy.

Detection of human activities based on sound is less costly than image and more usable than sensor technology. People's behavior can be determined by their voices or the sounds of the objects they use. Sound is very easy to reflect and transmit in its environment. Therefore, it can be easily measured from a distance. In addition, the signal includes the behavior of people who do not use any sensor devices.

Sound recognition systems are a method of learning and interpreting sound patterns with certain algorithms and extracting information from sound signals. Voice recognition methods are classified according to the audio signal used and the purpose. The most commonly used classes are determination of ambient sounds [52], music recognition [3], speaker recognition [38], emotional state detection [53].

In addition, there is sound event detection for detecting a certain event by modeling and acoustic scene definition for recognizing the environment created by using more than one sound together. It is aimed to detect whether there is a similar event in the environment by using an audio signal produced in sound event detection. For example, cases such as falling, breaking, hitting are classified [51]. Acoustic scene classifications aim to classify the environment, situation, or space by detecting multiple sounds in a certain environment [16]. For example, it is used to classify environments such as crime scenes, schools, restaurants, barbers, and cafes[13].

Home activities are related to both event detection and acoustic scene structure. Sounds of activities performed by people in the home may include one or more events. For this reason, the sounds produced are of different frequencies and decibels. For example, among the sounds made by a person who is cooking or cleaning the house, he can act by the event detection, and in the

behavior that fits the acoustic scene. However, indoor activities also have distinctive features from the acoustic scene or event detection. In the acoustic scene, events usually consist of more than one situation and environment. Domestic activities, on the other hand, are concerned with an environmental activity that takes place. For this reason, the classification of human activity sounds has its sound characteristic.

There are limited studies in the literature on the classification of human activities from sound signals.[54] proposed a method for the classification of human activities in his study. A dataset was created with 10 different human activity sounds collected on YouTube. The audio data in the dataset were extracted using the Log Mel filter bank. It was then classified using the residual convolution neural network. The findings showed that human activities were classified with an accuracy rate of 87.6%. [55] Proposed a method for detecting the indoor location of people from sound signals. It is aimed to determine the indoor location by recognizing the human behaviors during the activity taking place in the building at that moment. For this purpose, 11 different human activity sounds were used in 4 different rooms of the house. The feature was extracted by the statistical feature extraction method. It was then classified with 95% accuracy using the random forest model.

5. MATERIAL AND METHOD

The section explains in detail the various concepts behind the different stages of the proposed methods. The binary pattern is employed for feature generation. A total of 256 features were extracted from the signal. The model proposed here is used for feature extraction and classification. The explanations are given below.

5.1. Feature Extraction

The feature extraction stage is an important task in machine learning because the quality of the prediction depends on the quality of the feature employed. For our model, we adopted a one-dimensional local binary pattern for feature extraction.

5.1.1. LBP (Future extraction)

There are many different types of feature generators. The local binary pattern called LBP for short is one of such that is used extensively for manual feature generators in many studies. The main idea of the LBP is to compute the local representation of patterns from micro-structures to attain optimal global features such as meta-heuristic optimization methods. Local Binary Patterns are used to describe the texture and pattern of an image. it divides an image pixel into 3 × 3 sized overlapping blocks, as shown in figure 5.1 below. The local representation is obtained by comparing each pixel with its surrounding neighborhood of the 3 × 3 sized blocks [58] setting the center pixel as the threshold value.

The major aim of the local binary pattern is to find neighborhood relations using the signum function. The signum function is a primary function and it is considered as the kernel function of the LBP.

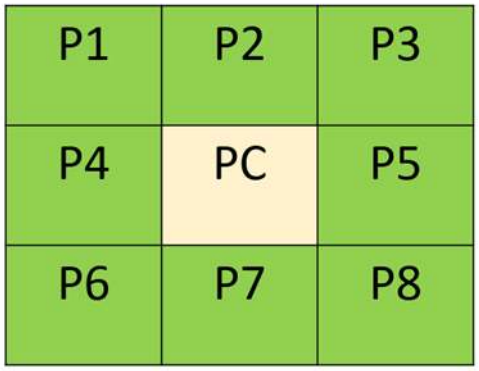

Figure 5.1. Neighborhood matrix with PC at the center pixel/value and other pixels as neighborhood pixels

Equation (5.1) Denotes the mathematical expression of the bit extraction (binary feature generation):

$$signum(t,d) = \begin{cases} 0, & t-d < 0 \\ 1, & t-d \geq 0 \end{cases} \quad (5.1)$$

The signum function in equation 5.1 above denotes a comparison function. Here, variable kernels can be used to generate binary features. The following steps explain intensely how LBP features are generated.

Step 1: Divide the image into 3 × 3 sized touching blocks. Here, W −2 × H−2 blocks are obtained from a W × H sized image where H and W are the height and width of the image.

Step 2: Obtain binary patterns from a 3 × 3 sized overlapping blocks:

$$bit(j) = signum(P_j, PC), j = \{1,2, \dots ,8\} \quad (5.2)$$

As seen from equation (5.2), eight generated bits from a block.

Step 3: We convert the obtained binary patterns to the decimal equivalent

$$map_{k,l} = \sum_{1}^{8} bit_j * 2^{j-1}, k = \{1,2, \dots , w-2\}, l = \{1,2, \dots , H-2\} \quad (5.3)$$

Step 4: Create a histogram of the map decimal value. The histogram created will have a size of $2^8 = 256$. The mathematical function for the creation of the histogram from the obtained features is given as:

$$hstgrm(map_{k,l}) = hstgrm(map_{k,l}) + 1 \quad (5.4)$$

where *hstgrm* is the histogram.

The histograms created are employed as a feature vector in the LBP.

So many studies have been presented on LBP-like feature generators because of their effectiveness. Some of the benefits are given below [56]:

- High discriminative power. Having the ability to generate informative/distinctive features.
- Computational simplicity
- Less complex mathematical structure. Hence simple to compute
- Wide area of application.

Figure 5.2**.** Shows the graphical example of the LBP

Local binary pattern LBP is two-dimensional. Its one-dimensional version is called the binary pattern

5.1.2. The Binary Pattern BP

This is a one-dimensional form of the LBP. It helps explain the benefits of the LBP on one-dimensional signals. The binary pattern operation involves dividing the signal into nine overlapping sized blocks of which the fifth value of Q3 is assigned as the center value since the LBP uses 3 × 3 sized blocks [48]. The computation of BP follows the steps below:

Step 0: Load a one-dimensional signal.

Step 1: Divide the one-D signal into a block length with 9 cells.

Step 2: select the center value of the block. Usually, the fifth value

Step 3: Compare the values with center value using the signum function to extract bits:

$$bit(j) = signum(v_j, center), j = \{1,2, \dots ,8\} \textbf{(5.5)}$$

To extract the features steps 3&4 of the LBP steps are implemented.

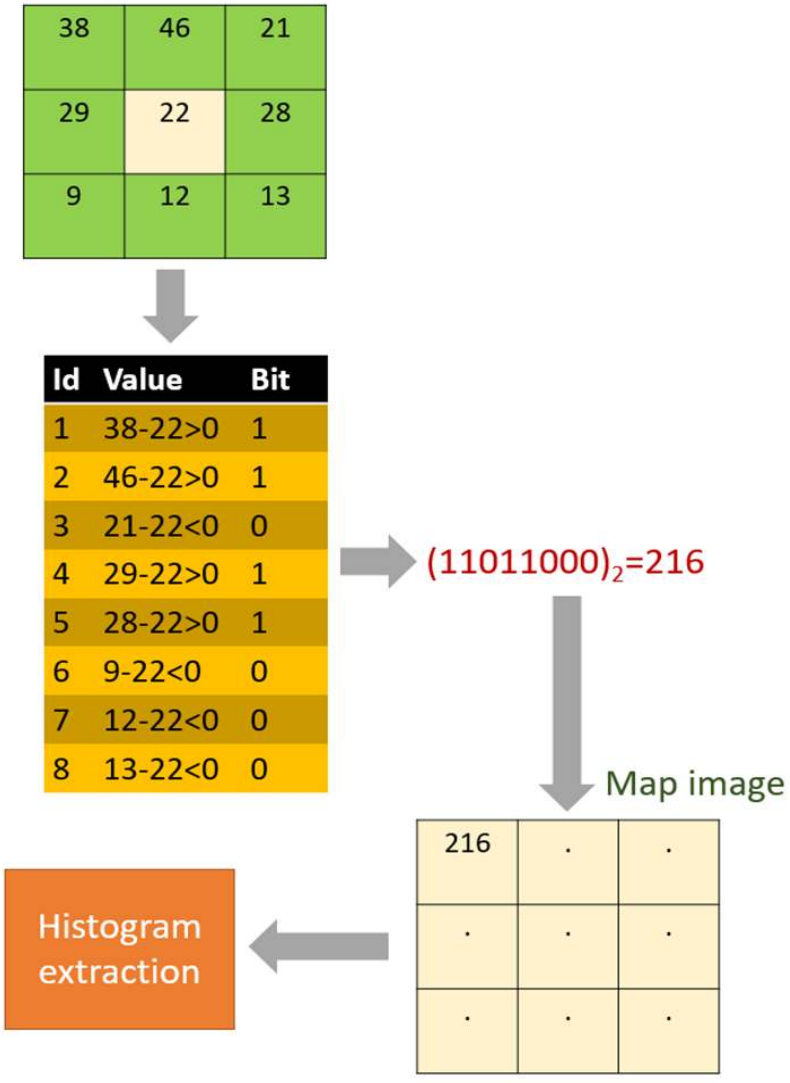

Figure 5.2. Snapshot of graphical representation of LBP

A summary of the BP is shown in Figure 5.3.

The center portion is given a yellow background and value 22. Others which neighbors are having a black background color. The center value (22) is the threshold and comparing it with the neighboring values using signum function, eight bits are generated. The bits are converted to a decimal value and a mapped signal is created. Histogram extraction is normally carried out at the final stage.

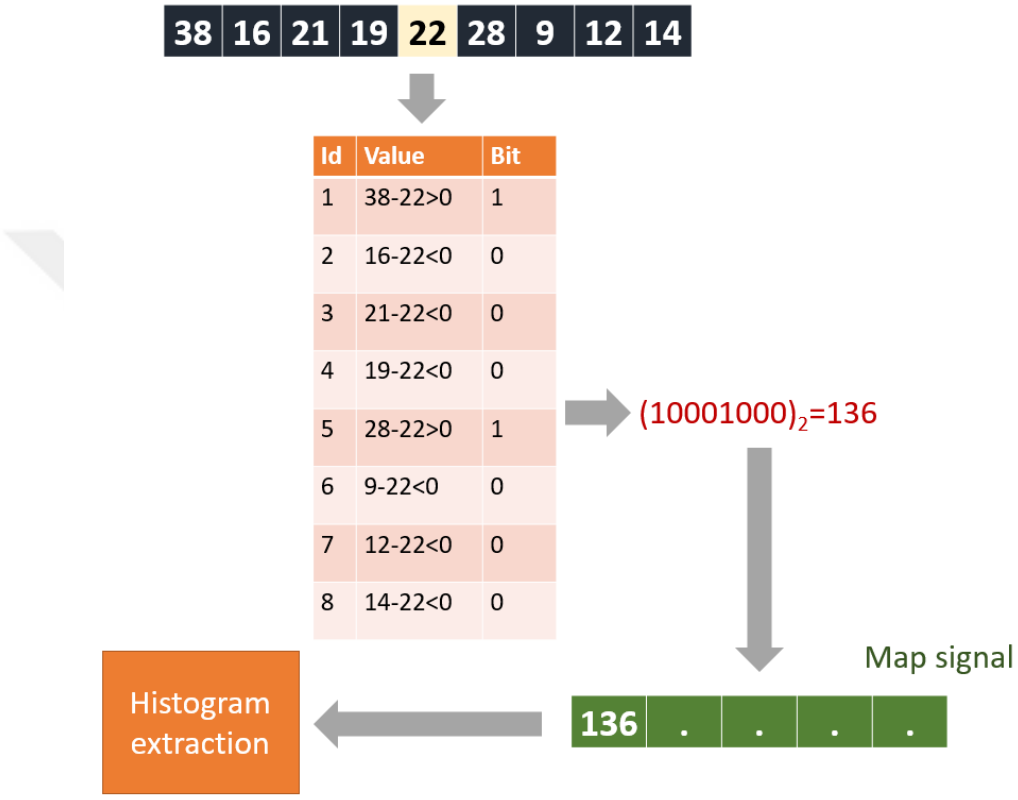

Figure 5.3. Summary of BP features extraction process

5.2. Support Vector Machine (SVM)

SVM is an optimization-based method developed by Vapnik in 1995 to solve classification problems[57]. Originally, a linear classification model was proposed. SVMs have proven effective in handling a variety of classification problems [58]. It finds the best hyper-plane of the class by finding the number of points on the edge of the class descriptor. Margin is the distance between the classes. The larger margin the more accurate the classification. Boundary data points are called support vectors.

In the future SVM used many kernels [59]. We have realized the ability to perform many classification operations through these kernels. The Frequently used are linear SVM kernels, polynomial order, radial basis function (RBF), Gauss, and Sigmoid for a prediction model [60]. Support Vector Machines are used for both regression and classification problems [61]. Has proven

to be efficient for smaller and larger datasets that cannot be handled [62]. This approach is well suited to solve problems in the form of linear and non-linear datasets [63]. It is often used in the literature [28].

5.3. Collected Dataset

A home location detection dataset was collected in this thesis. There are different user behaviors in each class of data. The data was collected mixed using a recording system and from YouTube. Sound samples to model behaviors in other videos have been determined. This dataset contains 4000 audio signals with eight classes. These classes are (1) Bathroom, (2) Bedroom, (3) Dining room, (4) Dressing room, (5) Kitchen, (6) Living room, (7) Toilet, and (8) washing room. In each class, there are 500 sounds. In this respect, a balanced dataset was collected. We obtained five hundred audio signals were from 50 different audios for each activity. Thus, we used other audio signals of different classes in the training data of the classifiers. The signal frequency is 48 kHz. The findings show that the proposed method effectively performs the classification sounds in-home activity. The information about the dataset we collected is in Table 5.1. as can be seen.

Table 5.1. Detail of collected sound dataset

Class No	**Location Name**	**Class Activity Content**	**Number of used Recording**	**Number of sample sound signal**
1	Bathroom	Showering	50	500
2	Bedroom	Sleeping, breathing, and snore sounds	50	500
3	Dining room	Eating, drinking, and other eating sounds	50	500
4	Dressing room	Folding clothes, measuring, and hand movements	50	500
5	Kitchen	Cleaning and water flushing	50	500
6	Living room	Singing, Talking, Studying, Music	50	500
7	Toilet	Flushing	50	500
8	Washing room	Washing Machine, Ironing	50	500

5.4. Proposed Method

To get results from this dataset a basic model has been presented. This model consists of two main steps and these are feature extraction using one-dimensional local binary pattern (1D-LBP) and classification using a support vector machine. The presented model is on the block diagram shown in Figure5.4 such as.

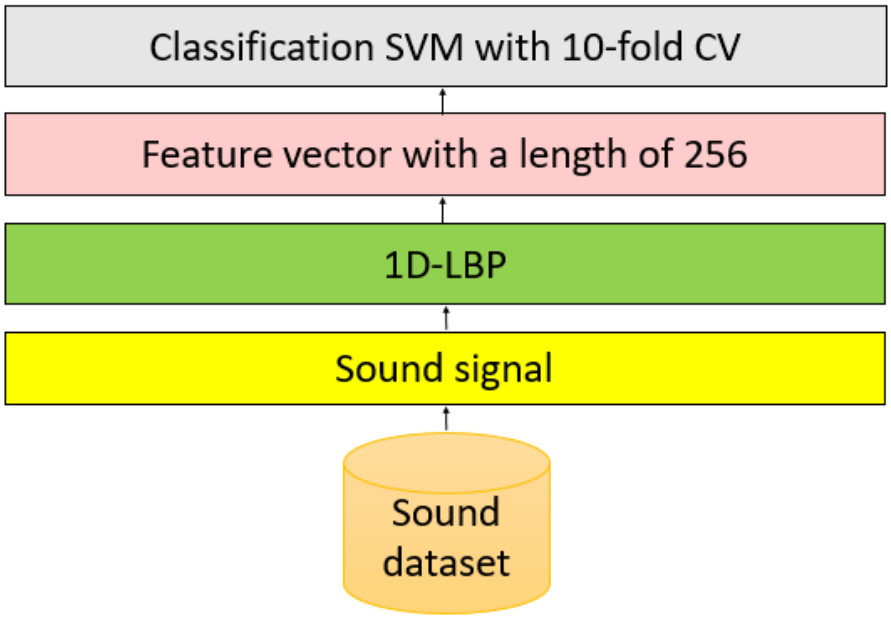

Figure 5.4. Block diagram of the used 1D-LBP based home location detection model using sound

5.4.1. Feature extraction

For this stage, the feature is generated, 1D-LBP based model for the feature was used to extract features. Steps of textural feature extraction are given below.

Moreover, the steps of the 1D-LBP-based models are given below.

Step 1: Read each sound to the collected sound dataset.

Step 2: Extract features deploying 1D-LBP.

Moreover, the definition of the 1D-LBP is in Figure 5.5 below.

Step 2.1: Divide sound signal into the overlapping block with a length of nine. They used overlapping block is denoted in Figure 5.5.

v1	v2	v3	v4	center	v5	v6	v7	v8

Figure 5.5. They used overlapping block to generate feature from a one-dimensional signal deploying 1D-LBP

Step 2.2: Extract the binary features from each block using a comparison function (signum function).

$$bit(i) = Signum(v_i, center) = \begin{cases} 0, v_i - center < 0 \\ 1, v_i - center \geq 0 \end{cases}, i \in \{1,2, \dots ,8\} \quad (5.6)$$

Where, $Signum(.,.)$ as the bit (binary feature) extraction function.

Step 2.3: Convert the generated bits to the decimal value to create a map signal.

$$map_{k,l} = \sum_{1}^{8} bit_j * 2^{j-1}, k = \{1,2, \dots , w-2\}, l = \{1,2, \dots , H-2\} \quad (5.7)$$

Step 2.4: Extract the histogram of the map signal created and use this histogram as a feature vector. As can be seen in Equation (1), the created map signal is coded using 8 bits. Therefore, the length of the generated histogram is 256 (=2^8).

Step 2.1-2.4 defines the used 1D-LBP for feature extraction.

Step 3: Classify the generated features (extracted 256 features from each sound using 1D-LBP) employing an SVM classifier. The used SVM is called Cubic SVM. The attributes of the used SVM classifier are;

Kernel function: 3rd-degree polynomial,

Kernel scale: Automatic,

Box constraint: One,

Standardize: True,

Coding: One-vs.-All,

Validation: 10-fold CV.

5.4.2. Classification

Here is the stage where classification occurs. We used SVM 3rd degree polynomial order. The SVM attributes used are as follows. A kernel was selected as Cubic (3rd-degree polynomial), we selected box constraint level (C) as one, and chose the multileveled classification method as one-vs-all and standardized as true. Test results were obtained and selected 10-fold cross-validation.

5.5. Results

The proposed method we used in our thesis is the 1D-LBP-based model for feature generation and the Cubic SVM method, tested on the collected dataset. We used 4000 audio signals with eight classes in which 500 sound signals are in each category. The worked method was implemented using a MATLAB programming environment. We used MATLAB2019 in this work.

We used SVM, MATLAB classification learner, and the calculated confusion matrix was in Table 5.1.

By deploying the proposed model, a confusion matrix is calculated and this matrix is denoted in Figure 5.6.

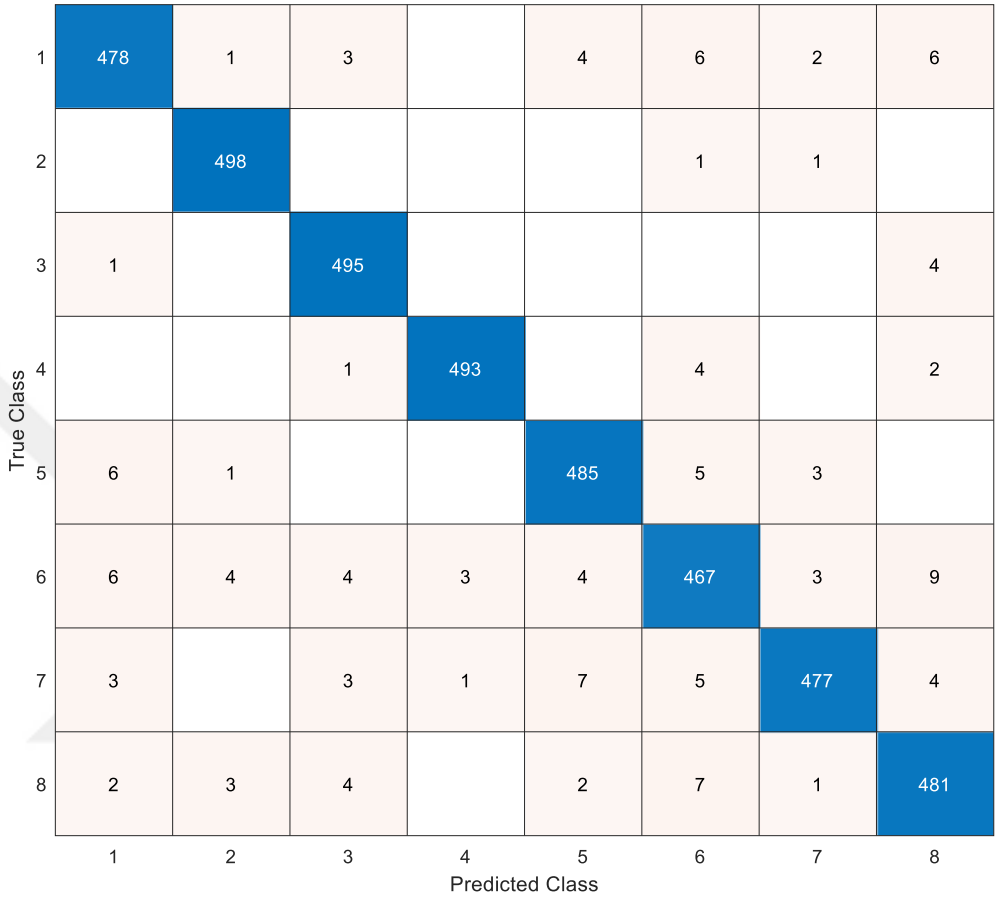

Figure 5.6. The calculated confusion matrix by using 1D-LBP and SVM based classification model

By using this confusion matrix, overall results which are accuracy, precision, recall, F1-score, geometric mean, and Mathew's correlation coefficients have been calculated and the calculated results have been tabulated in Table 5.2.

Table 5.2.Overall result

Performance metric	**Result (%)**
Accuracy	96.85
Precision	96.85
Recall	96.85
F1-score	96.85
Geometric mean	96.83
Mathew's correlation coefficients	96.40

Moreover, class-wise accuracies have been denoted in Figure 5.7.

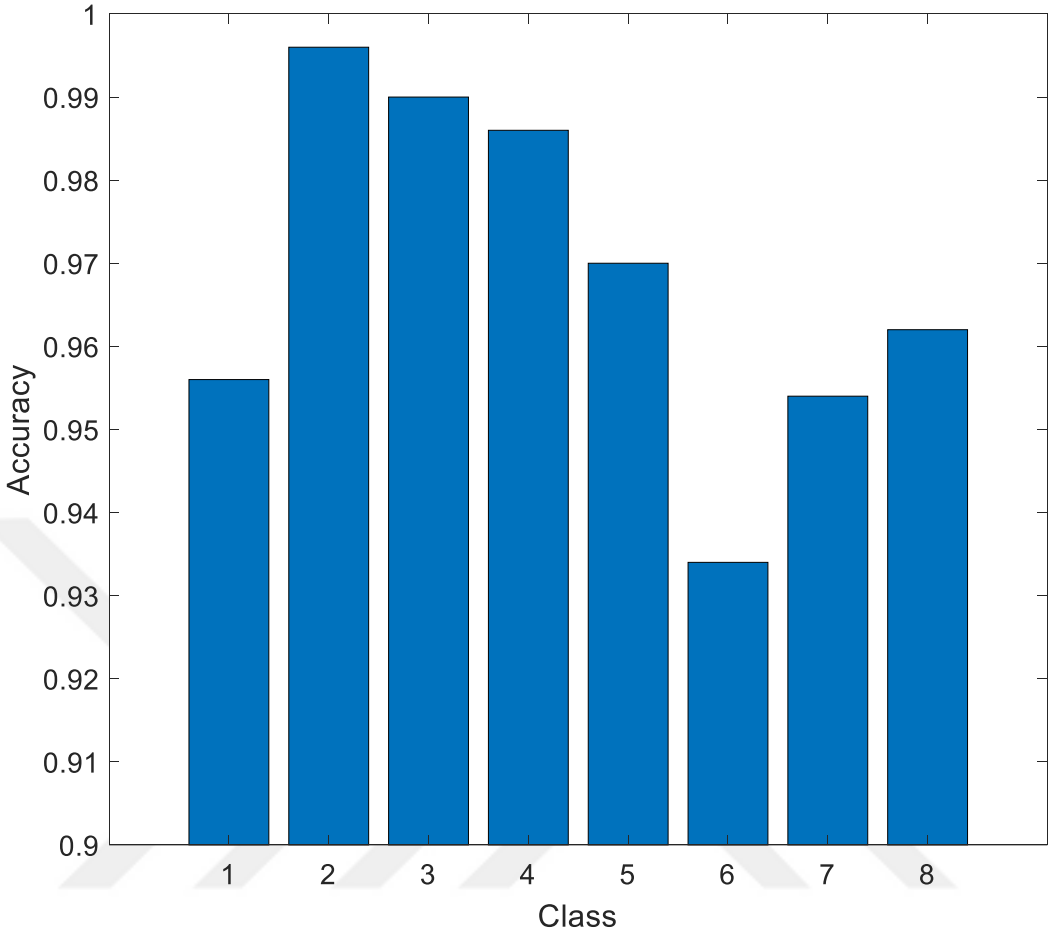

Figure 5.7. Class-wise accuracies have been denoted

As can be seen from Figure 5.7, the best-resulted class is Bedroom (2) and the worst resulting class is the living room (6).

6. CONCLUSIONS

This thesis presented an SVM called the Cubic-SVM method by using 1D-LBP feature extraction. This method consists of two stages, and these stages are feature generation and classification. We used a 1D-LBP-based model feature extraction method in the feature extraction stage, and the model extracts textural features. The feature extraction method generates 256 components in each sound. We forwarded the features to SVM, and 96.85% classification accuracy was calculated. The proposed method's success rate was also promising compared to other existing methods, and comparatively, results demonstrated that this method is successful for Sound location. The best-resulted class is Bedroom (2), and the only worst resulted class is a living room (6).

Our result shows that the technique to classify the sounds and identify the proper location, the right combination of the Machine Learning method with the sound archive system alone, produces good results. Making the machine understand the sound that has occurred in the environment will be helpful for both research and surveillance purposes. It should also be noted that training a machine to classify a particular sound proves that it is as capable as a human of predicting the environment.

REFERENCES

[1] M. Jung, S. Chi, Automation in Construction Human activity classi fi cation based on sound recognition and residual convolutional neural network, Autom. Constr. 114 (2020) 103177. https://doi.org/10.1016/j.autcon.2020.103177.

[2] W. Bian, J. Wang, B. Zhuang, J. Yang, S. Wang, J. Xiao, Audio-Based Music Classification with DenseNet and Data Augmentation, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), 2019. https://doi.org/10.1007/978-3-030-29894-4_5.

[3] K. Choi, G. Fazekas, M. Sandler, K. Cho, Transfer learning for music classification and regression tasks, in: Proc. 18th Int. Soc. Music Inf. Retr. Conf. ISMIR 2017, 2017.

[4] H. Li, S. Ishikawa, Q. Zhao, M. Ebana, H. Yamamoto, J. Huang, Robot navigation and sound based position identification, in: Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern., 2007. https://doi.org/10.1109/ICSMC.2007.4413757.

[5] M. Vacher, D. Istrate, L. Besacier, J.F. Serignat, E. Castelli, Sound detection and classification for medical telesurvey, in: Proc. IASTED Int. Conf. Biomed. Eng., 2004.

[6] L. Jing, B. Liu, J. Choi, A. Janin, J. Bernd, M.W. Mahoney, G. Friedland, A discriminative and compact audio representation for event detection, in: MM 2016 - Proc. 2016 ACM Multimed. Conf., 2016. https://doi.org/10.1145/2964284.2970377.

[7] P. Intani, T. Orachon, Crime warning system using image and sound processing, in: Int. Conf. Control. Autom. Syst., 2013. https://doi.org/10.1109/ICCAS.2013.6704220.

[8] Z. Ali, M. Talha, Innovative Method for Unsupervised Voice Activity Detection and Classification of Audio Segments, IEEE Access. (2018). https://doi.org/10.1109/ACCESS.2018.2805845.

[9] J. Ye, T. Kobayashi, X. Wang, H. Tsuda, M. Murakawa, Audio Data Mining for Anthropogenic Disaster Identification: An Automatic Taxonomy Approach, IEEE Trans. Emerg. Top. Comput. (2020). https://doi.org/10.1109/TETC.2017.2700843.

[10] M. Green, D. Murphy, Environmental sound monitoring using machine learning on mobile devices, Appl. Acoust. (2020). https://doi.org/10.1016/j.apacoust.2019.107041.

[11] H.D. Mehr, H. Polat, Human Activity Recognition in Smart Home with Deep Learning Approach, 7th Int. Istanbul Smart Grids Cities Congr. Fair, ICSG 2019 - Proc. (2019) 149–153. https://doi.org/10.1109/SGCF.2019.8782290.

[12] B.L. Giordano, S. McAdams, Sound source mechanics and musical timbre perception: Evidence from previous studies, Music Percept. 28 (2010) 155–168. https://doi.org/10.1525/mp.2010.28.2.155.

[13] O.K. Toffa, M. Mignotte, Environmental sound classification using local binary pattern and audio features collaboration, IEEE Trans. Multimed. 23 (2021) 3978–3985. https://doi.org/10.1109/TMM.2020.3035275.

[14] M. Jung, S. Chi, Human activity classification based on sound recognition and residual convolutional neural network, Autom. Constr. 114 (2020) 103177. https://doi.org/10.1016/j.autcon.2020.103177.

[15] M. Thakkar, S. Elias, A. Ashok, Speech Recognition Learning Framework for Non-Native English Accent, 2019 Int. Conf. Data Sci. Eng. ICDSE 2019. (2019) 84–89. https://doi.org/10.1109/ICDSE47409.2019.8971486.

[16] S. Rathor, R.S. Jadon, Acoustic domain classification and recognition through ensemble based multilevel classification, J. Ambient Intell. Humaniz. Comput. 10 (2018) 3617–3627. https://doi.org/10.1007/s12652-018-1087-6.

[17] D. Barchiesi, D.D. Giannoulis, D. Stowell, M.D. Plumbley, Acoustic Scene Classification: Classifying environments from the sounds they produce, IEEE Signal Process. Mag. 32 (2015) 16–34. https://doi.org/10.1109/MSP.2014.2326181.

[18] A. Khamparia, D. Gupta, N.G. Nguyen, A. Khanna, B. Pandey, P. Tiwari, Sound classification using convolutional neural network and tensor deep stacking network, IEEE Access. 7 (2019) 7717–7727. https://doi.org/10.1109/ACCESS.2018.2888882.

[19] P. Sangwan, D. Deshwal, N. Dahiya, Performance of a language identification system using hybrid features and ANN learning algorithms, Appl. Acoust. 175 (2021) 107815. https://doi.org/10.1016/j.apacoust.2020.107815.

[20] National Forensic Service, A Simplified Guide To Forensic Audio and Video Analysis, (2012) 19.

[21] R.C. Maher, Audio Forensic Examination: Authenticity, enhancement, and interpretation, Ieee Signal Process. Mag. [84]. (2009).

[22] What is Audio Forensics? Recordings used in Litigation, Audio Forensic Expert. (n.d.). https://www.audioforensicexpert.com/what-is-audio-forensics/.

[23] A. Mesaros, T. Heittola, T. Virtanen, M.D. Plumbley, Sound Event Detection: A tutorial, IEEE Signal Process. Mag. 38 (2021) 67–83. https://doi.org/10.1109/msp.2021.3090678.

[24] L. Sun, B. Zou, S. Fu, J. Chen, F. Wang, Speech emotion recognition based on DNN-decision tree SVM model, Speech Commun. 115 (2019) 29–37. https://doi.org/10.1016/j.specom.2019.10.004.

[25] J.H. McDermott, Auditory Preferences and Aesthetics: Music, Voices, and Everyday Sounds, Neurosci. Prefer. Choice. (2012) 227–256. https://doi.org/10.1016/B978-0-12-381431-9.00020-6.

[26] K.J. Devi, N.H. Singh, K. Thongam, Automatic Speaker Recognition from Speech Signals Using Self Organizing Feature Map and Hybrid Neural Network, Microprocess. Microsyst. 79 (2020) 103264. https://doi.org/10.1016/j.micpro.2020.103264.

[27] K.J. Devi, K. Thongam, Automatic speaker recognition from speech signal using bidirectional long-short-term memory recurrent neural network, Comput. Intell. (2020). https://doi.org/10.1111/coin.12278.

[28] F. Alías, J.C. Socoró, X. Sevillano, A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds, Appl. Sci. 6 (2016). https://doi.org/10.3390/app6050143.

[29] Z. Mushtaq, S.F. Su, Environmental sound classification using a regularized deep convolutional neural network with data augmentation, Appl. Acoust. 167 (2020) 107389. https://doi.org/10.1016/j.apacoust.2020.107389.

[30] T. Heittola, A. Mesaros, T. Virtanen, A. Eronen, Sound Event Detection in Multisource Environments Using Source Separation, (2011) 36–40.

[31] N. Oukrich, Daily Human Activity Recognition in Smart Home based on Feature Selection, Neural Network and Load Signature of Appliances, (2019). https://hal.archives-ouvertes.fr/tel-02193228.

[32] M.T. Jones, Artificial Intelligence: A System Approach, 2008. http://intelligence.worldofcomputing.net/category/knowledge-representation %0A.

[33] Z. Wang, Z. Liu, C. Zheng, Introduction to neural networks, Stud. Syst. Decis. Control. 34 (2016) 1–36. https://doi.org/10.1007/978-3-662-47484-6_1.

[34] G.P. Zhang, Neural networks for classification: A survey, IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. 30 (2000) 451–462. https://doi.org/10.1109/5326.897072.

[35] M. Batta, Machine Learning Algorithms - A Review , Int. J. Sci. Res. (IJ. 9 (2020) 381-undefined. https://doi.org/10.21275/ART20203995.

[36] S.M. Tahsien, H. Karimipour, P. Spachos, Machine learning based solutions for security of Internet of Things (IoT): A survey, J. Netw. Comput. Appl. 161 (2020). https://doi.org/10.1016/j.jnca.2020.102630.

[37] J. Vávra, M. Hromada, L. Lukáš, J. Dworzecki, Adaptive anomaly detection system based on machine learning algorithms in an industrial control environment, Int. J. Crit. Infrastruct. Prot. 34 (2021). https://doi.org/10.1016/j.ijcip.2021.100446.

[38] A. N. Jadhav, N. V. Dharwadkar, A Speaker Recognition System Using Gaussian Mixture Model, EM Algorithm and K-Means Clustering, Int. J. Mod. Educ. Comput. Sci. 10 (2018) 19–28. https://doi.org/10.5815/ijmecs.2018.11.03.

[39] D.S. Veena, N. M. V, R. J. V, S.T. S, Sound Classification System Using Machine Learning Techniques, Int. J. Eng. Appl. Sci. Technol. 5 (2020) 674–678. https://doi.org/10.33564/ijeast.2020.v05i01.120.

[40] T. Kobayashi, J. Ye, Acoustic feature extraction by statistics based local binary pattern for environmental sound classification, ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc. (2014) 3052–3056. https://doi.org/10.1109/ICASSP.2014.6854161.

[41] J. Ren, X. Jiang, J. Yuan, N. Magnenat-Thalmann, Sound-Event Classification Using Robust Texture Features for Robot Hearing, IEEE Trans. Multimed. 19 (2017) 447–458. https://doi.org/10.1109/TMM.2016.2618218.

[42] M. Huzaifah, Comparison of Time-Frequency Representations for Environmental Sound Classification using Convolutional Neural Networks, (2017) 1–5. http://arxiv.org/abs/1706.07156.

[43] J. Sharma, Advances in Deep Learning Towards Fire Emergency Application: Novel Architectures, Techniques and Applications of Neural Networks, n.d.

[44] N.K. Verma, A. Salour, Feature extraction, Stud. Syst. Decis. Control. 256 (2020) 121–173. https://doi.org/10.1007/978-981-15-0512-6_4.

[45] Y. Aytar, C. Vondrick, A. Torralba, SoundNet: Learning sound representations from unlabeled video, Adv. Neural Inf. Process. Syst. (2016) 892–900.

[46] N. Ganapathy, R. Swaminathan, T.M. Deserno, Deep Learning on 1-D Biosignals: a Taxonomy-based Survey, Yearb. Med. Inform. 27 (2018) 98–109. https://doi.org/10.1055/s-0038-1667083.

[47] D. Stowell, D. Giannoulis, E. Benetos, M. Lagrange, M.D. Plumbley, Detection and Classification of Acoustic Scenes and Events, IEEE Trans. Multimed. 17 (2015) 1733–1746. https://doi.org/10.1109/TMM.2015.2428998.

[48] E. Akbal, An automated environmental sound classification methods based on statistical and textural feature, Appl. Acoust. 167 (2020) 107413. https://doi.org/10.1016/j.apacoust.2020.107413.

[49] M. Lim, D. Lee, H. Park, Y. Kang, J. Oh, J.S. Park, G.J. Jang, J.H. Kim, Convolutional neural network based audio event classification, KSII Trans. Internet Inf. Syst. 12 (2018) 2748–2760. https://doi.org/10.3837/tiis.2018.06.017.

[50] H. Nam, B. Han, Learning Multi-domain Convolutional Neural Networks for Visual Tracking, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2016-Decem (2016) 4293–4302. https://doi.org/10.1109/CVPR.2016.465.

[51] A. Mesaros, T. Heittola, O. Dikmen, T. Virtanen, Sound event detection in real life recordings using coupled matrix factorization of spectral representations and class activity annotations, ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc. 2015-Augus (2015) 151–155. https://doi.org/10.1109/ICASSP.2015.7177950.

[52] P.K. Essandoh, F.A. Armah, E.K.A. Afrifa, A.N.M. Pappoe, Determination of Ambient Noise Levels and Perception of Residents in Halls at the University of Cape Coast, Ghana, Environ. Nat. Resour. Res. 1 (2011). https://doi.org/10.5539/enrr.v1n1p181.

[53] H. Aouani, Y. Ben Ayed, Emotion recognition in speech using MFCC with SVM, DSVM and auto-encoder, 2018 4th Int. Conf. Adv. Technol. Signal Image Process. ATSIP 2018. (2018) 1–5. https://doi.org/10.1109/ATSIP.2018.8364518.

[54] M.S. Seyfioğlu, A.M. Özbayoğlu, S.Z. Gürbüz, Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities, IEEE Trans. Aerosp. Electron. Syst. 54 (2018) 1709–1723. https://doi.org/10.1109/TAES.2018.2799758.

[55] L. Calderoni, M. Ferrara, A. Franco, D. Maio, Indoor localization in a hospital environment using Random Forest classifiers, Expert Syst. Appl. 42 (2015) 125–134. https://doi.org/10.1016/j.eswa.2014.07.042.

[56] T. Tuncer, S. Dogan, F. Ertam, A. Subasi, A dynamic center and multi threshold point based stable feature extraction network for driver fatigue detection utilizing EEG signals, Cogn. Neurodyn. 15 (2021) 223–237. https://doi.org/10.1007/s11571-020-09601-w.

[57] X. Huang, A. Maier, J. Hornegger, J.A.K. Suykens, Indefinite kernels in least squares support vector machines and principal component analysis, Appl. Comput. Harmon. Anal. 43 (2017) 162–172. https://doi.org/10.1016/j.acha.2016.09.001.

[58] S. Chen, M. Peng, H. Xiong, X. Yu, SVM Intrusion Detection Model Based on Compressed Sampling, J. Electr. Comput. Eng. 2016 (2016). https://doi.org/10.1155/2016/3095971.

[59] S. Amari, S. Wu, Improving support vector machine classifiers by modifying kernel functions, Neural Networks. 12 (1999) 783–789. https://doi.org/10.1016/S0893-6080(99)00032-5.

[60] K.A.A. Abakar, C. Yu, Performance of SVM based on PUK kernel in comparison to SVM based on RBF kernel in prediction of yarn tenacity, Indian J. Fibre Text. Res. 39 (2014) 55–59.

[61] R.G. Brereton, G.R. Lloyd, Support Vector Machines for classification and regression, Analyst. 135 (2010) 230–267. https://doi.org/10.1039/b918972f.

[62] H. Yu, J. Yang, J. Han, X. Li, Making SVMs scalable to large data sets using hierarchical cluster indexing, Data Min. Knowl. Discov. 11 (2005) 295–321. https://doi.org/10.1007/s10618-005-0005-7.

[63] S. Ghosh, A. Dasgupta, A. Swetapadma, A study on support vector machine based linear and non-linear pattern classification, Proc. Int. Conf. Intell. Sustain. Syst. ICISS 2019. (2019) 24–28. https://doi.org/10.1109/ISS1.2019.8908018.

CURRICULUM VITAE

Nura ABDULLAHI

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

PUBLICATIONS

1. International Research Journal of Innovative in Engineering and Technology.

"Using formative and summative assessment in data mining to predict student's final grades". 2020.

2. S. Aliyu, N. Abdullahi:"SQIX: Answering Queries over XML."(IEEE NIGERCON-2017).Federal University of Technology, Owerri, Nigeria.