Journal of Advanced Research in Signal Processing and Applications
Volume 2, Issue 2 - 2020, Pg. No. 8-10
Peer Reviewed Journal

Article

Cry Classifiacation

Drishti Bhati

Department of Computer Science, Global Institute of Technology, Jaipur, India.

I N F O

E-mail Id:
Drishtibahti786@gmail.com
Orcid Id:
https://orcid.org/0000-0001-6532-2834
How to cite this article:
Bhati D. Cry Classifiacation. *J Adv Res Sig Proc Appl* 2020; 2(2): 8-10.

Date of Submission: 2020-05-04
Date of Acceptance: 2020-05-15

A B S T R A C T

Infant cry classification is the collection of unstructured data. Unstructured data is the data which do not have any pre-defined structure such as audio, image and video. Dealing with audio is way more difficult, challenging and time consuming. Infant cry classification is use to identify the voice of child who is newly born. Cry of babies are so unpredicted that just with voice it is not possible to tell the reason behind their cries. It identify cry voice to detect the information regarding the state of an infant, such as hunger, pain or uncomfortable. Cries play an important role when children unable to speak, at that time their cry audio tells about their activity. In this paper method which used to find the reason of cry classification is Short Time Fourier Transform (STFT). This research progresses a system to categorise the infants cry sound by performing feature extraction using MFCC (Mel-Frequency Cepstral Coefficients). The coefficients are then passed on to Back Propagation Neural Network (BNN) with 8 neurons in the hidden layer, for classification of the audio samples into three class namely-hungry, anger and fear.

Keywords: Short Time Fourier Transform, Mel-Frequency Cepstral Coefficients, Back Propagation Neural Network

Introduction

Big data, being heterogeneous in nature, does not circumscribe bounds of data; it comprehends unstructured data which comprises of audios, texts, videos and so forth. Unstructured data is the culmination of myriad forms of info that whichever does not have a pre-defined data model nor it is defined in a predefined manner.[1] This amount to ambivalence and incertitude that makes it onerous to comprehend using conventional programs, that is opposed to data stored in fielded form in databases or annotated documents. The efforts and research analysis of past 30 years has been the subject of multitudinous in the infant cry classification. The thought that the information pertaining to the state of infant cry can be determined by their beseeching walls, as infant cry is the first and only factor with its caregiving environment. It is especially made for those parents who have little to no experience in parenting. It can also be used for the purpose of baby-sitting.

In such modern and high class technology era there was the need to develop an automated and non-invasive platform to look after the infants in resulting to satisfy their needs with the help of automation. The system will incorporate pre-processing by frame blocking and windowing followed by MFCC (Mel-Frequency Cepstral Coefficient) resultant being forwarded to BNN (Back propagation Neural Network)[2] and the final output being differentiated on the basis of voice type and will be classifying infant crying processed sounds in three categories: hungry, angry and fear.

There is more than one form in which the infant cry can be classified. They can cry due to the fire fear, thunder sound and also can cry due to the sound of other children crying. Sometimes heavy or strange voice of adults too can make them cry. Listening to such voices can increase their heart beat and can result them in crying. Sometimes situation can go otherwise too resulting in decrease in heart beat making it more dangerous to infant's health.[3]

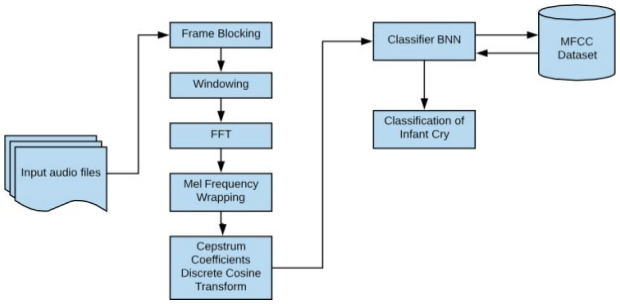
Keeping in mind about such unstructured data that it can get analysed the system is designed. Being the most interesting topic of research and experiment to provide automated result of infant cry we have made a system that tries to listen to the cry sound of infants and process the infant cry using BNN, providing the result with the reason of crying classified in three classes: hungry, angry and fear trying to show with max accuracy.[4]

Methodology

The proposed system consists of the Input audio files which are in. wav format which are the infant cry sounds. These cry sounds are then sent for pre-processing which involves the frame blocking and windowing. The frame blocking is the process of breaking down the signal into frames. Signal analysis is done on short span periods of time which are called frames to get the accurate values. Typical values for frame blocking are dividing the frames into 256 samples and taking the overlapping length as 100. This signal is then passed onto for windowing. Windowing is basically done in order to reduce the distortion of the signal. Finally the signal is then multiplied with the hamming windowing function and this windowing signal serves as an input to calculate the Mel Frequency Cepstral coefficients. The MFCC are calculated as follows:[6]

Algorithm :- Algorithm for Mfcc

- Take the Fourier transform of (a windowed excerpt of a signal).
- Map the powers of the spectrum obtained above onto the mel scale, using triangular over- lapping windows.
- Take the logs of the powers at each of the mel frequencies.



- Take the discrete cosine transform of the list of mel log powers, as if it were a signal
- The MFCCs are the amplitudes of the resulting spectrum.

The coefficients of MFCC are then normalized and sent to Back Propagation for further classification. The BNN takes the MFCC as the inputs and does the classification of the cries. More the number of coefficients, better the system will be able to classify the data. BNN comes under supervised learning algorithm, and generally consists of one input, hidden and output layer. The use of the neurons number in the hidden layer also affects the number of iterations at the training process. The output neurons will depict the classes and hence the infant cry sounds can be classified into three classes namely hungry, fear and anger.[5]

The proposed system is real time system which directly puts the audio which you want to test in the testing data itself, the only restriction being that the audio file should be in a .wav format. The audio sampled then goes through pre-processing, and after obtaining the windowed signal, the Mel-Frequency Cepstral coefficients are obtained. We have used 13 Mfcc coefficients in our system. An audio sample consists of various Mfcc feature vectors.[7]

Result

The proposed system is real time system which directly puts the audio which you want to test in the testing data itself, the only restriction being that the audio file should be in a .wav format. The audio sample then goes through preprocessing, and after obtaining the windowed signal,[8] the Mel-Frequency Cepstral coefficients are obtained. We have used 13 Mfcc coefficients in our system. An audio sample consists of various Mfcc feature vectors. *A script is run to get the feature vectors comprising the coefficients. The feature vectors (training data) obtained are then attached with the class indicator, 0 being for Hungry, 1 for Anger and 2 for Fear. Finally, our Back Propagation* algorithm generates the weights which are updated and then testing data is tested for giving us the desired results.[9-11]

Conclusion

In lieu of the current modus operand of classifying crying sounds, a methodology which is capable of classifying baby crying sound events, conferring to the pathological status of the infant for helping paediatricians during the diagnosis process, has been designed. We have presented the enlargement and application of Back Propagation Neural Networks and MFCC for the classification and discrimination between three types of infant cries: hungry, anger and fear. It can be concluded that, the ability system to do classification of infants cry sound is directly proportional to the number of coefficients of the MFCC.

References

1. Unravelling Unstructured Data: A Wealth of Information in Big Data..Amity Institute of Information Technology Amity University Uttar Pradesh, Noida, India, IEEE.
2. Agrawal M. The infant's cry in health and disease. *The National Medical Journal of India*.
3. Ioana-Alina Bnic, Cucu H, Buzo A. Drago Burileanu and Corneliu Burileanu Automatic Methods for Infant Cry Classification. *IEEE* 2016.
4. Chuan-Yu Chang, Jia-Jing Li. Application of Deep Learning for Recognizing Infant Cries. National Yunlin

University of Science Technology, Taiwan. *IEEE* 2016.
5. Thesis on Infant Cry Detection, University of Miami, 2016.
6. Rosita YD. Hartarto Junaedi Infants Cry Sound Classification using Mel-Frequency Cepstrum Coefficients Feature Extraction and Backpropagation Neural Network.
7. Pal P, Iyer AN, Robert E. Yantorno Emotion Detection from Infant Facial Expressions And Cries. 2015.
8. Petroni M, Malowany AS, Johnston CC et al. Stevens 3 Classification of Infant cry Vocalizations using Artificial Neural Networks (ANNs). Department of Electrical Engineering and Center for Intelligent Machines (CIM), McGill University.
9. Zhou X, Wang J, Hu H et al. Recognition of Infant's Emotions and Needs from Speech Signals.
10. Audio Pattern Recognition of Baby Crying sound events. 2015.
11. Muda L, Begam M, Elamvazuthi I. Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques.