ANKARA YILDIRIM BEYAZIT UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

ENVIRONMENTAL SOUND DATABASE DESIGN AND IMPLEMENTATION

M.Sc. Thesis by

Abdussamet Tanış

Department of Electronics and Communication Engineering

February, 2018

ANKARA

ENVIRONMENTAL SOUND DATABASE DESIGN AND IMPLEMENTATION

A Thesis Submitted to

The Graduate School of Natural and Applied Sciences
of Ankara Yıldırım Beyazıt University

In Partial Fulfillment of the Requirements for the Degree of Master of Science in Electronics and Communication Engineering, Department of Electronics and Communication Engineering

by
Abdussamet Tanış

February, 2018

ANKARA

M.Sc. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"ENVIRONMENTAL SOUND DATABASE DESIGN AND IMPLEMENTATION"** completed by **ABDUSSAMET TANIŞ** under the supervision of **PROF. DR. HÜSEYIN CANBOLAT** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Dr. Hüseyin Canbolat

Supervisor

Asst. Prof. Dr. Baran Uslu

Jury Member

Prof. Dr. Haldun Göktaş

Jury Member

Prof. Dr. Fatih V. ÇELEBİ

Director

Graduate School of Natural and Applied Sciences

ETHICAL DECLARATION

I hereby declare that, in this thesis which has been prepared in accordance with the Thesis Writing Manual of Graduate School of Natural and Applied Sciences,

- All data, information and documents are obtained in the framework of academic and ethical rules,
- All information, documents and assessments are presented in accordance with scientific ethics and morals,
- All the materials that have been utilized are fully cited and referenced,
- No change has been made on the utilized materials,
- All the works presented are original,

and in any contrary case of above statements, I accept to renounce all my legal rights.

Date:
Signature:
Name & Surname: Abdussamet Tanış

ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my supervisor, Prof. Dr. Hüseyin CANBOLAT for his tremendous support and motivation during my study. I also would like to thank Yüksel ARSLAN. His knowledge and precious recommendations constituted the milestones of this study. His guidance assisted me all the time of my research and while writing this.

2018, 19 February **Abdussamet Tanış**

ENVIRONMENTAL SOUND DATABASE DESIGN AND IMPLEMENTATION

ABSTRACT

This thesis proposes “An Environmental Sound Database Design and Implementation” for environmental sound recognition. Environmental sound recognition has become a hot topic in recent years. Environmental sound recognition can be used in security systems, in robot navigation, in internet search engines and more. It is necessary to establish a database which comprises the sounds to be recognized in advance to be able to recognize the environmental sounds. Sound database development is the first stage of all environmental sound recognition tasks. In this study, a database based relational database system is designed.. The sound databases designed and used so far are all constructed on the directory and file system of the operating system. For the first time, a sound database system was constructed using relational database and SQL (Structured Query Language) was used. After database design, the example sounds was loaded, a GUI (Graphical User Interface) as a desktop application and a web interface for accessing over internet was designed.

Keywords: Database design, environmental sound recognition, graphical user interface, relational database.

ÇEVRESEL SES VERİTABANI TASARIMI VE UYGULAMASI

ÖZET

Bu tezde, çevresel ses tanıma amaçlı "Bir Çevresel Ses Veritabanı Tasarımı ve Uygulaması" önerilmektedir. Çevresel ses tanıma son yıllarda öne çıkan bir konu haline geldi ve güvenlik sistemlerinde, robot navigasyonunda, internet arama motorlarında ve daha pek çok alanda kullanılabilmektedir. Çevresel sesleri tanımak için önceden tanınması gereken sesleri içeren bir veritabanı oluşturmak gereklidir. Ses veritabanı geliştirme, tüm çevresel ses tanıma görevlerinin ilk aşamasıdır. Bu çalışmada ilişkisel veritabanı sistemine dayalı bir veritabanı tasarlanmıştır. Şimdiye kadar tasarlanmış ve kullanılan ses veritabanları, işletim sisteminin dizin ve dosya sisteminde oluşturulmuştur. İlişkisel veri tabanı kullanılarak ilk kez bir ses veritabanı sistemi oluşturulmuştur ve bunda SQL'den (Yapılandırılmış Sorgu Dili) yararlanılmıştır. Veritabanı tasarımından sonra veritabanına, örnek sesler yüklenilmiş, masaüstü uygulaması olarak bir GUI (Grafik Kullanıcı Arayüzü) ve internet üzerinden erişmek için bir web arayüzü tasarlanmıştır.

Anahtar Sözcükler: Veritabanı tasarımı, çevresel ses tanıma, grafik kullanıcı arayüzü, ilişkisel veritabanı.

CONTENTS

NOMENCLATURE

AESR	Automatic Environmental Sound Recognition
ASR	Automatic Speech Recognition
CSV	Comma Seperated Values
DB	Database
DCASE	Detection and Classification of Acoustic Scenes and Events
ESC	Environmental Sound Classification
ESR	Environmental Sound Recognition
GUI	Graphical User Interface
HTML	Hypertext Markup Language
IEEE	The Institute of Electrical and Electronics Engineers
kHz	Kilo Hertz
MFCC	Mel Frequency Cepstral Coefficients
ML	Machine Learning
RWCP	Real World Computing Partnership
SNR	Signal to Noise Ratio
SQL	Structured Query Language

LIST OF FIGURES

Chapter 1

INTRODUCTION

Environmental Sound defined as all naturally occurring sounds except for speech and music [1]. When the reader is reading this sentence, the reader may hear a heater, a creaking voice, a computer humming, or a discussion of his neighbor. These are everyday sounds, and listening to them and commenting on why they come to the scene is a daily listening: a simple perception ability that can be done without having to use our technical abilities [2]. But it's not that simple for a recognition system.

Almost all research on auditory information was carried out by focusing on individual studies on acoustic signal processing, hearing processing and speech communication. However, the most important point to understand the sound scene is the close cooperation and integration of these functions. In order to recognize a specific sound, the system should localize the target sound among a large number of surrounding sound mixes and focus the sound. The collection of the audio scene data in real sound environments is indispensable for the investigation of the audio scene. The sound stage database contributes to the study of the sound stage understanding. [3].

Environmental Sound Recognition (ESR) is mainly based on machine learning (ML) and ML algorithms require first a training database. This database must comprise the sounds to be recognized and other related sounds. An ESR system needs the database during training, testing and in the production stage[4].

The applications of ESR can be listed as follows: In military, forensic and law enforcement domain there are studies on gunshot detection systems. In[5], a gunshot detection system is proposed.The system used a database amounted to 434 audio clips consisting of 332 gunshot and 102 outsider signals. All audio signals were sampled at 44.1 kHz with 8-bit quantization. In[6], the gunshot blast is used to identify the caliber of the gun. In [7] and [8] ESR is used for robot navigation. ESR can be used for home monitoring. It can be used to assist elderly people living in their home alone [9],[10]. In[11], it is used for home automation. . In [12] and [13], ESR is used for recognition of animal sounds.In [13], system used a large database with 1418 animal vocalizations annotated with the species and additional metadata, including recording conditions and type of vocalization for each audio file. In the surveillance area, it is used for surveillance of road [14], public transport [15], elevator [16] and office corridor [17].

There are a lot of papers in literature on environmental sound recognition[18][19][20][21]. Most of these papers mention about lock of a common environmental sound database and the difficulty of building a database. Although these articles mention the database or libraries that contain sound types used, there is no information about their details. Only the technical specifications of the voices used are given. During the development of an environmental speech recognition system, many sound clips, many of these sounds, and many algorithms that use these attributes will need to be studied. Once the system is operational and ready to be developed, new sound types and their new sound clips, new attributes and algorithms that should be recognized will need to be added. Establishing a database structure that will facilitate these studies will be very useful[22].

The objective of this thesis is design and develop novel Environmental Sound Database example for recognition systems' needs. . In this thesis, we present the design and pilot establishment of a database which will assists all researchers who want to establish an ESR system. This database employs relational database model which is not used for this task before. We explain in this study design and implementation details of the database, data collection and load process. Besides we explain the tools and developed graphical user interface for a desktop application and for the web. In addition a paper[4] about this topic was published in ASTESJ(Advances in Science, Technology and Engineering Systems Journal).

1.1 Literature Survey

There are some researches on building such databases. One of them http://www.auditorylab.org is developed by Carnegie Mellon University. It is main purpose to test the ability of human to understand events in the environment from sounds. http://www.desra.org is a general purpose environmental sound database. Researchers can access the database from the internet. http://www.daresounds.org is a database for auditory perception research.

Another one is http://www.desra.org . In [23] explains the purpose of this database, the design details and the scientific resources used. The use of experiential voice recognition ability and the technical details of the voices during database creation constitute the most important feature of the database. It has an interface that can be accessed bu users over the internet.

The database published at http://www.daresounds.org/ is described in [2]. This database is a database where the voices we hear in the complexity of everyday life are recorded with explanations and are accessible to the scientific world. Database is intended to provide a verification environment for auditory perception studies and environmental voice recognition systems.

A report describing an environmental sound database developed for a survey conducted at the US Land Forces Research Laboratory [24]. It was decided which features of the voices in the environmental sound database to be prepared should be documented, selection criterion for the voices to be selected in the database, and registration standards for the voices to be newly recorded.

One of the databases that have been mostly reverenced is Real World Computing Partnership's (RWCP) non-speech database [3]. The objective of the project is to provide common standard databases for research concerning real acoustic environments.

Detection and Classification of Acoustic Scenes and Events (DCASE) is an official IEEE Audio and Acoustic Signal Processing challenge. For challenge a database is prepared and it is also a resource for ESR researchers.

DCASE 2016 and DCASE 2017 challenges consists of four tasks [25] [26]. The goal of Task-2 is to detect sound events for example "bird singing", "car passing by" that are present within an audio segment.

Dataset for Environmental Sound Classification (ESC)[27] has tried to address scarcity of suitable and publicly available datasets issue by presenting a new annotated collection of 2 000 short clips comprising 50 classes of various common sound events, and an abundant unied compilation of 250000 unlabeled auditory excerpts extracted from recordings available through the Freesound project.

In [28], a dataset of annotated urban sound clips are recorded and taxonomy for urban sound sources is proposed. The dataset contains 10 sound classes with 18.5 hours of annotated sound event occurrences. The objective of the project is bring a solution to the lack of a common taxonomy and the scarceness of large, real-world, annotated data.

1.2 Motivation

The purpose of defining a computational sound event is to design a system that can achieve human-like performance in various hearing tasks [29]. Lyon refers to this as "Machine Hearing" [30] and explains what we can expect from a computer that people hear:

> “If we had machines that could hear as humans do, we would expect them to be able to easily distinguish speech from music and background noises, to pull out the speech and music parts for special treatment, to know what direction sounds are coming from, to learn which noises are typical and which are noteworthy. Hearing machines should be able to organise what they hear; learn names for recognisable objects, actions, events, places, musical styles, instruments, and speakers; and retrieve sounds by reference to those names."

As Lyon mentioned above, main task of a sound recognition system is organise what they hear. The first step to accomplish this is to design a database with correct, documented, organized references. It will be possible to recognize and organize the voices heard in the next steps of sound recognition systems.

Historically non-speech sound recognition has not received as much attention as automatic speech recognition (ASR). ASR has well established algorithms and databases while research has begun much earlier. Automatic ESR (AESR) is getting attraction since last two decades and environmental sound research is still in its fancy, as compared with the relatively well-established fields of speech,music and psychoacoustic research. For this reason, a well-established common standardized database for AESR studies still does not exist. This provides the motivation to develop novel methods to design database.

1.3 Organisation of Thesis

This thesis is organised as follows:

In Chapter 2, a thorough overview of the field of environmental sound and environmental sound recognition is provided.

Chapter 3 then introduces database requirements and design, followed by explain implementation of it. After that shows an example of load data to database.

In Chapter 4, shows development and use of an Graphical User Interface(GUI) that manages the implemented database. Describes tools and usage of GUI.

Chapter 5 then shows implement and use of implemented database’s web application.

Chapter 2

ENVIRONMENTAL SOUNDS

2.1 Overview

Environmental Sound has different definitions. Everyday listening[31],sound event[29],everyday sounds[32],everyday life sound [33] a few of them. In this thesis we use Environmental Sound usage.

The interpretation of the concept of environmental sound are various in the literature. As in [20], the wind in the trees can be caused by the sounds of nature in the nature, such as the sound of rain or the chirping of insects. In [1], Environmental Sound defined as all naturally occurring sounds other than speech and music. For instance dog barking ,car crashing, car horn , air conditioner humming etc. Environmental sound beyond the limits of theoretical and applied research in different areas is important. As a sound class, environmental sounds are worth examining because of their ability to transmit semantic information based on a complex acoustic structure. However, unlike speech and music of other large meaningful voices, in other words, the information conveyed by environmental sounds is not linguistic, as in speech, and is typically not designed for aesthetic value as in music. There are many practical and specific applications for environmental sounds; hearing rehabilitation for hearing aid users and buyers of cochlear implants, assessment of auditory and cognitive deficits in pre-language children, non-verbal diagnostic tools; sound control and design of acoustic environments; auditory symbols on computer screens; sonification, process of representing information by unknown sounds (see [34] for a recent review). However, knowledge base is far behind the naturally occurring daily voices, speaking and music classes. One of the main hurdles is that there is no standard database of free, tested, high quality, identifiable peripheral sounds ready for users to work with. [23].

2.2 Environmental Sound Recognition Overview

While research on sound recognition has traditionally focused on speech and music signals, the issue of environmental sound recognition (ESR) has received more attention in last years. ESR research has increased significantly in the last decade. [35].

Over the last decade, a significant amount of research has been done to model and identify environmental sounds. With environmental sounds, we have refer (that is, everyday sounds, except speech and music) to various voices that are natural and artificial. Environmental sound recognition (ESR) plays an important role in the near-term efforts to perfect machine audition [35]. With increasing sample-based search requests, such as content-based image and video search, ESR has begun to be used as an instrument in efficient voice search applications. [36]. ESR can be used for automatic tagging of audio files with identifiers for keyword-based audio reception[37]. By including ESR in the system, robot navigation can be advanced. [7], [8]. ESR can be used in the home monitoring environment to help elderly people living alone in their own home [9], [10] or for a smart home [11]. ESR, along with image and video analysis, find applications in surveillance [38], [39]. ESR can also be adapted for recognition by distinctive sounds of animal and bird species [12], [13].

Speech and music among the various types of sound signals are two categories that are examined extensively. The ESR algorithms were only a reflection of the speech and music recognition paradigms in the first period. However, due to the non-stationary nature of ambient sounds, these algorithms are insufficient for large-scale databases. [35].

The purpose of ESR systems is to recognize a sound event that results in an individual audio source in a continuous audio signal. The main concern is to detect the event in a given extract of the recording. The precise location of the event, i.e. the start and end times of the event, has a secondary prefix, so detection systems often have a rough temporal resolution. This is where ESRs differ from other audio-based recognition systems such as speech and music recognition. Previous ESR studies were typically performed for isolated sound events [40]. Each record has a single sound event, and the ESR is trained to detect the event at a specified time interval. ESR is performed on records collected from real-life environments. These environments have different acoustical properties due to the presence of human or nature activities. There are multiple overlapping audio events, so the detection of individual events is more challenging than the isolated recordings. An ESR system capable of detecting multiple events in the same situation is called a polyphonic ESR system. Figure 2.1 shows the polyphonic nature of sound events in real-scene situations. ESR systems first form an intermediate representation of the audio waveform in a record. This process is called feature extraction. Since irrelevant features make it difficult to detect because it can lead to loss of valuable information, it is very important to choose the right features to extract from an audio signal. There are strong signs of time in many cases of nature, such as wind and rain. For this reason, it is somewhat necessary to extract temporal domain knowledge and spectral domain knowledge in the environmental sound recordings. [41].

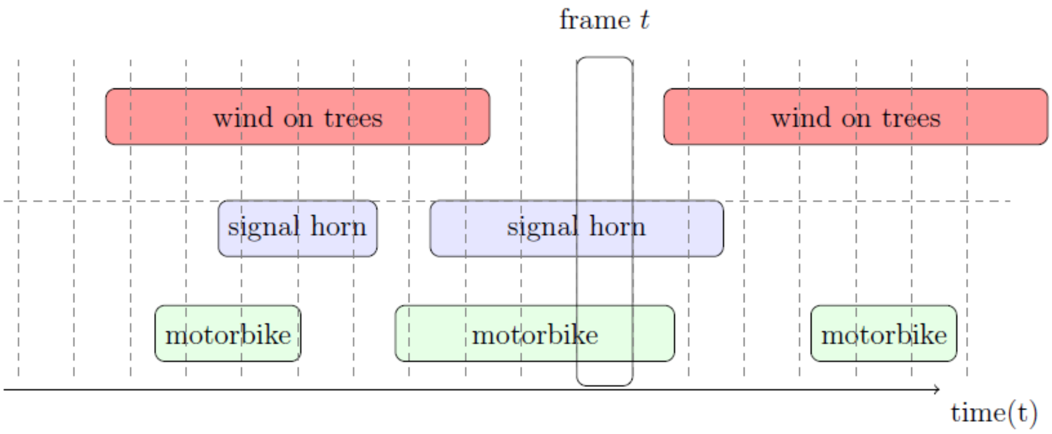


Figure 2.1: Overlapping sound events in a recording from a realistic environment. Frame t represents the short time frame from the recording where only signal horn and motorbike events are present.

2.2.1 General Challenges of Environmental Audio Classification and Recognition

In this section, main challenges that are faced in ESR and classification are pointed out. Unlike speech or music signals, environmental acoustic signals are difficult to model due to its high unpredictable nature. Speech or music can be categorized to structured sounds due to their formantic or harmonic structure characteristic whereas environmental sounds, on the other side, are typically unstructured, which have a broad noise-like flat spectrum and diverse variety of signal composition and are difficult to build models. Analysis of real world audio that consists of a rich mix of naturally occurring sounds such as the environmental sound is complex. As a result, classification and processing of environmental sound is generally more cumbersome compared with that of speech or music. The following are the general challenges that are faced during the design and implementation of ESR technique[42]:

- "**Overlap in time and/or frequency content** - Different sound events can happen at the same time which makes recognition of the type of sound event difficult. This leads to two challenging tasks: detection of individual sound events within the audio scene (segmentation) and classification. A system involved in the first task has as a goal to cluster mixed sound events from different sources into their corresponding source type, or try to segment the audio into pieces that represent a single occurrence of a specific event class by estimating the start and end time of each event and if necessary separating it from other overlapping events. The aim of the second task is to characterize and identify the type of sound event (e.g., label the environment in which the audio was recorded). Thus, the overlapping sound events that constitute a natural auditory scene (environmental sound) create an acoustic mixture signal that is more difficult to handle."

- "**Dynamic nature of environment** - Apart from containing a wide variety of sound classes, environmental sound has a dynamic nature, i.e., new sound types (classes) can appear and existing sound classes can disappear randomly at any time. Therefore, the ESR technique has to deal with and adapt to the dynamic nature of an environmental sound."

- "**Selection of feature set** - Audio features have a significant impact on the recognition accuracy. Thus, the definition and extraction of the right type of feature sets is a very important step in ESR. However, it is challenging step too. What feature types we define and how we use them depends on the type of application. For example, audio features used for audio classification in indoors might not perform well when used for the classification of types of sounds in outdoor areas."

2.3 Relationship Between Environmental Sound Recognition System Design and Database

In our published paper[4], we mentioned about this subject:

"Environmental Sound Recognition system design is started with the training phase. How a sound database is used in training phase of an ESR system is explained in Figure 2.2.

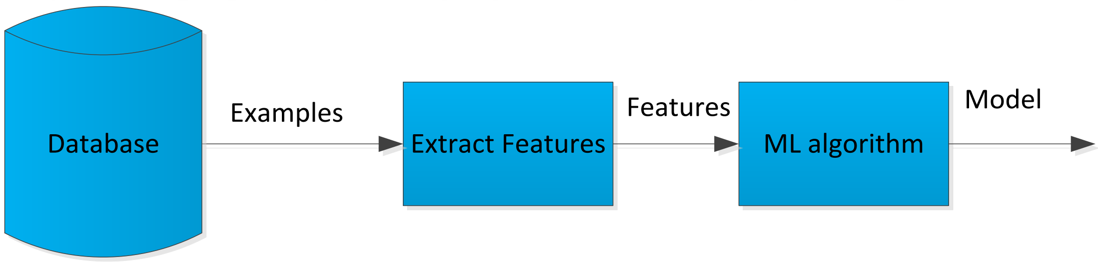


Figure 2.2 ESR system in training phase

During the development stage of an ESR system the desired ML algorithm must be trained with the sounds to be recognized. The database provides the sound clips to be recognized and other sound clips which are negative examples. After training, a model is developed and this model is used for testing. In Figure 2.2, the database provides positive and negative examples; features are inputs to the ML algorithm, ML algorithm using these features produce the model.

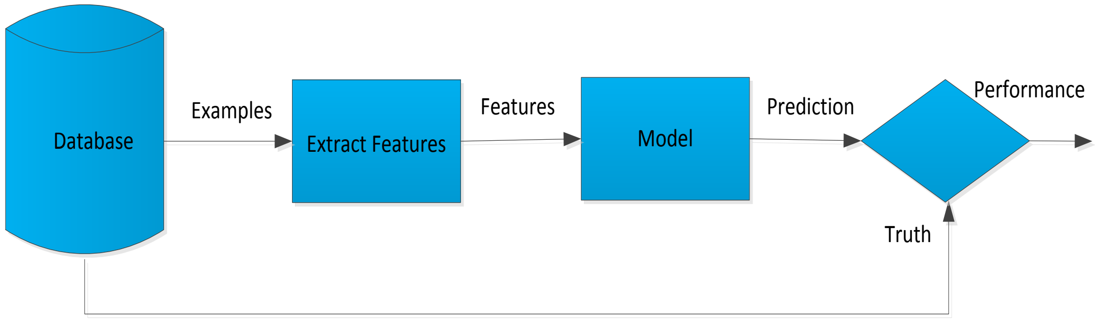


Figure 2.3 ESR system in testing phase

In Figure 2.3, testing phase of an ESR system is shown. In testing phase database provides the positive and negative examples, model produces the predictions about the examples provided and at last predictions are compared with the truth provided by the database and a performance rate is reached. According to this performance rate, ML algorithm or the feature set or other parameters of the ML algorithm may be needed to change. Then the whole training and testing phase start again. This process continues until an acceptable performance rate is reached.

After model creation and testing, this model is used in the production phase (Figure 2.4). In production phase, sounds come from the environment, model makes the predictions about the classes of sounds and we actually may not know the real truth. In production stage we may update sounds in database, update the model even we may change the ML algorithm. In each case we need the database.”

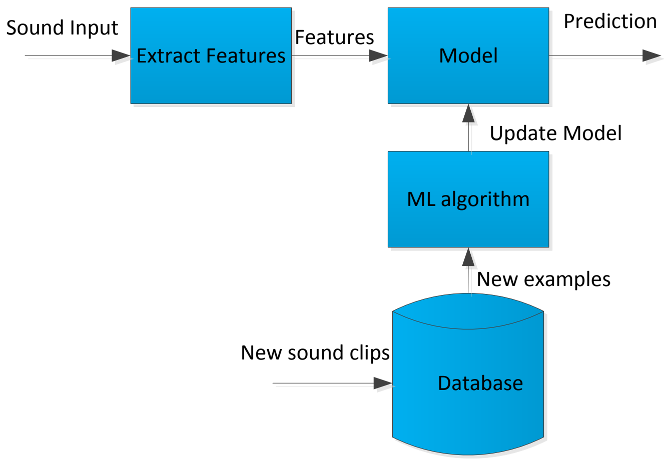


Figure 2.4 ESR system in production phase

2.4 Applications

A considerable amount of research has been made towards modeling and recognition of environmental sounds over the past decade. By environmental sounds, we refer to various quotidian sounds, both natural and artificial (i.e. sounds one encounters in daily life other than speech and music). Environmental sound recognition (ESR) plays a pivotal part in recent efforts to perfect machine audition[35]. With a growing demand on example-based search such as content-based image and video search, ESR can be instrumental in efficient audio search applications [36]. ESR can be used in automatic tagging of audio files with descriptors for keyword-based audio retrieval [37]. Robot navigation can be improved by incorporating ESR in the system [7], [8]. ESR can be adopted in a home-monitoring environment, be it for assisting elderly people living alone in their own home [9], [10] or for a smart home [11]. ESR, along with image and video analysis, find applications in surveillance [38], [39]. ESR can also be tailored for recognition of animal and bird species by their distinctive sounds [12], [13].

2.5 Other Database Examples and Problems

In this section we will review the mostly referenced databases in ESR related papers. The structure of the databases will be explored; pros and cons of the structure will be argued if applicable. In our published paper[4]:, we talked about this subject

“One of the databases that have been mostly reverenced is Real World Computing Partnership’s (RWCP) non-speech database [3]. Using a standard single microphone, acoustic signals of about 100 types of sound sources were measured in an anechoic room as the dry source. By using the impulse response of three kinds of rooms, convolution with the 48 kHz sampling signal was carried out to reconstruct the sound signal in respective rooms. The database structure (structure of directories) and navigation to sounds is explained and provided by an HTML page as shown in Figure 2.5 for dry sources. As an ESR developer we must examine the structure of HTML file and find the desired sound clips. Maintenance of the data and usage is not so easy with this kind of structure.

```
/nospeech ¡Ä Current directory
 +---/drysrc ¡Ä¡Ä¡Ä¡Ä Dry source
      +---/a1 ¡Ä¡Ä¡Ä¡Ä Collision sound source (wood)
      |     +---/cherry1 Beating handheld wooden board (small) B with a wooden stick
      |     |     +--- index.htm
      |     |     +--- cherry1.gif       Graph image of waveform
      |     |     +--- cherry1.jpg       Graph image of spectrogram
      |     |     +--- cherry1.spg       Spectrogram data
      |     |     +--- cherry1.wav       Sound for replay
      |     |     +---/48khz
      |     |     |     +--- 002.raw    Dry source (48kHz)
      |     |     |     :                ¡§
      |     |     |     +--- 100.raw    Dry source(48kHz)
      |     |     |
      |     |     +---/16khz
      |     |           +--- 002.raw    Dry source(16kHz)
      |     |           :                ¡§
      |     |           +--- 100.raw    Dry source(16kHz)
      |     |
      |     +---/cherry2         Beating handheld wooden board (small) B with a wooden stick
      |     +---/cherry3         Beating handheld wooden board (medium) B with a wooden stick
      |     +---/magno1          Beating handheld wooden board (small) C with a wooden stick
```

. **Figure 2.5** RWCP dry source sound clips structure

Detection and Classification of Acoustic Scenes and Events (DCASE) is an official IEEE Audio and Acoustic Signal Processing challenge. For challenge a database is prepared and it is also a resource for ESR researchers.

DCASE 2016 challenge consists of four tasks [25]. The goal of Task-2 is to detect sound events for example 'bird singing', 'car passing by' that are present within an audio segment. To be prepared for the challenge two datasets, train and development are given. Train is used to create the model and development is to test the model. For training all sound files are in one directory and a readme file explains the details, such as sampling frequency, quantization bit depth, etc. For development dataset there is an annotation text file for each sound clip. Each .txt file contains information about the onset, offset, and event class for each event in the scene, separated by a tab.

DCASE 2017 Task-2 dataset contains .yaml files for annotations [26]. .yaml files (Figure 2.6) can be read by a Matlab command, so it is easy to work with these structured files.


```
- annotation_string: "mixture_devtrain_glassbreak_000_d89a6bc22fba3b6efcb7b618ebed92cd.wav\t\
    25.27983994244287\t26.03983994244287\tglassbreak"
  bg_classname: cafe/restaurant
  bg_path: audio/a109_60_90.wav
  ebr: -6.0
  event_class: glassbreak
  event_length_seconds: 0.7600000000000002
  event_path: glassbreak/257642.wav
  event_present: true
  event_start_in_mixture_seconds: 25.27983994244287
  mixture_audio_filename: mixture_devtrain_glassbreak_000_d89a6bc22fba3b6efcb7b618ebed92cd.wav
  scaling_factor: 0.07767960012836954
  segment_end_seconds: 4.12
  segment_start_seconds: 3.36
```


Figure 2.6 DCASE 2017 task-2 development dataset yaml file for glass break

Dataset for Environmental Sound Classification (ESC) contains two databases. There are 10 classes and each class has 40 clips in ESC-10 dataset. ESC-10 is subset of ESC-50 which contains 50 classes and each class has 40 clips. There is a readme file in which there is a line for each clip explaining the details of the clip for each dataset [27].

In [28], a dataset of annotated urban sound clips are recorded and taxonomy for urban sound sources is proposed. The dataset contains 10 sound classes with 18.5 hours of annotated sound event occurrences. The dataset contains a CSV file (Figure 2.7) explaining the details of each recording.

```
slice_file_name,fsID,start,end,salience,fold,classID,class
100032-3-0-0.wav,100032,0.0,0.317551,1,5,3,dog_bark
100263-2-0-117.wav,100263,58.5,62.5,1,5,2,children_playing
100263-2-0-121.wav,100263,60.5,64.5,1,5,2,children_playing
100263-2-0-126.wav,100263,63.0,67.0,1,5,2,children_playing
```

Figure 2.7 CSV file explaining Urbansound database

Another ESR database is http://www.desra.org. In [23], the aim of this database, details of the design and the sources used are explained. It is designed as multi-purpose database. The database contains variety of sounds from different events with thorough background information for each sound file. It is accessible from the Internet (The database is not fully functional at the moment). The database was designed considering for admin tasks and general user level tasks. Web front end provides the functionality for user level tasks

http://www.auditorylab.org is [43] another database. This database was constructed by Carnegie Mellon University to examine the human ability to use sounds to understand what events are happening in the environment. All the sounds in the database are recorded in a controlled way. The laboratory and the recording media used are technically detailed. In the construction of database sounds are grouped by the event that makes the sound. Sounds of events like impact, rolling can be downloaded from this database.”

Chapter 3

ENVIRONMENTAL SOUND DATABASE DESIGN AND IMPLEMENTATION

3.1 Before to Database Design

Before designing a database, the requirements of the database must first be determined. We need to know what researchers expect from a database and what features they need. In order to demonstrate these requirements, the expectations of the people working with the environmental sound and the difficulties they face must be considered.

One of the main methodological difficulties in the execution of environmental sound research is to find and select the appropriate alerts to a specific project. It is unclear which stimulus will be included most appropriate in the test, given the almost infinite variety of environmental sounds and lack of a reliable perceptual-based classification, or even a description of a representative set. Other difficulties stem from the inadequacy of normative behavioral data, the methods used in previous research, variability in stimuli, the quality of stimuli, and the lack of criteria to reach agreement on evaluating a particular experimental relative fit. As a result, researchers entering the field will often need to repeat environmental sound sources from scratch and develop their own rules in selecting stimulants. [23].

Depending on the aims of a particular research, it is desirable and practical to present environmental sounds to listeners, either naturally produced or, in particular, sounding for a project. However, it is not always possible to produce real-time sound or to record specific areas of stimuli, especially when try to measure the environmental sound perception of listeners, because of the nature of sound and the large financial resources needed to make quality recordings. In such cases, researchers tend to turn to already available compilations of environmental sound recordings.

Utilizing existing environmental sound recordings as stimulants allows researchers to have consistent record quality and more duties on how to deliver stimuli and allows them to explore more environmental sound perception in a single study. Using freely available voices for other researchers allows for the duplication of findings and comparisons between studies. Other advantages are saving time, money and necessary staff and making environmental sound drills more accessible and accessible in general.

3.1.1 Selection of Environmental Sounds

Sounds to be used in environmental sound studies should be carefully selected. Practical considerations regarding sound quality and database resources should be evaluated in the experimental design process later, although they are extremely important. Since environmental sound perception is related to a large number of theoretical and applied fields, theoretical constructs and methodological requirements are used in relation to the population (eg children, adults, elderly, hearing impaired, hearing impaired, animal or machine listeners) and methods used in practice (eg perception, identification, or similarity rating). Previous research in the area of interest may be the most valuable guide when choosing specific types of sounds for a study. In general, recorded environmental sounds that have been employed in research thus far can be divided into the following four categories[23]:

- Isolated multiple sound tokens of the same physical sound source, or of the same type of sound source, have been used in research concerned with listeners' ability to perceive a particular property of the same soundproducing object or event. For instance, researchers asked listeners whether a bottle bounced or broke after a fall [44], whether listeners could judge the configuration of two palms during clapping from the sound of the claps [45], or whether the gender of a walker could be determined from the sound of his/her footsteps [46].

- Inventories of variable lengths comprising isolated environmental sounds produced by different sources have been used to investigate identification accuracy [47,48], naming agreement across different listener populations [49], and acoustic parameters and cognitive factors [32,50,51] involved in their perception.

- Temporal sequences of environmental sounds produced by different sources that do not overlap in time with each other have been used to investigate the context effects of the preceding and subsequent sounds [52] and the ease of learning different sound sequences [53].

- Mixtures of environmental sounds (a.k.a. auditory scenes), in which sounds from different sources may overlap with one another in time and in frequency, have been used to evaluate the accuracy of source identification by human and machine listeners [54].

3.1.2 Questions for the "Suitable Database"

In[23], Shafiro and Gygi ask some questions for choose suitable database for the project to be used. The following questions are intended to serve as a general guide for evaluating existing databases and the sounds they contain. Although answers to some questions may overlap with others with respect to the information they provide about the sounds in databases, together they will ensure that most crucial aspects of stimulus selection will have been considered:

- *"How many environmental sound types preented in a database*? It is important to check that all types of sounds selected for a project are represented. Simply going by the number of sounds in a database can be misleading, because most or all of the sounds may be various recordings of the same type of sound source (e.g., singing birds or car crashes). However, it is also preferable to have alternative tokens of the same sound sources, since some tokens may be of better quality than others. In addition, having multiple tokens of the same type of sound provides the researcher with an opportunity to test for possible differences in perceiving different tokens of the same environmental sound. ”

- *“How well documented are the recording equipment and setup?* Knowing what kind of recording equipment was used can help to evaluate the quality of recordings and can indicate the likelihood of certain recording artifacts, such as a high signal-to-noise (S/N) ratio or distortions in the sounds’ spectra. It is also important to know how the microphone or microphones (for stereo recordings) were mounted, the reverberation characteristics of the space where recordings were made, the distance from the sound source (near or far field), and whether the recorded signal already reflects the head related transfer function. Having recording information can also help to replicate the original study, if necessary. Unfortunately, the details of recording setup usually are not available for most existing environmental sound databases, possibly because this information is of little interest to most consumers.”

- *“How thoroughly are the sound sources described?* Other information of potential value to the researcher is a detailed description of the sound source. For instance, a sound in a database may be labeled simply airplane, or it can describe the airplane’s model, whether it is a prop or a jet, or whether it is approaching, taking off, or flying by. Knowing these details may be necessary for projects concerned with differences in perception of different sounds produced by the same type of sound source.”

- *“What was the original purpose for making a particular recording? (b) Who is the intended consumer of recordings in a database?* Recording quality may differ considerably depending on the purpose of the recordings. Although most existing recordings can be used for a variety of purposes, when thorough documentation is lacking, the purpose of making a recording or its intended application can be an indirect indication of the emphasis placed on recording quality. For instance, sounds recorded for use as special effects in computer games or computer animations may have a low sampling rate in order to minimize processing load and storage demands. These sounds would be less likely to have adequate representation of high-frequency information. In addition, Foley sounds (specifically designed effect and imitation sounds frequently used in film and television, which are often included in sound effects collections) may sometimes be provided instead of real-world recordings. Although perceptually such sounds may closely resemble the sounds of actual objects and events (e.g., substituting the sound of coconut halves for that of horse hoofs), the information they contain may be distorted or represent only a caricature of the original sound. Recordings of actual soundproducing objects should always be preferred, unless the comparison of Foley sounds with real sounds is being investigated [55]."

- *"What editing has been performed after a recording was made?* Similarly, sounds recorded for entertainment and special effects are more likely to have been edited after the recording was made. Edited recordings may not represent some information contained in the original sound. Thus, it is important to know what, if any, editing has been performed, and how the originally recorded signal has been modified."

- *"When were recordings made, and what was their original storage medium?* Some of the sound recordings that are now available in a digital form have been transferred from magnetic tapes and vinyl records recorded several decades ago. These recordings still contain the limitations imposed by the original media (e.g., smaller dynamic range, higher S/N ratio) or artifacts specific to particular media (e.g., record needle noise). Although this does not necessarily warrant their dismissal from the list of potential stimuli, these limitations need to be appraised during the stimulus selection process."

- *"What is the format in which recordings are available now?* The question of format is relevant to the issues of compatibility with the stimulus presentation software to be used in the study. Although most standard digital audio formats (e.g., .wav, .aiff, .au., or .mp3) can be easily converted one into another, for other formats conversion tools are not always readily available (e.g., .snd for Neuroscan or many proprietary file formats). In addition, some commonly used digital audio formats involve methods of file compression that may sacrifice information in the signal for increased storage capacity (e.g., MPEG or RealAudio). Recordings available in compressed file formats may still be useful for a particular project, depending on its goals. However, this question needs to be addressed before stimuli are chosen.”

- *“Have recordings in this database been used in research before?* Examination of previous research that employed sounds in a database can be a valuable source of additional information about the quality of the recordings. Although this does not ensure that the sounds in a database are of high quality and previous research might have had different requirements for its stimuli, positive and negative aspects of the recordings might be described in reports or discussed with researchers involved in earlier work with the same recordings. Moreover, if behavioral data have been previously collected on some recordings, it can be supplemented by collecting similar kinds of data on other recordings in the same database or by examining other behavioral measures, using the same stimuli as before.”

- *“What is the price range for obtaining selected recordings?* Prices for existing environmental sound recordings vary widely. Sounds from a number of electronic databases can be obtained free of charge (e.g., www.findsounds.com), whereas others may cost from several hundreds to over a thousand dollars (e.g., www. sound-ideas.com;). Higher priced databases often provide more thorough descriptions of the sounds they contain and are, generally, better documented. However, this may not always be the case. It may also be possible to find high-quality sounds without spending a lot of money.”

- *“Will it be possible to share these recordings with other researchers?* Some commercially available recordings cannot be freely redistributed to others, even if only for research purposes. Frequently, recordings advertised as “royalty free” by their manufacturers may be shared with others on a royalty-free basis only if synchronously embedded into a movie track, animation sequence, or computer game. Thus, unless special arrangements can be made

> with companies selling such recordings, other researchers who would be interested in using these recordings as stimuli will likely need to invest in acquiring rights to the whole database. This becomes an important consideration when recordings from more expensive databases are used. The high cost of a database can become an obstacle for further research work on the stimuli by making it prohibitively expensive for other researchers to use the same recordings as stimuli."

From the above questions, we have identified the database requirements, which are the first stage of design.

3.2 Database Design

The databases reviewed in Section 2.5 are prepared just for resources for the development of ESR applications. Many of them provide the data and the files for correctly handling the data. In this context they are valuable resources for all ESR researchers. Desra [23], provides extra tools such as a web graphical user interface for searching and testing. In our published paper[4], we talked about database design:

"In the development process of an ESR system, some main functions training, testing and production explained in Section 2.3. During this development process, many sound clips are used; new sound clips may be added or deleted. We deal with many features extracted from these sound clips; we create models using different ML algorithms. Then we compare the models; try to find best features and best ML algorithms. This loop continues. Our first goal in this database design is to help researchers as much as possible to ease the burden of handling data. The second goal is to help maintenance of the data. Our data is sound clips, features, algorithms and models. Addition, deletion, searching, annotation, backup and restore can be thought as the maintenance task.

3.2.1 Taxonomy of Environmental Sounds

We need a taxonomy to be able to store, search and retrieve the data from the database. During literature review we see two kinds of taxonomy. In the first taxonomy, the sounds can be grouped by the event that makes the sound. In Figure 3.1, grouping of sounds defined in [43] is seen.

title	author	type	modified
LEGAL NOTICE	Auditory Lab User	Page	2009-06-24 21:16
Impact Events	Auditory Lab User	Folder	2009-06-25 10:15
Deformation Events	Auditory Lab User	Folder	2009-06-25 10:16
Rolling Events	Auditory Lab User	Folder	2009-06-25 10:17
Air Events	Auditory Lab User	Folder	2009-06-25 10:17
Liquid Events	Auditory Lab User	Folder	2009-06-25 10:18
Experiment Stimuli	Auditory Lab User	Folder	2009-06-25 10:19

Figure 3.1 Grouping of sounds based on sound producing events [43]

Another hierarchical grouping is seen in Figure 3.2. This is also classified in the first taxonomy. This is the grouping defined in [31] based on sound producing events and the listeners' description of the sound.

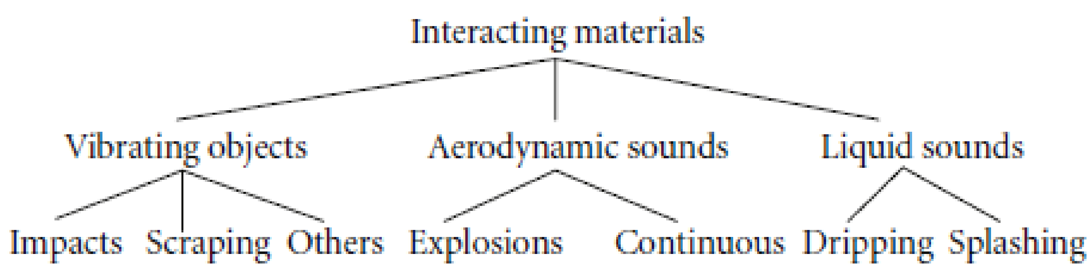


Figure 3.2 Grouping of sounds defined in [31]

Second taxonomy is based on the sound source [28][56]. The subset of the taxonomy defined in [56] is taken into consideration for urban sounds given in [28].

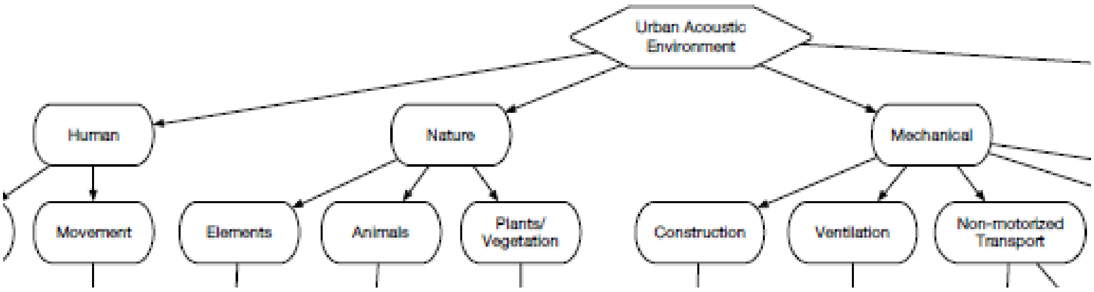


Figure 3.3 A part of the urban sound taxonomy defined in [28]

In our database design we decided to use taxonomy of the sounds like the one described in [28]. In [28] one of the aims is to go to as much detail as possible going down to low level sound sources. In our database design we don't need the steeps between main branch and the leaves. For example instead of human→ movement→footsteps, we use just as human→ footsteps. In Figure 3.4, how the sounds are grouped in our database design is explained."

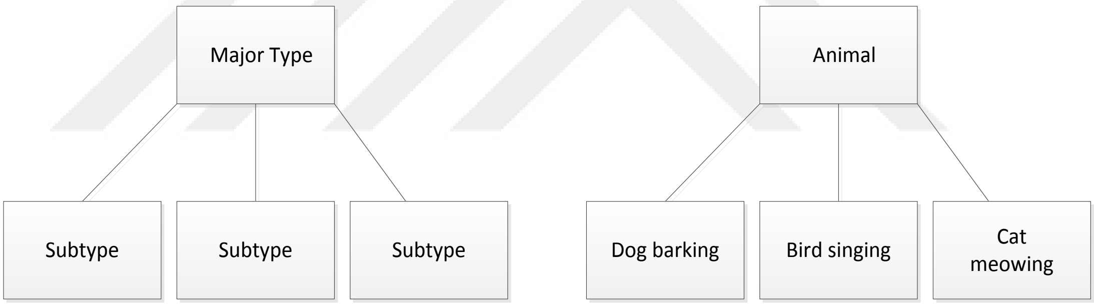


Figure 3.4 Environmental sounds grouping in the database

3.3 Database Requirements

We need to first derive the expected functions and properties of the database, that is, the requirements and then realize them. We can sort the database requirements as follows:

3.3.1 Non-Functional Requirements of the Database

Sound Clips:

- There will be sound clips in the database. These sound clips will contain environmental sounds.

Exp: As an example you go into forest and you can make a recording in the same clip such as a dog barking, birds' sound and a hunter's gunshot.

- So each sound clip can have more than one environmental sound.
- These environmental sound clips can belong to different major environmental sound types.
- Origin of these sound clips, such as own recording, from an internet site or from another database, should be entered to DB.
- Recording details should be entered.
- Size should be entered.
- File type should be entered.
- File path should be entered.

Environmental Sound Types:

- There should be major types, such as human, nature, music.
- There should be subtypes, such as dog bark, gunshot etc.
- Each subtype should belong to one major type.

Environmental Sound Subtypes:

- Subtype unique name should be entered.
- Subtype major type to which it belongs to should be entered.

Environmental Sound Clips:

- These are extracted sound clips from sound clips.
- These clips should have a subtype.
- It is known that from which sound clip it is extracted.
- Start and end sample index or start/end time should be known.

Features:

- Features should be extracted from environmental sound clips.
- There should be unique feature name.
- Each feature should have its own properties.
- Exp: MFCC can have variable size filter bank.
- Each property has its property id.
- (The features and properties can be implemented in two tables. One features table and a table for each feature with the feature name as the table name.)

Feature Extractor:

- Features are extracted from environmental sound clips by way of feature extractors.
- We should know the path of the feature extractor.
- There should be a help file explaining the used options for the feature.
- There should be a unique name for the feature extractor.

Environmental Sound Subtype Features:

- The extracted feature path should be entered.
- The extracting script name should be known.
- The feature name should be entered.
- The same feature from the same environmental sound clip can be extracted with different options of the feature.

Noise Clip Type:

- For each noise clips there should be type, such as traffic, crowd, etc.
- These type names should be known.

Noise Clips:

- Noise clips should be entered to database.
- The path of these clips should be known.
- Recording details should be entered.
- Noise clip type should be entered.
- Size and file type should be entered.

Sampling:

- Recording details should be entered to this table such as, sampling rate, quantization bit depth, etc.
- Each recording type should be unique id.

Sound In Noise:

- To measure the performance of detection or recognition algorithms in noise, there will be some scripts which will embed environmental sound clips into noise clips. The resulting file details should be entered in database. Regarding these file;
- Noise file name should be known
- Embedded environmental sound clip name should be known
- Start index of the environmental sound clips should be known
- End index of the environmental sound clips should be known
- Start time of the environmental sound clips should be known
- End time of the environmental sound clips should be known
- SNR level of the file should be known
- Only one environmental sound clip should be embedded into one noise clips.
- Same environmental sound clip should be entered into same noise clip with different SNR levels.

3.3.2 Functional Requirements of the Database

- There should be scripts which will take algorithms as arguments and extract the environmental sounds from sound clips. These extracted clips should be entered to the database as explained in B.
- There should be scripts which will embed environmental sound into noise clips at desired Signal to Noise Ratio (SNR) level.
- There should be scripts which will extract features from environmental sound clips and store to the database.
- The scripts, their help files, paths should be stored in the database.

3.4 Database Implementation

Unlike previous studies, we decided to use the relational database method when implementing our database. The reason for this decision is the existence of data classes with intimate relationships that are intertwined in the environmental sound databases. Having to scroll through these data classes or enter the same data many times during new information entry brings a huge loss of time and effort.

In addition, one of the biggest problems in databases is that the information obtained or existing information about the audio data is not well documented or categorized, even if they are not categorized regularly.

The implemented database will include some data which can be stored by way of data types found in the standard database software and also sound files. These sound files will not be stored to database instead they will be stored in the operating system file structure and the path to this file is just recorded in the database.

Microsoft SQL 2008 Database Server is used to create tables. The tables and the relationships between them are shown in Figure 3.5. The scripts providing functionality are coded in Matlab 2011a. The files are stored in “mat” type when required.

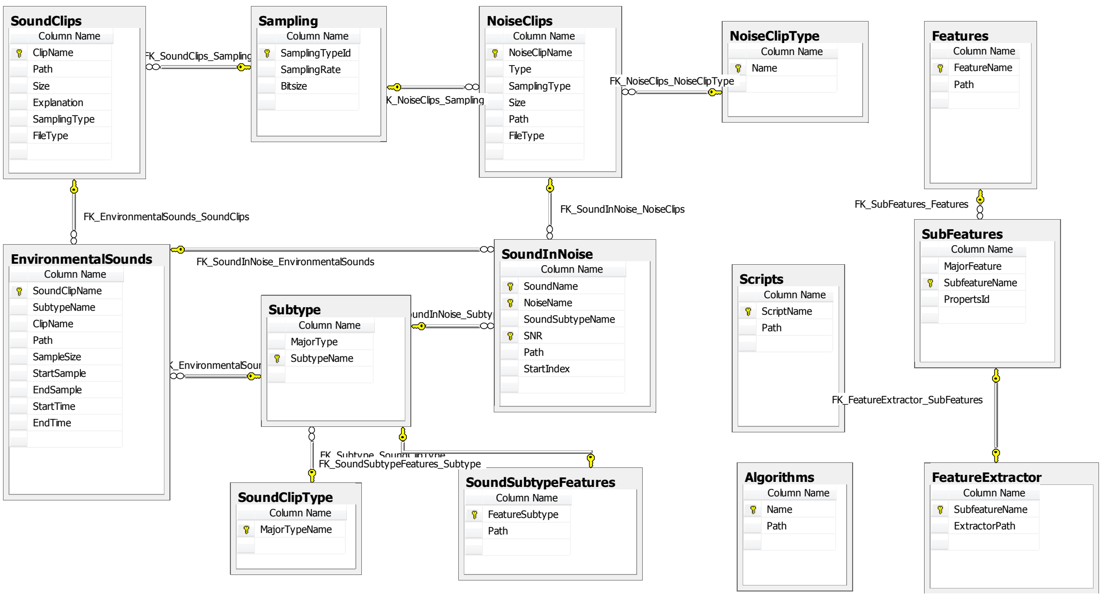


Figure 3.5 Database Table Diagram

For functional requirements the following scripts are coded by Yüksel Arslan using Matlab.

- **Environmental Sound Detector:** This script is an interface between algorithms that extracts environmental sounds from the sound clips. Different algorithms can be used here for extraction of the environmental sounds. The algorithms must conform to this script interface definition. This script takes the algorithm name that will be started and the environmental sound types which will be searched are given as arguments.
- **Environmental Sound Embedder:** This script takes the type of the environmental sound, noise type, SNR level and at last the number of required record count to be created. The script merges the environmental sound clip with the noise clip at the desired SNR level and records it to the database table. The scenario finds the current SNR value from the formula ($20\log_{10}(Vs/Vn)$). This value is subtracted from the desired SNR value to obtain the difference. The environmental sound clip is strengthened to the value obtained. The SNR is adjusted according to this scenario.
- **Feature Extractor:** This script acts as an interface between feature extractors. It takes the path of the feature extractor from the features table, feature name and the environmental sound type from command line. After extracting features, it is saved as a mat file.

- **Data Loader:** To load data from other databases this script is used to interface with data loading scripts.

3.5 Example of Load Data to Database from Another Database ‘Urban Sounds’

After database and scripts implementation , to test the functionality of the database, it was necessary to load various data into the database. Since creating new sound clips is not possible in terms of time, effort, and possibility, we used the ‘URBANSOUND8K’ dataset available on the internet. This dataset can be requested for academic studies by filling out a form from link [57].

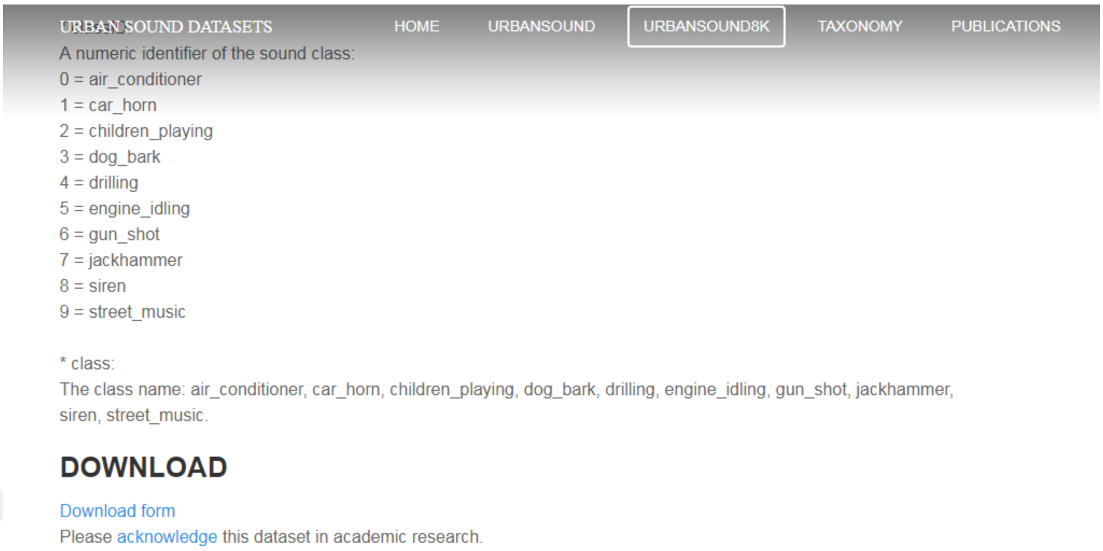


Figure 3.6 Urban Sound Datasets Web Page

This dataset contains 8732 labeled sound excerpts (<=4s) of urban sounds from 10 classes: air_conditioner, car_horn, children_playing, dog_bark, drilling, enginge_idling, gun_shot, jackhammer, siren, and street_music. The classes are drawn from the urban sound taxonomy[28]. For a detailed description of the dataset and how it was compiled please refer this paper.

All excerpts are taken from field recordings uploaded to www.freesound.org. The files are pre-sorted into ten folds (folders named fold1-fold10) to help in the reproduction of and comparison with the automatic classification results reported in the article above.

	A	B	C	D	E	F	G
1	slice_file_name,fsID,start,end,salience,fold,classID,class						
2	100032-3-0-0.wav,100032,0.0,0.317551,1,5,3,dog_bark						
3	100263-2-0-117.wav,100263,58.5,62.5,1,5,2,children_playing						
4	100263-2-0-121.wav,100263,60.5,64.5,1,5,2,children_playing						
5	100263-2-0-126.wav,100263,63.0,67.0,1,5,2,children_playing						
6	100263-2-0-137.wav,100263,68.5,72.5,1,5,2,children_playing						
7	100263-2-0-143.wav,100263,71.5,75.5,1,5,2,children_playing						
8	100263-2-0-161.wav,100263,80.5,84.5,1,5,2,children_playing						
9	100263-2-0-3.wav,100263,1.5,5.5,1,5,2,children_playing						
10	100263-2-0-36.wav,100263,18.0,22.0,1,5,2,children_playing						
11	100648-1-0-0.wav,100648,4.823402,5.471927,2,10,1,car_horn						
12	100648-1-1-0.wav,100648,8.998279,10.052132,2,10,1,car_horn						
13	100648-1-2-0.wav,100648,16.699509,17.104837,2,10,1,car_horn						
14	100648-1-3-0.wav,100648,17.631764,19.253075,2,10,1,car_horn						
15	100648-1-4-0.wav,100648,25.332994,27.197502,2,10,1,car_horn						
16	100652-3-0-0.wav,100652,0.0,4.0,1,2,3,dog_bark						
17	100652-3-0-1.wav,100652,0.5,4.5,1,2,3,dog_bark						

Figure 3.7 UrbanSound8K Dataset metadata csv file

In addition to the sound excerpts, a CSV file(Figure 3.7) containing metadata about each excerpt is also provided.

8732 audio files of urban sounds (see description above) in WAV format. The sampling rate, bit depth, and number of channels are the same as those of the original file uploaded to Freesound (and hence may vary from file to file).

After downloading the dataset to the computer, we wrote a matlab script that extracts data from the csv file that contains the metadata of each sound clip and transfers it to the database. Also script saves the sound clips in computer files as wav files with sampling rates of 16000 Hz and bitsize = 16 bits.

In this script, we added audio clips and metadata to both the ‘Sound Clips’ and ‘Environmental Sounds’ tables because they met the requirements of the two tables. Despite the fact that the types of audio clips are obvious, the requirement of the Sound Clips table is met by the feature of the clip containing more than one o sound example(For instance a dog bark clip have 5 times dog bark). On the other hand, these clips met the requirement of the 'Environmental Sounds' table because they are created by cutting a part of the original clip.

```
end

 FileType = 'wav';
 myMat = dir(imppath1);
 sizeBull = myMat.bytes;
 data = {FilenameGundb,imppath1, sizeBull,Explanation,1,FileType};
 colnames = {'ClipName','Path', 'Size', 'Explanation', 'SamplingType','FileType'};
 tablename = 'SoundClips';
 fastinsert(conn,tablename,colnames,data);

 FilenameGundb = ('US_');

 end
```

Figure 3.8 A part of matlab script

After running this(Figure 3.8) script, 8732 sound clips were added to the database along with the metadata. As mentioned in previous chapters , sound files were not be stored to database instead they were stored in the operating system file structure and the path to this file is just recorded in the database.

SoundClipName	SubtypeName	ClipName	Path	SampleSize	StartSample	EndSample	StartTime	EndTime
US_air_condition...	air_conditioner	US_air_condition...	D:\Yüksek Lisans...	705644	1	64000	0	4
US_air_condition...	air_conditioner	US_air_condition...	D:\Yüksek Lisans...	705644	1	64000	0	4
US_air_condition...	air_conditioner	US_air_condition...	D:\Yüksek Lisans...	705644	1	64000	0	4
US_air_condition...	air_conditioner	US_air_condition...	D:\Yüksek Lisans...	705644	1	64000	0	4
US_air_condition...	air_conditioner	US_air_condition...	D:\Yüksek Lisans...	705644	1	64000	0	4
US_air_condition...	air_conditioner	US_air_condition...	D:\Yüksek Lisans...	256044	1	63999	0	3,99999999
US_air_condition...	air_conditioner	US_air_condition...	D:\Yüksek Lisans...	256044	1	63999	0	3,99999999
US_air_condition...	air_conditioner	US_air_condition...	D:\Yüksek Lisans...	256044	1	64000	0	4
US_air_condition...	air_conditioner	US_air_condition...	D:\Yüksek Lisans...	256044	1	64000	0	4,00000000
US_air_condition...	air_conditioner	US_air_condition...	D:\Yüksek Lisans...	256044	1	64000	0	4,00000000
US_air_condition...	air_conditioner	US_air_condition...	D:\Yüksek Lisans...	256044	1	64000	0	4,00000000

Figure 3.9 Environmental Sounds Table

Thus, an example showing the use of the database was made. In the next sections, we need to have a database with actual data in order to improve the web application and user interface more accurately.

3.6 Our Sound Recordings

In addition to the Urban Sounds, we have added 24 voice recordings to our database. In the picture below there is a table of information about these environmental sound clips. All clips are recorded for this study.

Clip Name	Clip Type	Environment
bird_cock_clip	Animal	Farm
Birds_tweeting	Animal	Farm
car_washing	Mechanical	Outdoor
Cardboard_cutting_machine	Mechanical	Factory
Cardboard_production_machine	Mechanical	Factory
Cock_crowing	Animal	Farm
computer_fan_humming	Mechanical	Office
dog_barking	Animal	Outdoor
Drill_1	Mechanical	Indoor
Drill_2	Mechanical	Indoor
Drill_3	Mechanical	Indoor
Drill_4	Mechanical	Indoor
Fire_control_system_buzzer	Mechanical	Office
Helicopter_engine	Mechanical	Airport
Kitchen_robot	Mechanical	Indoor
School_Student_noise	Human	Outdoor
Water_boiling_sound	Nature	Indoor
Mall_crowd_1	Noise	Indoor
Mall_crowd_2	Noise	Indoor
Out_of_mosque_crowd	Noise	Outdoor
traffic_1	Noise	Outdoor
Traffic_ivedik	Noise	Outdoor

Figure 3.10 Our records table

Chapter 4

GRAPHICAL USER INTERFACE AND TOOLS

4.1 Overview to GUI

In the previous section, the implementation of the database using Microsoft SQL Server was shown. As an example of usage, ‘URBANSOUND8K’ dataset was added to the database using matlab script.

However, it is necessary to know the SQL language to use the database. Even if the SQL language is known, there is a significant loss of time and effort due to the large number of tables and data that can be edited and checked in the database.

As a solution to this problem, it is planned to develop an interface and web application that will simplify the use of the database and reduce it to the user level.

As a first step, the requirements of the interface were determined:

- Administrators can use the GUI for meta data entrance, deletion and update such as major types, subtypes, recording details.
- Administrators can use the GUI for data load manually or using the data loader script.

Graphical user interface was developed with these requirements in mind.

The GUI is developed using Microsoft Visual Studio with C# language. GUI has admin utilities. In Figure 4.1 it is seen the part of GUI providing admin operations. These admin operations are additions and deletions of major types, subtypes and sampling information. Besides the GUI provides the administrators load data one by one or using a script.

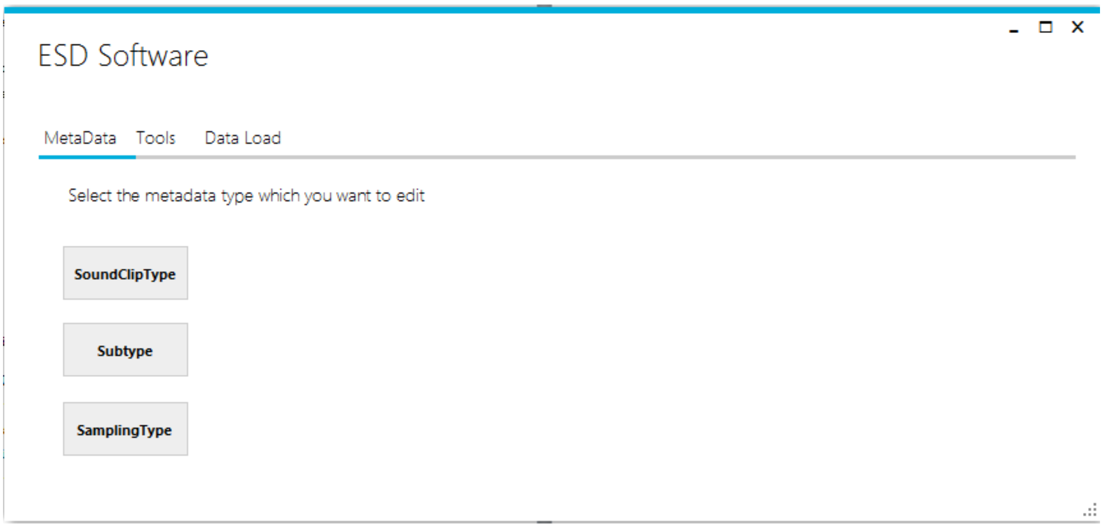


Figure 4.1 GUI – Metadata Load Screen

4.2 Meta Data Edit from Database

In Meta Data screen(Figure 4.1) , major type , subtype and sampling type data can be added , deleted and updated.

4.2.1 Edit Sound Clip Type

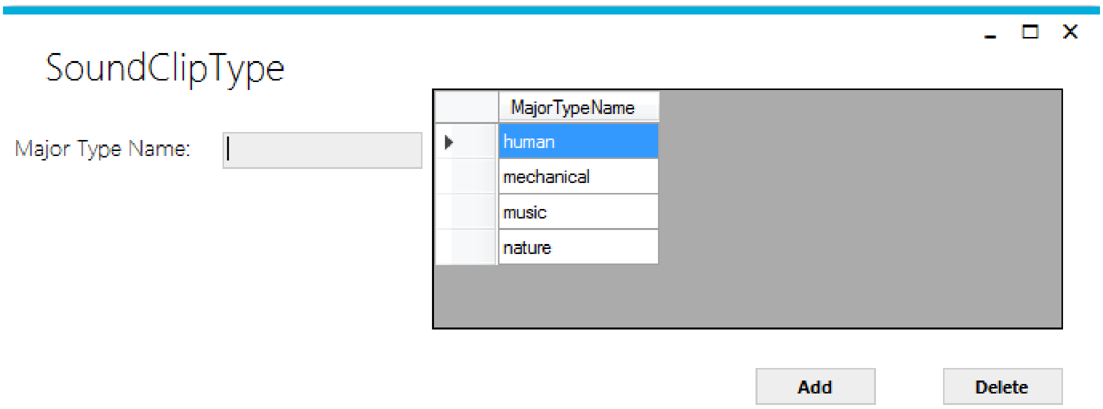


Figure 4.2 GUI- Edit Sound Clip Type Screen

In this screen , major sound clip types like human, nature etc. can be added, deleted and updated.

4.2.2 Edit Subtype

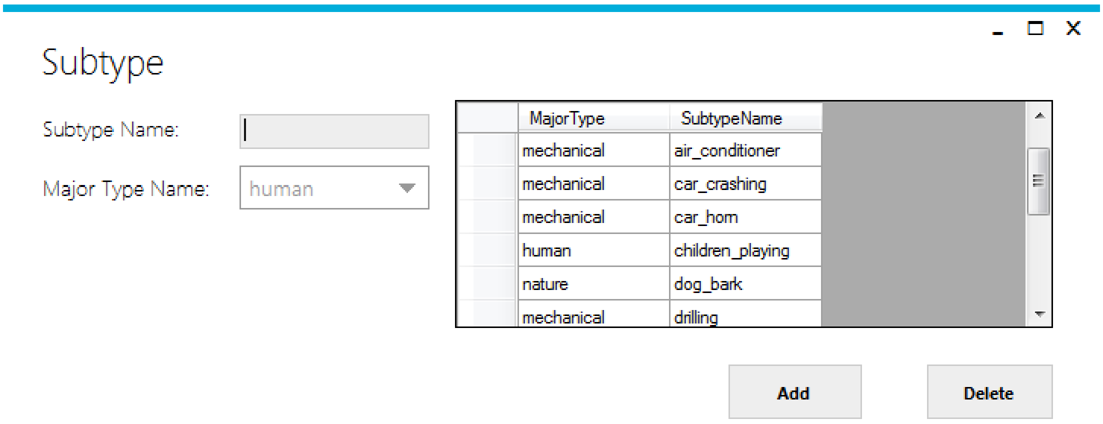


Figure 4.3 GUI- Edit Subtype Screen

In this screen , clip subtypes like car crashing, dog barking etc. can be added, deleted and updated.

4.2.3 Edit Sampling Type

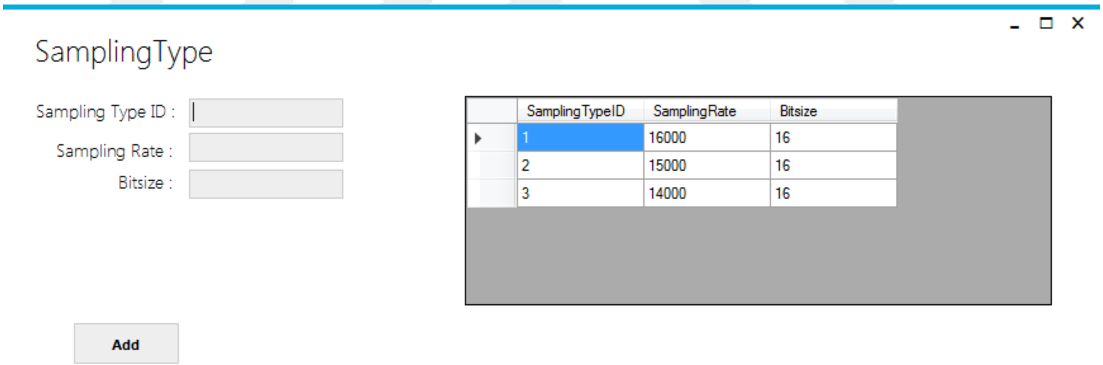


Figure 4.4 GUI- Edit Sampling Type Screen

In this screen , sampling types with sampling rate and bitsize can be added.

4.3 Sound Data Edit From Database

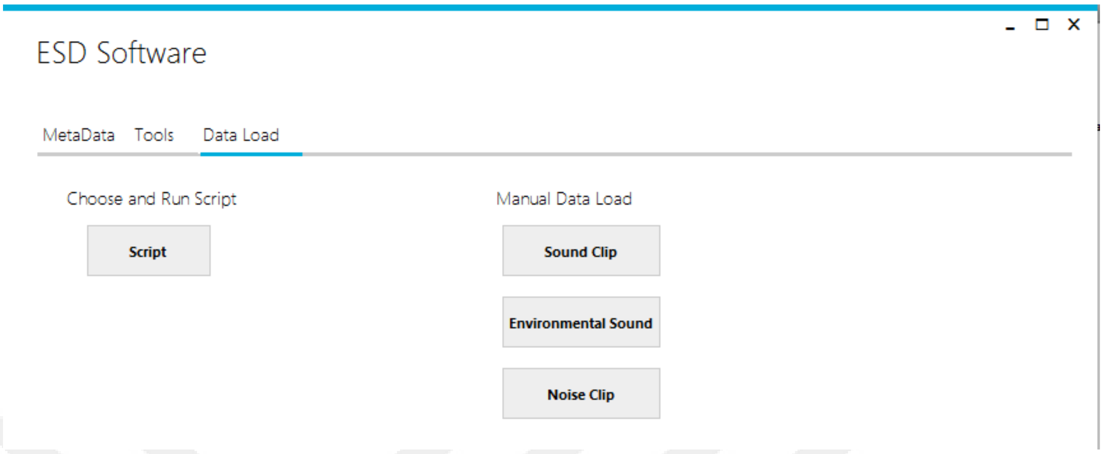


Figure 4.5 GUI – Sound Data Load and Edit Screen

In Data Load screen , Sound Clip ,Environmental Sound Clip and Noise Clip data can be added deleted and updated. In Figure 4.6, on the left side of the screen are the areas to be filled with the features of the clips, and on the right side are the existing records on the table. Sound Clip and Noise Clip screens are similar to the Environmental Sound screen.

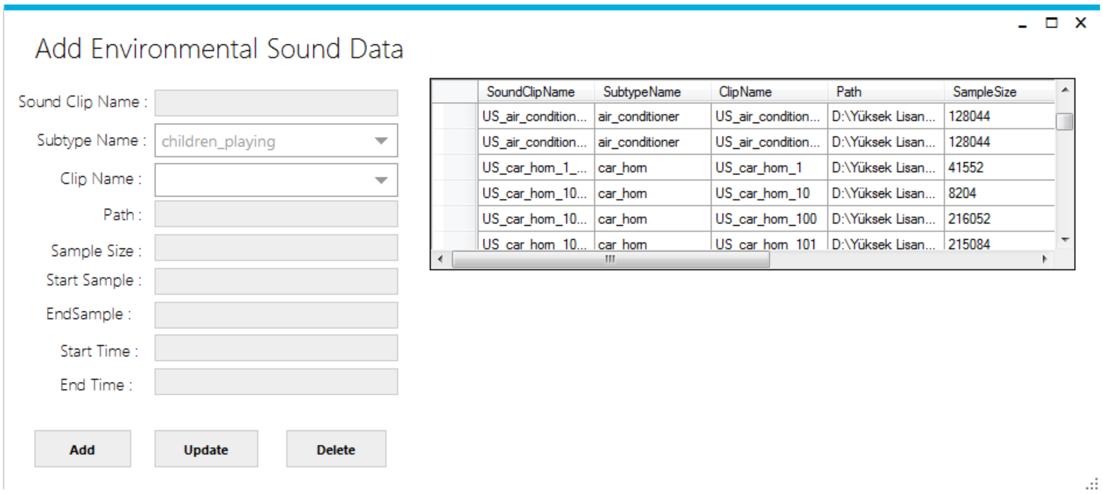


Figure 4.6 Edit Environmental Sound Table Screen

Besides the GUI provides the users load data one by one or using a matlab script (Figure 4.7).

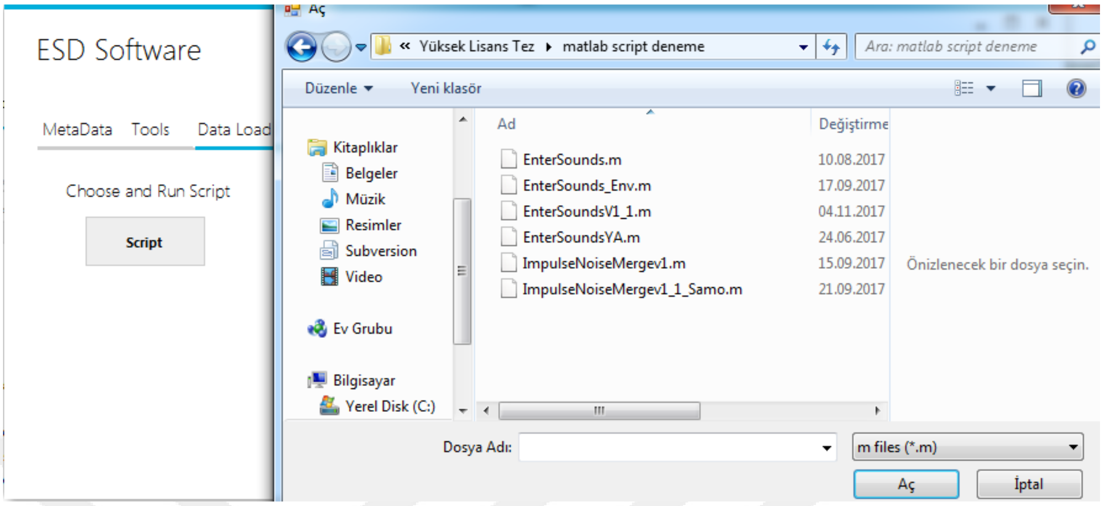


Figure 4.7 GUI- Loading data using matlab script-script selection

Chapter 5

WEB APPLICATION

5.1 Overview

In the previous chapter, a graphical user interface(GUI) was implemented to simplify the use of the database and reduce it to the user level. However, this GUI was limited to admin-level access to the database. There were no controls for other researchers or users other than Admin.

Also, the fact that the developed application is a pc interface makes it easy to use the database, but it is difficult for all researchers and users to access it. Even if the application is reached, the database must be available on the user's computer in order to run.

The main aim of the thesis was to create a common resource for environmental sound research and make it easily accessible. When these two problems are taken into consideration, it seems that these targets are removed.

In this section, a solution to these problems would be to implement a web application that has the admin controls in the interface as well as the tools that other users can use the database. This will ensure that all users have easy access to the database.

5.2 Front End Functions of Web Application

The Web Application is developed using Microsoft Visual Studio with Asp.Net language. Web application has admin and user utilities. Login is required to use the application. In Figure 5.1 it is seen the main page of application where all functions can be accessed.

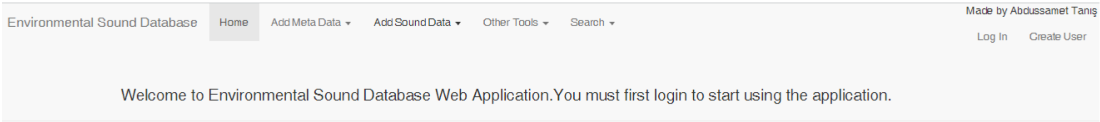


Figure 5.1 Web Application – Main Page

There are four main functions that are essential to the functioning of the database. They are user enrollment and access; add metadata and sound data to the database; the search engine; and testing.

User Levels: There will be two different user levels. Admin would have all rights to the database and be limited to people working on the project. They would be able to add or delete meta and sound data to database.

Other users can be search sound or noise clips and downlad them. Also they can be test their scripts.

Adding / Editing Data: As mentioned, admin will be able to freely upload sounds, edit all sound data. Only admin will be able to add or delete sound and meta data.

Search Engine: All Users can search the database to see the sound clips with desired type.

Test: All users can be test their algorithms on the sound clips given by the database.

5.2.1 User Enrollment

Registered users are required to use the application. Access to other parts of the application is not allowed without login as user(Figure 5.2). There will be two different user levels. Admin would have all rights to the database and be limited to people working on the project. They would be able to add or delete meta and sound data to database.

Other users can be search sound or noise clips and downlad them. Also they can be test their scripts.

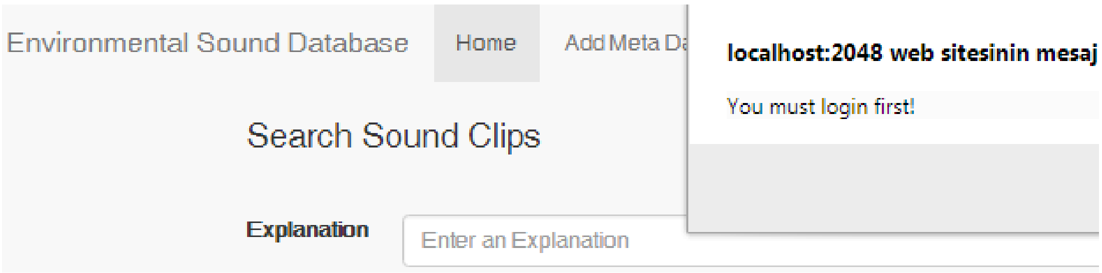
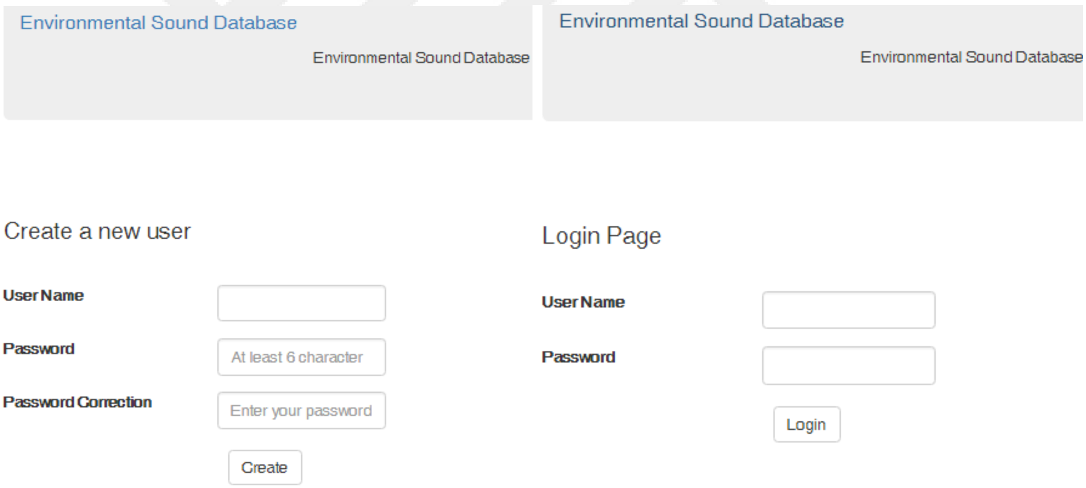


Figure 5.2 Web Application - Attempting to reach the app without login

A new user can be created by entering the user name and password from the user creation page(Figure 5.3(a)). Once the user has been created, login can be done from the login screen(Figure 5.3(b)).



(a) (b)

Figure 5.3 Web Application- (a) Create User Page (b) Login Page

5.2.2 Adding / Editing Data (Admin Controls)

There are two types of admin control in the application.

- Edit Metadata : Which is manage metadatas about clips like major-subtypes, sampling types etc.
- Edit Sound Data: Which is manage sound clips data wtih using existed metadatas and additional datas.

5.2.2.1 Edit Metadata

There are three types of metadata can be editing in the application. First of them is edit sound clip types. Sound clip types are major type of sound clips. As we mentioned in chapter three , sounds are grouped in our database design like Figure 5.4.

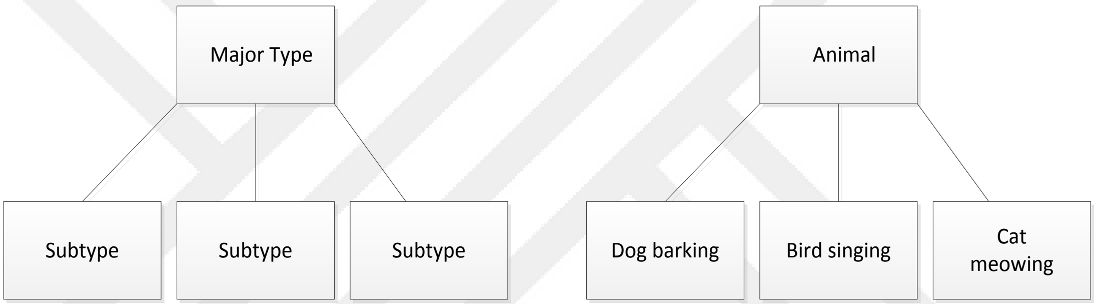


Figure 5.4 Environmental sounds grouping in the database

In edit sound clip type page admin can add or delete major types. Below the page ,existing major types are listed(Figure 5.5) .

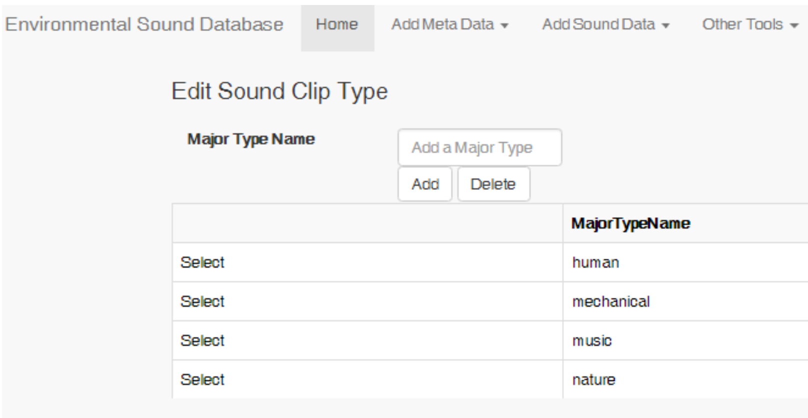


Figure 5.5 Web Application- Edit Sound Clip Type Page

The second of the metadata that can be edited is sound subtypes. These are subtypes of major types. Each subtype has to be linked to a major type like thunder → nature.

In edit subtype page admin can add or delete subtypes. Below the page ,existing subtypes and, majortype that they are linked to listed(Figure 5.6).

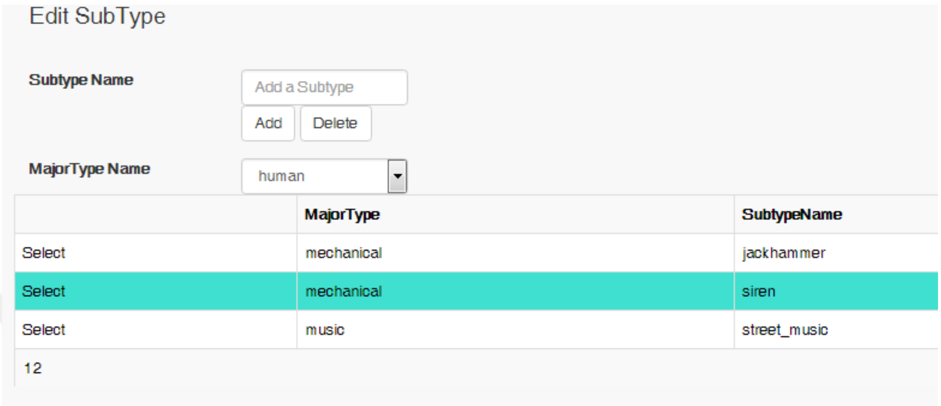


Figure 5.6 Web Application- Edit Subtype Page

The last of the metadata that can be edited is sound sampling types. Sampling types of sounds are contains two features; sampling rate and bitsize. Sampling rate is the number of samples of a sound that are taken per second to represent the event digitally. The more samples taken per second, the more accurate the digital representation of the sound can be. Much like the sample rate defines the frequency response as it divides up the horizontal axis of the waveform, the bitsize defines the dynamic range of the sound as it describes the amplitude of the waveform at each sample point.

In edit sampling type page admin can add new sampling types. Below the page ,existing sampling types and, features of them are listed(Figure 5.7).

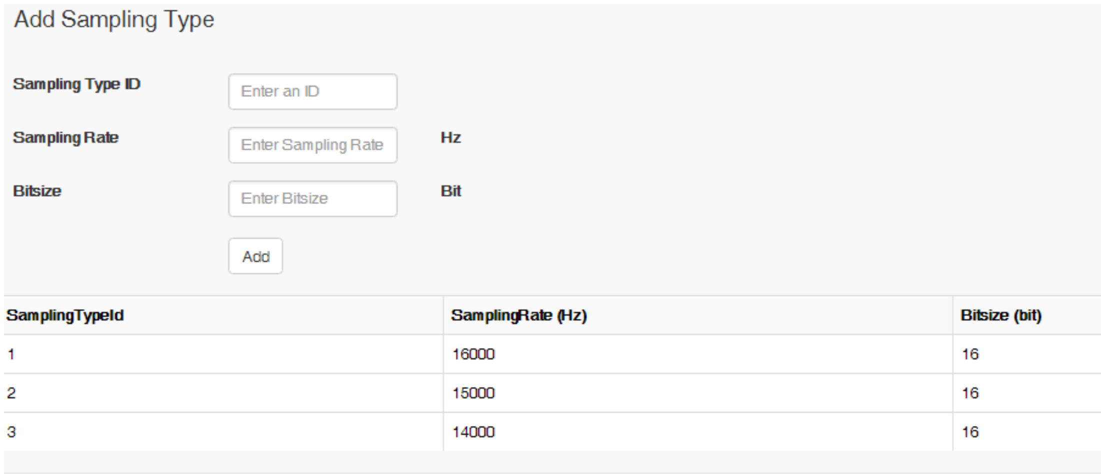


Figure 5.7 Web Application- Edit Sampling Type Page

5.2.2.2 Edit Sound Data

There are three types of sound data can be editing in the application. First of them is edit sound clips. Sound clips are contains environmental sounds. Each sound clip have more than one environmental sound. These environmental sound clips can belong to different major environmental sound types. As an example you go into forest and you can make a recording in the same clip such as a dog barking, birds' sound and a hunter's gunshot. Origin of these sound clips can be own recordings or from an internet site or from another database.

In edit sound clip data page(Figure5.8), admin can add or new sound clip data by entering the required features or delete existing sound clips.. These features are:

Clip Name: Name of the clip.

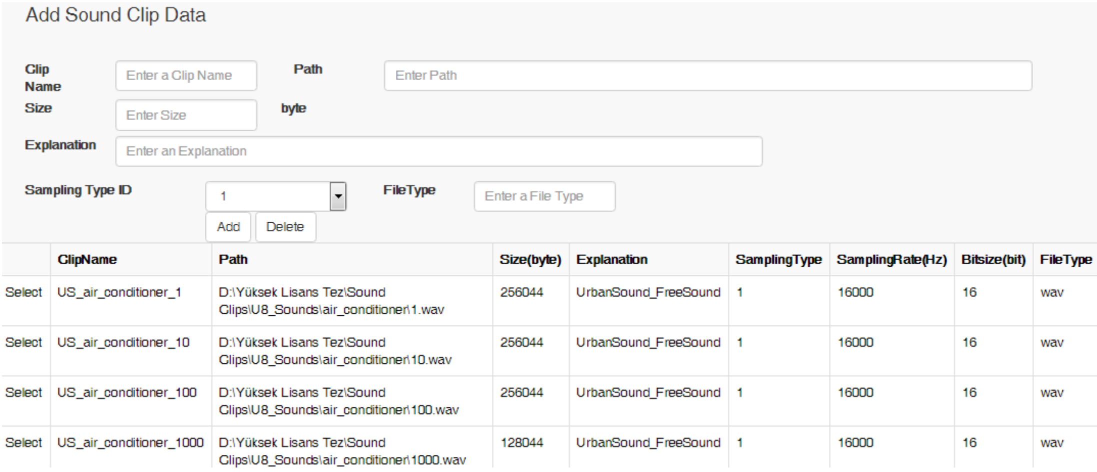
Path: Path of the clip on the server.

Size: Size of clip.

Explanation: Information about the clip like origin of it, recording details etc.

Sampling Type ID: Sampling type of clip. This type is taken from sampling types table.

File Type: File type of clip like wav or mp3 file.



	ClipName	Path	Size(byte)	Explanation	SamplingType	SamplingRate(Hz)	Bitsize(bit)	FileType
Select	US_air_conditioner_1	D:\Yüksek Lisans Tez\Sound Clips\U8_Sounds\air_conditioner\1.wav	256044	UrbanSound_FreeSound	1	16000	16	wav
Select	US_air_conditioner_10	D:\Yüksek Lisans Tez\Sound Clips\U8_Sounds\air_conditioner\10.wav	256044	UrbanSound_FreeSound	1	16000	16	wav
Select	US_air_conditioner_100	D:\Yüksek Lisans Tez\Sound Clips\U8_Sounds\air_conditioner\100.wav	256044	UrbanSound_FreeSound	1	16000	16	wav
Select	US_air_conditioner_1000	D:\Yüksek Lisans Tez\Sound Clips\U8_Sounds\air_conditioner\1000.wav	128044	UrbanSound_FreeSound	1	16000	16	wav

Figure 5.8 Web Application – Edit Sound Clip Data Page

The second sound data that can be edited is environmental sounds. Environmental sound clips should be extracted from sound clips defined in as mentioned above. Each of these sound clips should have a subtype and it should be known from which sound clips it is extracted.

In edit environmental sound data page(Figure5.9), admin can add or new sound data by entering the required features or delete existing sound clips.. These features are:

Sound Clip Name: Name of the environmental sound clip.

Subtype Name: Subtype of the environmental sound like car crash or dog bark.

Clip Name: Name of the clip which the environmental sound is extracted.

Path: Path of the clip on the server.

Sample Size: Size of environmental sound clip.

Start Sample: Starting sample of the sound clip which the environmental sound is extracted.

End Sample: Ending sample of the sound clip which the environmental sound is extracted.

Start Time: Starting time of the sound clip which the environmental sound is extracted.

End Time: Ending time of the sound clip which the environmental sound is extracted.

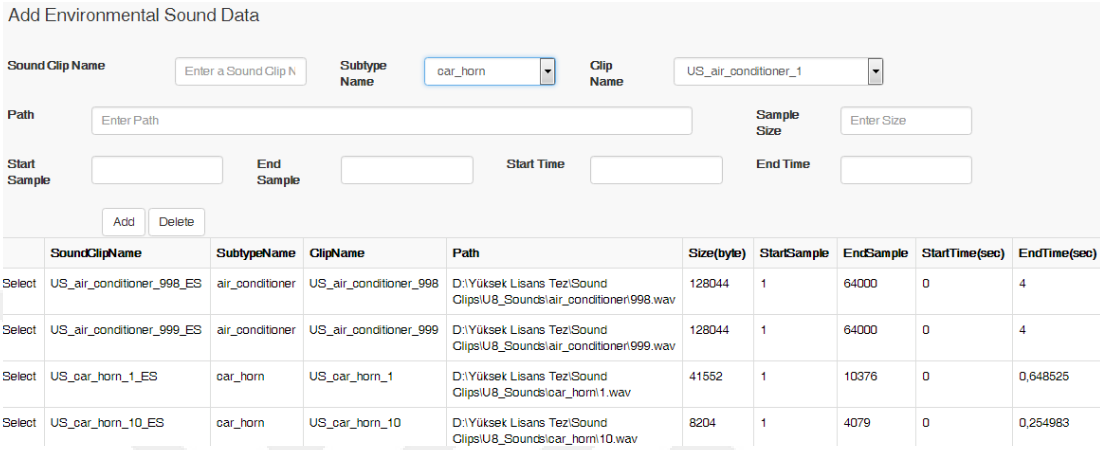


Figure 5.9 Web Application – Edit Environmental Sound Data Page

Last sound data that can be edited is noise clips. Noise clips are background clips like traffic . Each noise clip should have a type.

In edit noise clip data page(Figure5.10), admin can add or new sound data by entering the required features or delete existing noise clips' data.. These features are:

Noise Clip Name: Name of the noise clip.

Noise Type: Type of the noise clip like traffic.

Sampling Type ID: Sampling type of noise clip. This type is taken from sampling types table.

Size: Size of clip.

Path: Path of the clip on the server.

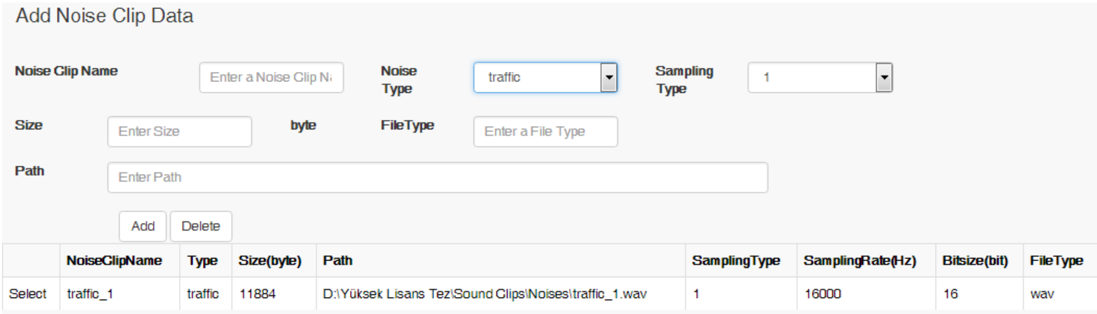


Figure 5.10 Web Application – Edit Noise Clip Data Page

5.2.3 Search Engine

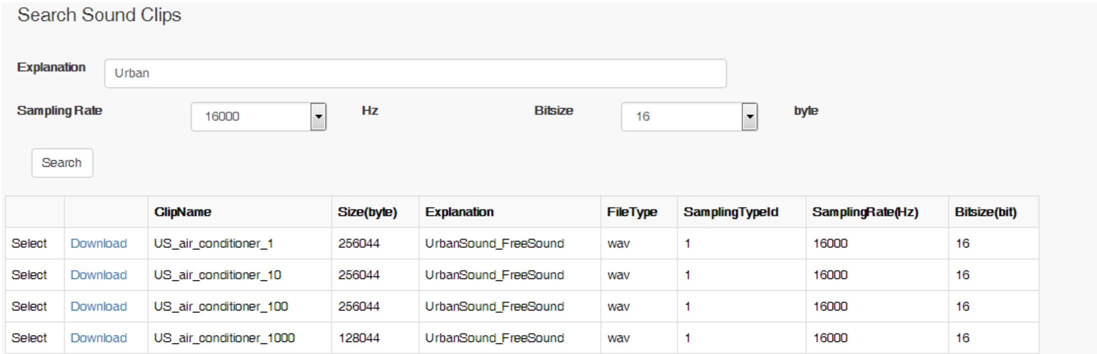
Another tool in the web application is to search the database. End users can use the search feature. There are 3 types of clips that can be searched.

- Sound Clips
- Environmental Sound
- Noise Clips

The clips listed in the search result can be selected and downloaded.

5.2.3.1 Search Sound Clips

Sound clip search can be done using three different filters. These filters are *explanation*, *sampling rate* and *bitsize*. When the filter entries are left blank, the results are filtered according to the other filters without any filtering. That is, when all filters are left empty, all available clips are listed.



Search Sound Clips

Explanation: Urban

Sampling Rate: 16000 Hz   Bitsize: 16 byte

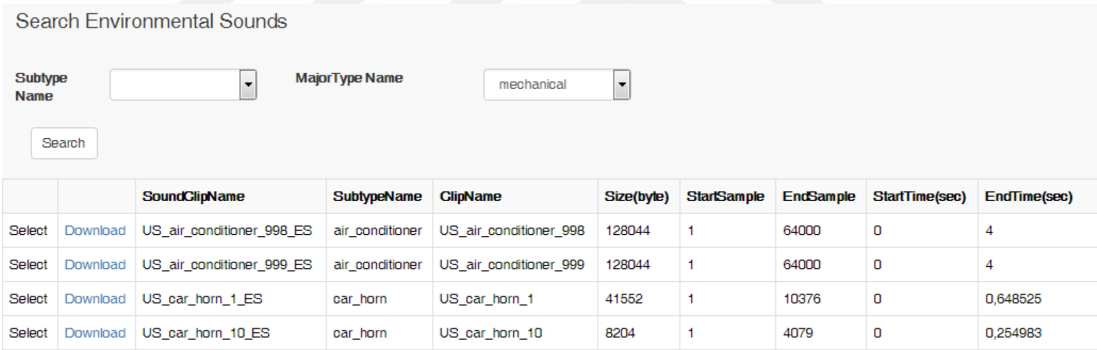
Search

		ClipName	Size(byte)	Explanation	FileType	SamplingTypeId	SamplingRate(Hz)	Bitsize(bit)
Select	Download	US_air_conditioner_1	256044	UrbanSound_FreeSound	wav	1	16000	16
Select	Download	US_air_conditioner_10	256044	UrbanSound_FreeSound	wav	1	16000	16
Select	Download	US_air_conditioner_100	256044	UrbanSound_FreeSound	wav	1	16000	16
Select	Download	US_air_conditioner_1000	128044	UrbanSound_FreeSound	wav	1	16000	16

Figure 5.11 Web Application – Search Sound Clips Page

5.2.3.2 Search Environmental Sounds

Environmental sound search can be done according to the major type or subtype of clips. As an example user can be choose ‘human’ major type and all clips in the subtypes that are the subset of the selected major type are listed like baby crying or children playing.



Search Environmental Sounds

Subtype Name: MajorType Name: mechanical

Search

		SoundClipName	SubtypeName	ClipName	Size(byte)	StartSample	EndSample	StartTime(sec)	EndTime(sec)
Select	Download	US_air_conditioner_998_ES	air_conditioner	US_air_conditioner_998	128044	1	64000	0	4
Select	Download	US_air_conditioner_999_ES	air_conditioner	US_air_conditioner_999	128044	1	64000	0	4
Select	Download	US_car_horn_1_ES	car_horn	US_car_horn_1	41552	1	10376	0	0,648525
Select	Download	US_car_horn_10_ES	car_horn	US_car_horn_10	8204	1	4079	0	0,254983

Figure 5.12 Web Application- Search Environmental Sounds Page

5.2.3.3 Search Noise Clips

Noise clip search can be done using three different filters. These filters are *noise type*, *sampling rate* and *bitsize.*

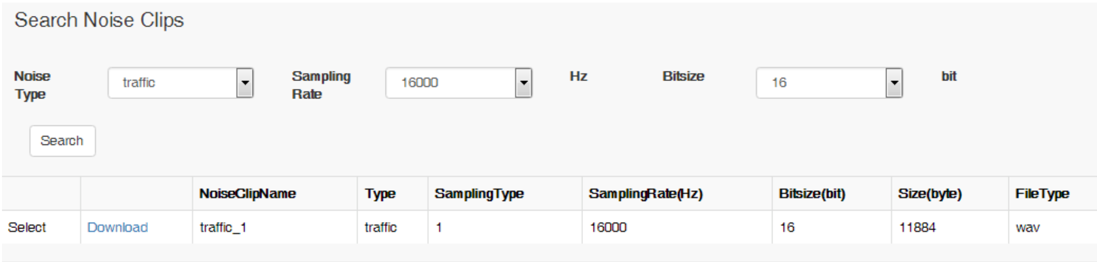


Figure 5.13 Web Application – Search Noise Clips Page

5.2.4 Test

End users can search the database with noise clip type, sound clip type and with SNR interval then select a clip and copy it to their own computer. After finding the start index of the embedded environment sound clip, write to the edit box on the test page and clicking the test button for test their algorithm correctness. The test page for end user tools is seen in Figure 5.14.

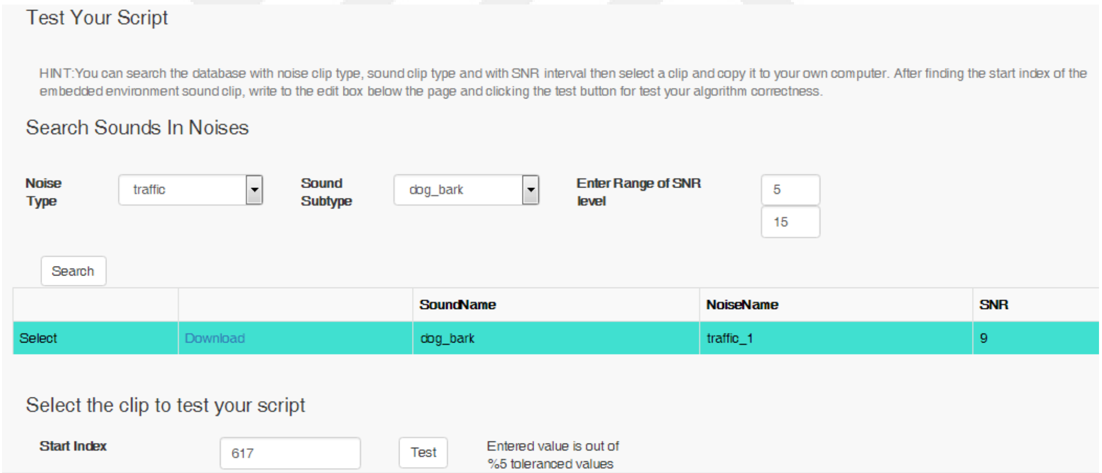


Figure 5.14 Web Application – Test Your Script Page

Chapter 6

CONCLUSION

The lack of a common database for environmental sound recognition is an important obstacle in front of the researchers. The development of and ESR system is a tough process during which someone have to deal with lots of sound clips with different types, algorithms and models. Also, finding and selecting suitable sounds in such databases can be difficult because of the great variety of sounds present, poor documentation, questionable recording quality, and required purchasing costs.

This thesis has focussed explain a relational database model which will make the data handling easier. The database developed is different from other counterparts which are just providing the data. The relational database model described here provides easy maintenance as well as easy usage.

After implementation , environmental sound clips added to database. Some of these sound clips taken from other databases, some of them was recorded by us for this study.

In addition, a computer interface was created for easy use of the created database. Later, it was developed and turned into a web application so that everyone can access it.

This database is intended for ESR purposes, but other areas dealing with environmental sounds can use it.

In this study, a prototype of the target database was created. Also some environmental sound recordings was added to database. In the following studies, this database can be improved by adding more data from commonly used public databases and a common database for ESR research activities can be created by adding more functionality.

REFERENCES

[1] Gygi B, Kidd GR, Watson CS. Similarity and categorization of environmental sounds. Perception and Psychophysics.;69(6):839–855, 2007

[2] M.W.W. van Grootel, T.C. Andringa_, and J.D. Krijnders , DARES-G1: Database of Annotated Real-world Everyday Sounds, Proceedings of the NAG/DAGA Meeting, Rotterdam, 2009

[3] S. Nakamura, K. Hiyane, F. Asano, T. Nishiura, T. Yamada, “Acoustical Sound Database in Real Environments for Sound Scene Understanding and Hands-Free Speech Recognition,” Proceedings of 2nd ICLRE, pp. 965-968, 2000

[4] Y. Arslan, A. Tanış and H. Canbolat , “A Relational Database Model and Tools for Environmental Sound Recognition”, Advances in Science, Technology and Engineering Systems Journal(ASTESJ), vol. 2, no. 6, pp. 145-150, 2017

[5] T. Ahmed, M. Uppal and A. Muhammad, “Improving Efficiency and Realibility of Gunshot Detection Systems”, IEEE, ICASSP 2013

[6] P. Thumwarin, T. Matsuura and K. Yakoompai, "Audio forensics from gunshot for firearm identification", Proc. IEEE 4th Joint International Conference on Information and Communication Technology Electronic and Electrical Engineering Tailand, pp. 1-4, 2014.

[7] S. Chu, S. Narayanan, C.-C. J. Kuo, and M. J. Mataric, “Where am I? Scene recognition for mobile robots using audio features,” in Multimedia and Expo, 2006 IEEE International Conference on. IEEE, pp. 885–888, 2006

[8] N. Yamakawa, T. Takahashi, T. Kitahara, T. Ogata, and H. G. Okuno, “Environmental sound recognition for robot audition using Matching- Pursuit,” in Modern Approaches in Applied Intelligence. Springer, pp. 1–10, 2011

[9] J. Chen, A. H. Kam, J. Zhang, N. Liu, and L. Shue, “Bathroom activity monitoring based on sound,” in Pervasive Computing. Springer, pp. 47–61, 2005

[10] M. Vacher, F. Portet, A. Fleury, and N. Noury, “Challenges in the processing of audio channels for ambient assisted living,” in e-Health Networking Applications and Services (Healthcom), 2010 12th IEEE International Conference on. IEEE, pp. 330–337, 2010

[11] J.-C. Wang, H.-P. Lee, J.-F. Wang, and C.-B. Lin, “Robust environmental sound recognition for home automation,” IEEE Trans. Autom. Sci. Eng., vol. 5, no. 1, pp. 25–31, 2008.

[12] R. Bardeli, D. Wolff, F. Kurth, M. Koch, K.-H. Tauchert, and K.-H. Frommolt “Detecting bird sounds in a complex acoustic environment and application to bioacoustic monitoring,” Pattern Recognition Letters, vol. 31, no. 12, pp. 1524–1534, 2010.

[13] F. Weninger and B. Schuller, “Audio recognition in the wild: static and dynamic classification on a real-world database of animal vocalizations,” in acoustics, speech and signal processing (ICASSP), 2011 IEEE international conference on. IEEE, pp. 337–340, 2011

[14] P. Foggia, A. Saggese, N. Strisciuglio, M. Vento, and N. Petkov. Carcrashes detection by audio analysis in crowded roads. In Advanced Video and Signal Based Surveillance e (AVSS), 2015 12th IEEE International Conference on, pages 1-6, Aug 2015.

[15] J. Rouas, J. Louradour, and S. Ambellouis, "Audio Events Detection in Public Transport Vehicle," Proc. of the 9th International IEEE Conference on Intelligent Transportation Systems, 2006.

[16] R. Radhakrishnan and A. Divakaran, “Systematic acquisition of audio classes for elevator surveillance,” in Image and Video Communications and Processing 2005, vol. 5685 of Proceedings of SPIE, pp. 64–71, March 2005.

[17] P. K. Atrey, N. C. Maddage, and M. S. Kankanhalli, “Audio based event detection for multimedia surveillance,” in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’06), vol. 5, pp. 813–816, Toulouse, France, May 2006.

[18] K. Lupatka, P. Zwan, A. Czyzewski, “Dangerous sound event recognition using support vector machine classifiers”, Advances in Multimedia and Network Information System Technologies Volume 80 of the series Advances in Intelligent and Soft Computing pp 49-57, 2010

[19] A.Okamoto, K.Hayashida, M.Nakayama, T.Nishiura, “A detection ofdangersounds based on variable-state hidden Markov models”, Signal Processing in Acoustics, 2013

[20] S. Chu, S. Narayanan, and C.-C. J. Kuo, “Environmental sound recognition with time– frequency audio features,” IEEE Trans. Audio, Speech, Lang. Process., vol. 17, no. 6, pp. 1142–1158, 2009.

[21] M. Lojka, M. Pleva, E. Kiktova, J. Juhar, A. Cizmar. "Efficient acoustic detector of gunshots and glassbreaking." Multimedia Tools and Applications, 75(17):10441–10469, 2016

[22] Yüksel Arslan and Hüseyin Canbolat, A sound database development for environmental sound recognition, 2017 25th Signal Processing and Communications Applications Conference (SIU) Year: 2017

[23] Brian Gygi and Valeriy Shafiro, Development of the Database for Environmental Sound Research and Aplication (DESRA) :Design, Fuctionality and Retrieval Considerations, EURASIP Journal on Audio, Speech, and Music Processing Volume 2010.

[24] K. Dickerson, J. R Gaston, and S. McCarty-Gibson, Parameterizing Sound: Design Considerations for an Environmental Sound Database, Army Research Labarotory, 2015.

[25] http://www.cs.tut.fi/sgn/arg/dcase2016/

[26] http://www.cs.tut.fi/sgn/arg/dcase2017/

[27] K. J. Piczak, "ESC: Dataset for environmental sound classification", Proceedings of the ACM International Conference on Multimedia, 2015.

[28] Justin Salamon, Christopher Jacoby, Juan Pablo Bello, A Dataset and Taxonomy for Urban Sound Research, Proceedings of the 22nd ACM international conference on Multimedia, November 03-07, 2014.

[29] Dennis, J. W. Sound Event Recognition in Unstructured Environments using Spectrogram Image Processing. Unpublished master's thesis, Nanyang Technological University, 2014.

[30] R. F. Lyon, \Machine Hearing: An Emerging Field," IEEE Signal Processing Magazine, vol. 27, pp. 131{139, Sept. 2010.

[31] Gaver WW. What in the world do we hear?: An ecological approach to auditory event perception. Ecological Psychology.;5(1):1–29, 1993

[32] Ballas JA. Common factors in the identification of an assortment of brief everyday sounds. Journal of Experimental Psychology: Human Perception and Performance.;19(2):250–267,1993

[33] Leila Abdoune and Mohamed Fezari. 2016. Everyday life sounds database: telemonitoring of elderly or disabled. Journal of Intelligent Systems 25, 1, 71—84, (2016)

[34] B. Gygi and V. Shafiro, "Environmental sound research as it stands today," in Proceedings of the Meetings on Acoustics, vol. 1, p. 050002, 2008.

[35] S. Chachada and C.-C. J. Kuo. Environmental sound recognition: A survey. APSIPA Transactions on Signal and Information Processing, 3:e14, 2014.

[36] T. Virtanen and M. Hel´en, "Probabilistic model based similarity measures for audio query-by-example," in Applications of Signal Processing to Audio and Acoustics,IEEE Workshop on. IEEE, 2007, pp. 82– 85, 2007

[37] S. Duan, J. Zhang, P. Roe, and M. Towsey, "A survey of tagging techniques for music, speech and environmental sound," Artificial Intelligence Review, pp. 1–25, 2012.

[38] M. Cristani, M. Bicego, and V. Murino, "Audio-visual event recognition in surveillance video sequences," IEEE Trans. Multimedia, vol. 9, no. 2, pp. 257–267, 2007.

[39] R. Sitte and L. Willets, "Non-speech environmental sound identification for surveillance using self-organizing-maps," in Proceedings of the Fourth conference on IASTED International Conference: Signal Processing, Pattern Recognition, and Applications, ser. SPPR'07. Anaheim, CA, USA: ACTA Press, pp. 281–286, 2007

[40] O. Gencoglu, T. Virtanen, H. Huttunen. (2014). "Recognition of Acoustic Events Using Deep Neural Networks," in European Signal Processing Conference (EUSIPCO 2014).

[41] Cakir, E. Sound Multilabel Sound Event Classification with Neural Networks. Unpublished master's thesis, Tampere University of Technology, 2014.

[42] Gebremedhin, T. A. An Environmental Audio–Based Context Recognition System Using Smartphones. Unpublished master's thesis, University of Twente, 2014.

[43] http://www.auditorylab.org

[44] Warren, W. H., & Verbrugge, R. R. . Auditory perception of breaking and bouncing events. Journal of Experimental Psychology: Human Perception & Performance, 10, 704-712, 1984

[45] Repp, B. H. The sound of two hands clapping. Journal of the Acoustical Society of America, 81, 1100-1109, (1987).

[46] Li, X., Logan, R. J., & Pastore, R. E. Perception of acoustic source characteristics: Walking sounds. Journal of the Acoustical Society of America, 90, 3036-3049, (1991)

[47] Marcell, M. M., Borella, D., Greene, M., Kerr, E., & Rogers, S. Confrontation naming of environmental sounds. Journal of Clinical & Experimental Neuropsychology, 22, 830-864, (2000).

[48] Vanderveer, N. J. Ecological acoustics: Human perception of environmental sounds. Dissertation Abstracts International, 40, 4543B, (1979).

[49] Fabiani, M., Kazmerski, V. A., Cycowicz, Y. M., & Friedman, D. Naming norms for brief environmental sounds: Effects of age and dementia. Psychophysiology, 33, 462-475, (1996).

[50] Gygi, B., Kidd, G. R., & Watson, C. S. Spectral-temporal factors in the identification of environmental sounds. Journal of the Acoustical Society of America, 155, 1252-1265, (2004).

[51] Shafiro, V. Perceiving the sources of environmental sounds with a varying number of spectral channels. Unpublished doctoral dissertation, City University of New York, (2004).

[52] Fowler, C. A. Sound-producing sources as objects of perception: Rate normalization and nonspeech perception. Journal of the Acoustical Society of America, 88, 1236-1249, (1990).

[53] Ballas, J. A., & Howard, J. H. Interpreting the language of environmental sounds. Environment & Behavior, 19, 91-114, (1987).

[54] Ellis, D. P. Prediction-driven computational auditory scene analysis. Unpublished doctoral dissertation, MIT, Cambridge, MA, (1996).

[55] Heller, L. M., & Wolf, L. When sound effects are better than the real thing. Journal of the Acoustical Society of America, 111, 2339, (2002).

[56] A. L. Brown, J. Kang, and T. Gjestland. Towards standardization in soundscape preference assessment. Applied Acoustics, 72(6):387-392, 2011.

[57] http://serv.cusp.nyu.edu/projects/urbansounddataset/

CURRICULUM VITAE

PERSONAL INFORMATION

Name Surname : Abdussamet Tanış

Date of Birth : 01/04/1992

Phone : 05376971979

E-mail : abdussamettanis@gmail.com

EDUCATION

High School : Özel Ankara Öncü College

Bachelor : Sakarya University Electrical and Electronical Engineering

Master Degree :

WORK EXPERIENCE

Research and Development Engineer : Eyza Teknoloji – (2015 – ongoing)

TOPICS OF INTEREST

- Microprocessors
- Programming
- Digital Signal Processing