International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT2173163

Infant Monitoring System for Deaf Parents

Akshay Mohite*, Prof. Dhiraj Jadhav
Department of Computer Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, India

Article Info
Volume 7, Issue 3
Page Number: 527-534

Publication Issue :
May-June-2021

Article History
Accepted : 12 June 2021
Published : 22 June 2021

ABSTRACT

Infant Monitoring System for Deaf Parents that aims to convert Baby crying sounds which are not distinguishable by deaf people into recognizable alerts. The project proposes an intelligent system which is capable of identifying the cry of baby using mic and send the alerting messages directly to the concern person's mobile phone number using GSM Modem. This project makes use of a mic, GSM Modem, and Raspberry Pi processor. Raspberry Pi processor automatically reads the crying sound of the baby using mic, which is interfaced to the USB driver of the Raspberry Pi, after audio file is recorded it is processed with ANN, filtered with MFCC then the processor will detect sound consist of baby cry. If baby cry detected then predefined alert message will be sent to the predefined phone number that is stored in SIM card via GSM modem.

Keywords : Raspberry Pi, classification, ANN, MFCC, GSM, SMS, and LED.

I. INTRODUCTION

The hearing sense is one of the five senses essential for human life. To acquire language, learn how to speak, and gain general knowledge, the hearing sense is required. Mainly, due to hearing, individuals can communicate and socialize with the world around them. However, deaf people cannot communicate or socialize effectively with normal people around them because of their hearing loss which affects many aspects of their lives. Overall parenting is a challenging responsibility. It becomes crucial if parents are either hearing impaired or visually impaired. Due to this, many vital signs of infant caring are unnoticed this is one of the reasons of increase infant mortality. This infant caring can be improvised by infant monitoring by generating alerts for the hearing-impaired parents.

In India, "hearing handicapped" as defined by the Rehabilitation Council of India Act, 1992, is – "hearing impairment of 70 dB and above, in better ear or total loss of hearing in both ears" [1]. According to National Association of the Deaf, there are 18 Million Deaf Individuals in India. In the past decade, infant mortality rate in India has risen considerably. The main cause observed was Sudden Infant Death Syndrome. This was worse in the case of blind and deaf parents as at times many vital signs like Respiration, Blood pressure, Temperature, Heart rate of the infants left unnoticed. Because of the hearing impairment, they used to co-sleep with their baby(s) or tie a string around their wrist to the baby's wrist. The deaf parent will wake up knowing that the baby is either moving or crying if the baby moved. Some deaf parents put their arm or leg near the baby in the cradle. Some deaf parents would have a trained hearing dog that would alert them of the baby crying [2]. However, these ways do not help in many cases and it would be life threatening to infant also. Many deaf parents try to treat their hearing impairment to take care of their baby in way that is more alert. Hearing impairment treatments include hearing aids and cochlear implants [3].

The objective of hearing aids is to enhance the hearing sense through sound amplification [4]. Although hearing aids are satisfactory in many cases, they fail to help people with severe or profound hearing weakness. On the other hand, cochlear implants are small electronic devices that surgically placed under the skin behind the ears [5]. Cochlear implants are more effective than hearing aids because they are capable of providing a sense of hearing for people who suffer from sever or profound deafness. Hearing aids provide very less improvement in case of severe deafness and cochlear implants are expensive and have some medical issues like bleeding, infection, device malfunction, facial nerve weakness, ringing in the ear, dizziness [6].

This paper proposes an Infant monitoring system for deaf parent that helps deaf people to get appropriate alerts in various situations. For example, when deaf parents are not around the infant and infant starts crying then they will get SMS alerts on their phone their phone will get vibrated. In case if they don't have phone with them they will get alerted using LED's which are integrated in the system. Proposed system is Audio recognition system capable of distinguishing normal alerts baby crying, doorbell ringing, TV noise, baby laughing, other household noises etc. Next, proposed system converts these sounds into recognizable alert that the deaf person can feel such as mobile phone vibration.

To reduce infant mortality in case of deaf parents an Infant monitoring system is proposed.

II. LITERATURE SURVEY

In [7], Aktas et al. introduced a real-time infant monitoring system for hard of hearing parents, consisting of sensors (finger heartbeat, body temperature, humidity and sound detection), a microcontroller and android-based mobile devices (smartphone and smartwatch). They proposed a system that monitors both physiological data collected from infants and creates alarms for abnormal conditions. The designed and implemented system, developed on the Arduino Leonardo board, have been used in order to collect data from the sensors and to create alarm cases by evaluating these data. For notification of alarms to the parent, they used low-cost vibrating smartwatch compatible with android-based smartphones. From the implementation results, they observed that the data collected from the sensors are monitored in real-time and based on abnormal conditions a related alert is notified successfully [7].

Nair et al. designed IOT based infant monitoring system consists of mainly two PCB boards and an IOT platform. The first PCB board is attached to the infant's body while the other PCB Board is attached to the blind and deaf parent's wrist with a vibrator module in it. The PCB board related to infant has four sensors as LM 35, pulse sensor, activity sensor MPU 6050 and sound sensor CZN 15E. It also includes Arduino Uno and a NodeMCU module. All the sensor values are collected by the Arduino board and is send over to NodeMCU module. The NodeMCU is calibrated with threshold of temperature, pulse, sound and axis of motion. All these values are send to the Blynk IOT platform where the medical surveillance monitors the values and sends an alert signal to the parent in case of emergency. They used various technological resources in an integrated manner to detect the various health issues at an early stage and provide alerts in the form of vibrations to the disabled parents. The vital signs of an infant such as temperature, pulse, sound and motion are measured and recorded using various sensors and the data obtained are continuously logged into the cloud website, so doctors can access system database anytime. IOT platforms following MQ Telemetry Transport (MQTT) protocol is selected to develop an automated alerting system without the help of a medical assistance [8].

Ramzi R. Saifan implemented a Deaf Assistance Digital System (DADS) for speech and sound recognition. DADS helps deaf people to live safely and become actively involved in the society. Moreover, DADS can rescue deaf people from a lot of dangers that surround them such as penetration, theft accidents, and car and fire accidents. DADS achieves its objectives by converting the sonic alarms that cannot be heard by deaf people (e.g., car horn or cautionary words like "be careful") into appropriate alerts that are recognizable by deaf people (e.g., vibration and pictures). DADS consists of two engines, namely, speech recognition engine and sound recognition engine. The speech engine is responsible for recognizing critical cautionary words and is implemented using Soundex algorithm which achieves up to 85% accuracy for one occurrence and up to 100% with three occurrences. On the other hand, the sound engine is responsible for recognizing four categories of sound (i.e., vehicle, baby, normal, and loud sounds). In this paper, they investigated three approaches to implement the sound engine classification, namely, K-Nearest Neighbour (KNN), Neural Network (NN) and Decision Tree. Our experimental results showed that using a three layer NN with 60 nodes in the hidden layer achieves the highest classification accuracy of 91.7%. Hence, DADS implementation leverages NN classification in the sound engine [9].

It is observed that following are the major alerting conditions of an infant which requires attention.

III. PRAPOSED METHOD AND MATERIALS

In this paper to identify baby cry sound with respect to surrounding sounds can be done by extracting features by Mel-Frequency-cepstral-coefficient (MFCC) and sound classification using Artificial Neural Network (ANN).

A. Building a Dataset:

First step was getting a useable dataset to building a dataset. Proposed system used the Donate-A-Cry corpus for positive samples of sounds of babies crying. The dataset has nearly 1000 sound clips of 7 seconds in length each. The corpus didn't have negative samples, so I used the Environmental-Sound-Classification (ESC50) dataset which has about ~2000 samples of 5 secs length. As the ECS50 dataset contains sounds of babies crying, we have to remove those clips from the negative samples. Divide the dataset into train and validation (90:10).

B. Feature Extraction:

The process of forming speech signals starting from the larynx (where the vocal cords located) and ended up in the mouth. Speech or voice signals categorized into voiced and unvoiced. Unvoiced is a condition where the state of the vocal cords do not vibrate. Voiced is a condition where the state of the vocal cords vibrate and produce a pulse of the glottis. Pitch is known as the fundamental frequency of the glottis. When we speak we generate multiple sound waves with multiple frequencies simultaneously [10].

Time-domain representation of the speech signal shows us the loudness (amplitude) of the sound wave changing with time. The silence is represented by

Amplitude 0. (From the definition of the sound waves- This amplitude is nothing but the amplitude of air particles which are oscillating due to the pressure change in the atmosphere because of sound.

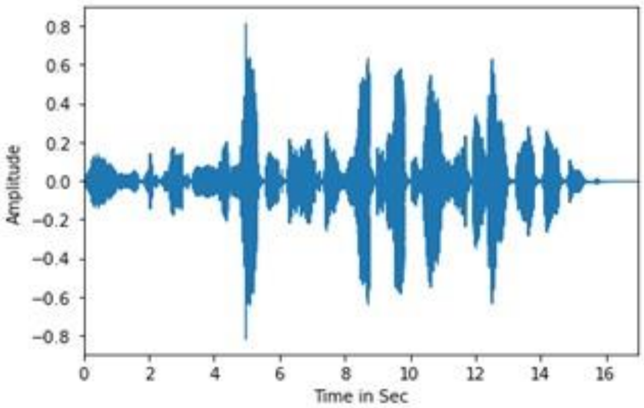

Figure 1. Baby Cry Signal

As they only talk about the loudness of audio recording, these amplitudes are **not very informative**. To better understand the audio signal, it is necessary to transform it into the **frequency-domain. The frequency-domain** representation of a signal is used to find out what different frequencies are present in the signal. Fourier Transform is a mathematical concept that can convert a continuous signal from time-domain to frequency-domain. It also gives the magnitude of the each frequency present in the signal [11].

$$X(k) = \sum_{n=0}^{N-1} x(n).\, e^{-i2\pi kn/N} \quad \ldots\ldots (1)$$

$$x(n) = \sum_{k=0}^{N-1} X(k).\, e^{i2\pi kn/N} \quad \ldots\ldots (2)$$

Where N is number of time samples we have. n is current sample we're considering (0 .. N-1). x(n) is value of the signal at time n k is current frequency we're considering (0 Hertz up to N-1 Hertz). X_k is amount of frequency k in the signal (amplitude and phase, a complex number).

Figure 2. Speech Recognition System

Speech recognition mainly focuses on training the system to recognize an individual's unique voice characteristics. Mel Frequency Cepstral Coefficients called MFCC is the most popular feature extraction technique as it is less complex in terms of implementation and more effective and robust under various conditions [12]. MFCC is designed based on the knowledge of human auditory system. It is a standard method for feature extraction in speech recognition. MFCC involves steps such as Pre-emphasis, Framing, Windowing, FFT, Mel filter bank, computing DCT.

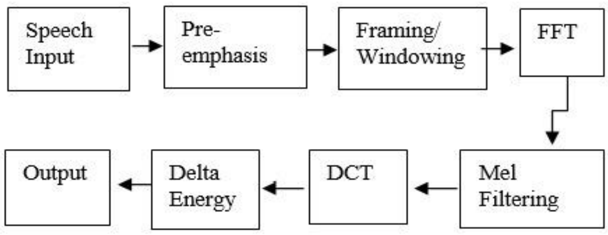

Figure 3. MFCC Block Diagram

Pre-emphasis increases energy of the signal at higher frequencies. Frame blocking is the process where voice signal is segmented into small duration block of 20-30 ms known as frames. Speech is time varying signal so framing is required. After that every frame is multiplied by Hamming window in order to keep continuity of signal. Window is used for decrease spectral distortion.

$$Y(n) = X(n) * W(n) \quad \ldots\ldots (3)$$

Where W (n) is the window function.

Fast Fourier Transform (FFT) is a mathematical algorithm which calculate Discrete Fourier Transform (DFT) of a given signal. Difference between Fourier Transform (FT) and FFT is that FT considers continuous signal where as FFT considers Discrete signal. FFT converts time domain signal into frequency domain signal.

In Mel filtering process we magnitude frequency response is multiplied by a set of 20 triangular band pass filters in order to get smooth magnitude spectrum. The size of features involved also get reduces. DCT (Discrete Cosine transform) extract signals main information and peak. We apply DCT on the 20 log energy $E_k$obtained from the triangular band pass filters to have L Mel-scale cepstral coefficients. DCT formula is shown below

$$C(m) = \sum_{k=1}^{N} cos\left(\frac{m(k-0.5)\pi}{N}\right) E(k) \quad \ldots. m = 1, 2, \ldots L \qquad \ldots\ldots. (4)$$

Where N = number of triangular band pass filters, L = number of Mel-scale cepstral coefficients. And N=20 and L=12. The frequency domain is transformed into a time-like domain called quefrency domain by using DCT. These features are called as the Mel-scale cepstral coefficients. For speech recognition, we can use MFCC alone but for better performance, we can add the log energy and can perform delta operation as well. Using delta energy the energy within the frame can be calculated. At the end we will get MFCC features of given speech signal [13].

Mel-Filter bank of input signal is as shown in below Figure 4.

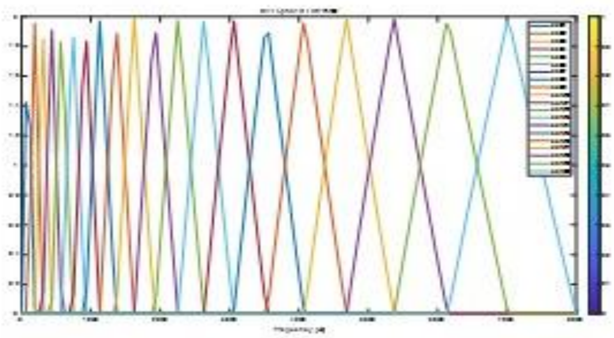

Figure 4. Mel-Filter bank of input signal

C. Classification:

Artificial Neural Network (ANN):

Artificial Neural Networks (ANNs) developed in 1940 on the basis of the workings of the human brain with numerical logic adapted to computer equipment. ANN is able to recognize activities based on data. Also it has the ability to make decisions about data that has not been studied. ANN shows great advantages in sound recognition process. The structure of the ANN is shown in Figure 5.

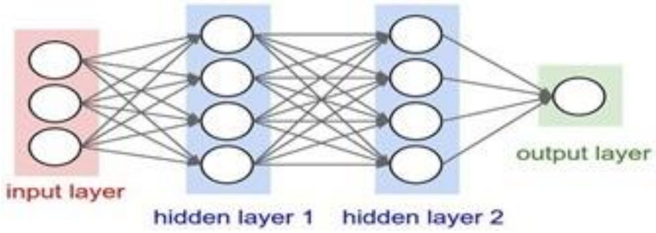

Figure 5. Structure of ANN

It has mainly two stages: feedforward and backpropagation. Through backpropagation we update the weights and bias of each other. By adjusting weights it learns what signal is important what not. Sound recognition needs less layers than the image recognition, typically 3-4 layers are enough. Weights are randomly initialized to small numbers close to 0. Then we input our observation from features extracted to input layer. Then forward propagation takes place until we get final prediction.

$$C = \frac{1}{2}\sum(y - Y)^2 \qquad \ldots\ldots (5)$$

C is cost function, y is predicted value, and Y is actual value. After gating predicted value, it will be compared with actual value and measure the error, it is called as cost function. Then error will be back propagated and the weights are updated until we get smallest cost function. When the whole training set is passed through the ANN it will be called as epoch.

D. Deployment:

The detailed block diagram of baby cry detection with Artificial Neural Network (ANN) by using Raspberry Pi is as shown in Figure.6.

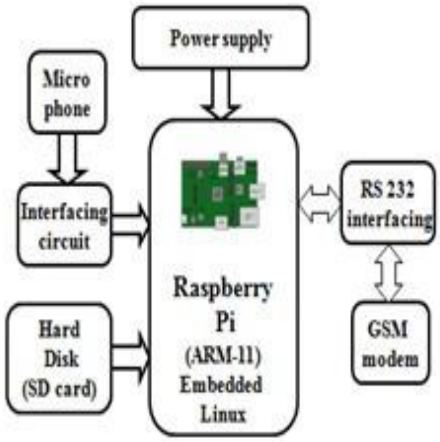

Figure 6. Block Diagram

The Raspberry pi is developed in UK by Raspberry pi foundation. It is credit card sized single board computer. The Raspberry Pi has a Broadcom BCM2835 system on a chip (SoC), which includes an ARM1176JZF - S700 MHz processor, Video Core IV GPU, and was originally shipped with 256 megabytes of RAM, later upgraded to 512 MB. It does not include a built-in hard disk or solid state drive, but uses an SD card for booting and long-term storage.

In this project GSM modem and mic is interfaced with ARM1176JZF – S processor. Input has taken from mic and fed to the Raspberry pi processor. If baby cry detected in input sound then SMS alert will be sent to concern parent also LED which connected to the system will be lightened to alert parent in case if they don't have phone with them.

IV. RESULTS AND DISCUSSION

According to the simulations, the time domain and frequency domain representation of the input signal is as shown in Figure 7 and Figure 8 respectively.

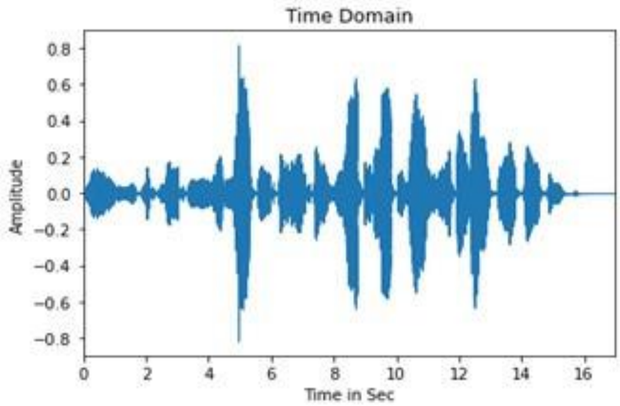

Figure 7. Input signal in Time Domain

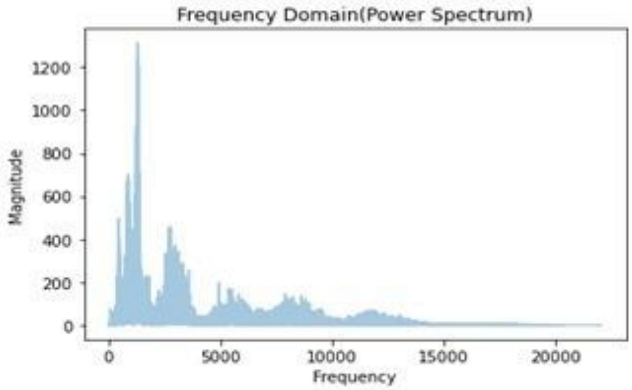

Figure 8. Input signal in Frequency Domain

Power spectrum modified through Mel ceptral filter:

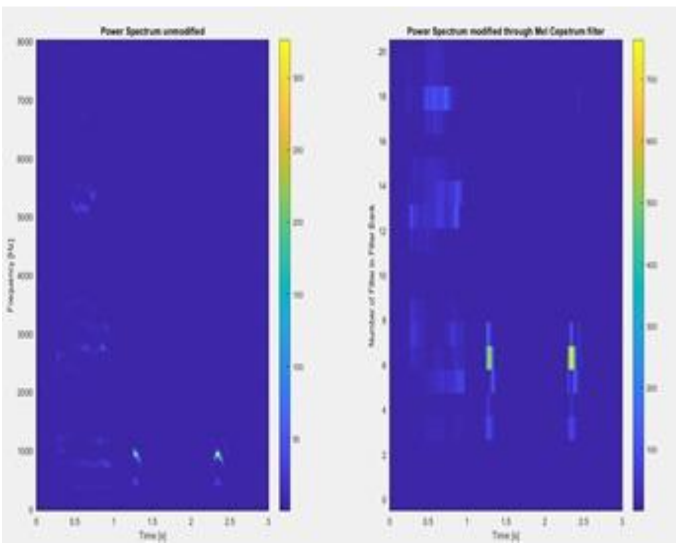

Figure 9. Modified Power Spectrum

The dataset used contains 2000 samples. The Frequency domain is shown in Fig. 8. The Power spectrum modified by Mel ceptral filter is shown in Fig. 9. Total 20 features are extracted by MFCC for each input signal. The accuracy obtained by trial and error is 85% using Artificial Neural Network.

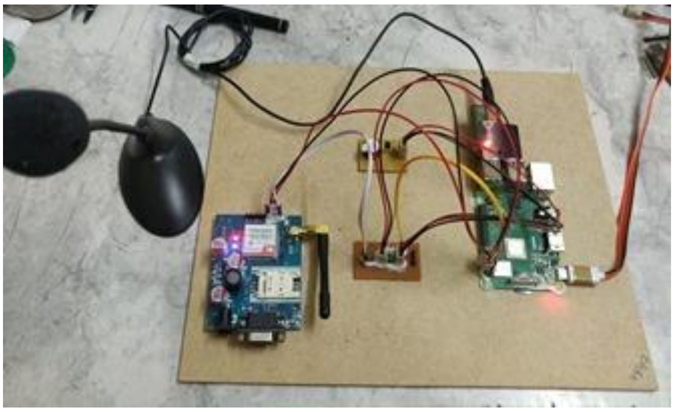

Figure 10. System PCB

Figure 11. Alerted Message

During the system implementation screenshots of system was taken. Alert message sent by hardware have been captured. Test results are given in Figure 10 and Figure 11, it shows successful implementation of the system.

V. CONCLUSION

Crying is one of the major means of infants to communicate with their surroundings, and it is intended to point out any distress or attachment needs to their caregivers. From every-day applications to academic research, Automatic detection of a baby cry in audio signals can be used for various purposes. In this project, a deep-learning based algorithm for automatic detection of baby cry in domestic audio recordings is developed. The algorithm, based on a simple Artificial Neural Network (ANN) algorithm representation of the recordings.

VI. FUTURE SCOPE

This project "Infant Monitoring System for deaf parents" is mainly intended to design a system which is capable of identifying the cry of baby using mic and send the alerting messages directly to the concern person's mobile phone number using GSM Modem.

This technology can be implemented at homes, day care centre, schools, hospitals etc… This project make use of a mic, GSM Modem, and Raspberry Pi processor. User can send the SMS messages to the modem that is connected to the ARM1176JZF-S 700 MHz processor. The ARM1176JZF-S 700 MHz processor automatically reads the crying sound of the baby using mic and the predefined alert message is sent to the predefined phone number that is stored in SIM card via GSM modem.

This project can be extended using high efficiency GPS receiver and a GSM module. The GSM module gives the intimation of the baby cry details to the predefined number through SMS along with location.

The GPS module can also give the location of the baby details in case of emergencies. Camera can be added through which the live video streaming can be monitored when the baby cries being detected using wireless Wi-Fi technology.

VII. REFERENCES

[1]. The Rehabilitation Council of India Act, 1992, Ministry of Law, Justice and Company Affairs (1992): (No.34 of 1992), NewDelhi, 1992. http://www.rehabcouncil.nic.in/engweb/rciact.pdf [Last accessed on 2021 Feb].

[2]. “How do deaf parents hear their baby cry?” Browsed from https://www.rehabcouncil.nic.in/engweb/rciact.pdf. [Last accessed on 2021 Feb]

[3]. Jürgens, T., Clark, N. R., Lecluyse, W., & Meddis, R., “Exploration of a physiologically-inspired hearing-aid algorithm using a computer model mimicking impaired hearing”, International Journal of audiology, 55(6), 346-357, 2016.

[4]. Cox, R. M., & Alexander, G. C. “The abbreviated profile of hearing aid benefit” Ear and hearing, 16(2), 176-186, 1995.

[5]. Balkany, T. J., Hodges, A. V., Eshraghi, A. A., Butts, S., Bricker, K., Lingvai, J., & King, J. “Cochlear implants in children--a review”, Acta oto-laryngologica, 122(4), 356-362, 2002.

[6]. “What to expect from Cochlear implant” Browsed from https://stanfordhealthcare.org/medical-treatments/c/cochlear-implants/what-toexpect.html Dated 13th Jan 2021.

[7]. Aktas, F., Kavus, E., & Kavus, Y. "A Real-Time Infant Health Monitoring System for Hard of Hearing Parents by using Android-based Mobil Devices” Istanbul University-Journal of Electrical & Electronics Engineering, 17(1), 3107-3113,2017.

[8]. Nair, A. R., John, E. M., Menon, L. R., Teresa, N., & Unnikrishnan, L. “IoT Based Infant Monitoring System”

[9]. Saifan, R. R., Dweik, W., & Abdel-Majeed, M. “A machine learning based deaf assistance digital system”, Computer Applications in Engineering Education, 26(4), 1008-1019, 2018.

[10]. H.Subramanian, “Audio Signal Classification,” M. Tech Credit Semin. Rep., Pp. 1–17, 2004.

[11]. “Understanding Audio data, Fourier Transform, FFT and Spectrogram features for a Speech Recognition System” browsed from https://towardsdatascience.com/understanding-audio-data-fourier-transform-fft-spectrogram-and-speech-recognition-a4072d22852011

[12]. C.Poonkuzhali, R. Karthiprakash, S.Valarmathy and M.Kalamani, “An Approach to feature selection algorithm based on Ant Colony Optimization for Automatic Speech Recognition”, International journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 11(2), and 2013.

[13]. Parwinder Pal Singh, Pushpa Rani, "An Approach to Extract Feature using MFCC", International Organization of Scientific Research, Vol.04,pp. 21-25,August 2014.

Cite this article as :

Akshay Mohite, Prof. Dhiraj Jadhav, "Infant Monitoring System for Deaf Parents", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7, Issue 3, pp.527-534, May-June-2021. Available at doi : https://doi.org/10.32628/CSEIT2173163
Journal URL : https://ijsrcseit.com/CSEIT2173163