International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 03 | Mar -2020 ISSN: 2582-3930

Infant Cry Analysis and Detection by using Feature Extraction Method

Vidhyalakshmi.R M.E[1], Priya Sharmini.M.K[2], Saraswathi.G [3], Kumaran.R[4]
[1] *Assistant Professor/Department of Computer Science and Engineering*
[2, 3, 4] *UG Student/Department of Computer Science and Engineering*
Adhiyamaan College of Engineering,Hosur 635109

Abstract—Recently, plenty of research has been directed towards natural language processing. However, the baby's cry, which serves as the first means of communication for infants, has not yet been extensively explored, because it's not a language that can be easily understood. Since cry signals carry information about a babies' wellbeing and will be understood by experienced parents and experts to an extent, recognition and analysis of an infant's cry isn't only possible, but also has profound medical and societal applications. During this paper, we obtain and analyze audio features of infant cry signals in time and frequency domains .Based on the related features, we can be able to classify given cry signals to specific cry meanings for cry language recognition. Features extracted from audio feature space include linear predictive coding (LPC), linear predictive cepstral coefficients (LPCC), Bark frequency cepstral coefficients (BFCC), and Mel frequency cepstral coefficients (MFCC). Compressed sensing technique was used for classification and practical data were accustomed design and verify the proposed approaches. Experiments show that the proposed infant cry recognition approaches offer accurate and promising results.

I. INTRODUCTION

CRYING is that the primary means of communication for infants. Experts, including experienced parents, pediatricians and child care specialists, can often distinguish infant cries though training and skill [1]. However, it's difficult for new parents and inexperienced pediatricians and caregivers to interpret infant cries. Hence, differentiating cries with various meanings supported related cry audio features is of great importance [1], [2].Prior works on infant cry analysis have either investigated the difference between normal and pathological (deaf or hearing disabled infants) cries, or they have attempted to differentiate conditional cries like pain from immunization shots, fear from jack-in-the box toys, or frustration from head restraints [3]. Previously, in [4], [5], we proposed preliminary approach which can recognize cry signals of specific infant. However, only limited normal cry signals such as hunger, a wet diaper and a focus are studied, and the algorithms work just for specific infants within the studying a controlled lab environment. Nevertheless, an abnormal cry will be related to severe or chronic illness, so the detection and recognition of abnormal cry signals are of great importance. Compared with normal cry signals, abnormal cry signals are more intense, requiring further evaluation [6]. An abnormal cry is usually associated with medical problems, such as: infection, abnormal central systema nervosum, pneumonia, sepsis, laryngitis, pain, hypothyroidism, trauma to the hypopharynx, vocal cord paralysis, etc. Therefore, approaches which can identify and recognize both normal and abnormal cry signals in practical scenarios is of maximum importance. In this paper, we propose a unique cry language recognition algorithm which can distinguish the meanings of both normal and abnormal cry signals in a very noisy environment. Additionally, the proposed algorithm is individual crier independent. Hence, this algorithm can be widely utilized in practical scenarios to acknowledge and classify various cry features. The proposed algorithm are often want to interpret a babies 'needs, providing parents with an appropriate thanks to sooth infants. Furthermore, it can help parents or infant caregivers avoid misunderstanding of their babies' cries thereby reducing their own stress. It also helps prevent ill-treatment and neglect. Moreover, analyzing infant cries provides a non-invasive diagnostic of the condition of the infant without using invasive tests [7]. Using an infant's cry as a diagnostic tool plays an important role in various situations: tackling medical problems in which there's currently no diagnostic tool available (e.g. sudden infant death syndrome (SIDS), problems in developmental outcome and colic), tackling medical problems in which early detection is feasible only by invasive procedures (e.g. chromosomal abnormalities), and at last tackling medical problems which can be readily identified but would benefit from an improved ability to define prognosis, (e.g. prognosis of future developmental outcome in cases of prematurity and drug exposure [8]).In our model, cry signals are output signals from the vocal tract system, which is additionally called the linear system. The stimuli signal, which excites the linear system, is that the airflow from an infants' lungs [9]. almost like digital speech signal processing, we use a time-varying Fourier transform to review the spectral properties of cry signals. Therefore, we are able to identify the difference between vocal tract systems and input signals, which are related with different cry reasons. During this paper, short-time Fourier transform (STFT) is employed to analyze the cry signals. Recently, speech recognition and acoustic signal classification techniques are widely employed in many areas such as manufacturing, communication, consumer electronic products and treatment [10] [12]. Speech recognition is a signal processing procedure that transfers speech signal waveforms during a spatial domain into a series of coefficients, called a feature, which might be recognized by the pc [10], [13]. Since infant cry signals are time-varying no stationary random signals which are just like speech signals. The stimuli for infant cry signal is that the same

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 03 | Mar -2020 ISSN: 2582-3930

because the stimuli for voiced speech signal. During this paper, we use techniques originally designed and employed in automatic speech recognition to detect and recognize the features for infant cry signals, and use compressed sensing to investigate and classify those signals. Figure 1 shows the procedures of cry signal recognition which consists of the subsequent steps:

Step 1: Cry unit detection
Step 2: Feature extraction
Step 3: Analysis and classification

Fig. 1. Block diagram for infant cry recognition.

This paper is organized as follows. Section II introduces anatomy of infant-related cries and the physiology of cry signals. In Section III, Discrete Fourier analysis is proposed and cry detection techniques are presented. Section IV presents methodology, and proposes a compressed sensing model to recognize and classify infant cry signals. Experimental results are presented in Section V. Finally, in Section VI, we conclude the paper.

II. INFANT CRY MODELLING AND CATEGORIZATION

A. Physiology of Infant

Cry From a physiological point of view, increasing alertness and decreasing crying, as part of the sleep/wakefulness cycle, suggests that there is also a balanced exchange between crying and attention. The change from sleep/cry to sleep/alert/cry necessitates the development of control mechanisms to modulate arousal. An infant should increase arousal gradually to keep up states of attention for extended periods. The infant cry is that the results of complex interactions between anatomic structures and physiologic mechanisms. These interactions involve the central system nervous, the respiratory system, the peripheral system nervous, and a spread of muscles [14]. Newborns differ from each other in their response to different stimuli. There are two main physiological states which infants can switch between: a sleep state and an awake state. Within the sleep state, infants fall into two categories either the quiet sleep or the active sleep category. On the opposite hand, the awake state is characterized by four main behaviors: drowsy, quiet alert, active alert, and crying.

Physiological changes can easily affect an infant's cry behavior directly. In the first few weeks after birth, crying contains a reflexive-like quality and is possibly tied to the regulation of physiological homeostasis because the neonate is balancing internal demands with external demands [15]. As physiological processes stabilize, periods of alertness and attention increase, which place additional demands on regulatory functions. Crying can occur when the system becomes overloaded due to external stimulation. Crying is additionally considered as a mechanism for discharging energy or tension. The need for tension reduction is very acute every now and then of major developmental upheavals and shifts. Unexplained fussiness and sudden increases in crying occur between 3 and 12 weeks old thanks to maturational changes in brain structure and shifts within the organization of the central systema neervosum. Physiological and anatomical changes that occur around 1 to 2 months end in more control over vocalization, thus crying becomes more differentiated. At the age of 7¡9 months there is a second bio-behavioral shift characterized by major cognitive and affective changes that also are thought to reflect central systema nervosum reorganization. Crying now occurs for additional reasons, like fear and frustration [15].

B. Catalog of Cry Signal

The cry production mechanism in infants resembles the speech production process in adults [9]. First, external or internal stimuli will stimulate the infant's brain. Then, the nervous system will transmit the brain's commands to speech and respiratory muscles which control the ejection of air from the lungs to the vocal tract, changing the vocal tract status [14]. As a result, a distinct acoustic sound is uttered. The vibration of vocal cords and muscle movements ends up in a change in air pressure. The cord vibration fundamental frequency is termed the pitch. Similar to speech signals, infant sounds also can be defined as voiced or unvoiced excitations supported different utterance mechanisms. Voiced excitations occur within the larynx and involve vocal cord vibration while unvoiced excitations involve air turbulences of occlusion caused by the mouth, tongue, teeth, or lips. Crying serves several useful purposes for infants. Crying is a way for infants to communicate when they are hungry or uncomfortable [15]. Crying helps them keep out intensive stimuli, such as: sights, sounds, and other sensations. Additionally, it helps infants release tension. Sometimes, crying even helps babies eliminate of excess energy. Normal cries can be thanks to hunger, a requirement for a diaper change or a need to be held. However, there are additionally cries associated with something more severe (abnormal cries), like a hair tourniquet (a piece of hair wrapped very tightly around a finger or toe), an obstruction within the intestine, or pain and sickness. Understanding and identifying the different reasons for various infant cries, especially the abnormal cries can help parents or caregivers choose the correct healthcare service and reduce the risk of health impairment for infants. The following are some common reasons for infant crying [16], [17], hunger, stomach problems, desirous to sleep, a grimy diaper, needing to be hold and etc.

III. CRY SIGNAL TIME FREQUENCY ANALYSIS AND DETECTION

After obtaining cry signals, we analyze the recorded signals by using waveform and time frequency analysis. Then we conduct signal detection and segmentation for later pattern extraction. Signal detection processes instances of voiced activity instead of spending computational time during silent periods. To accurately detect potential periods of voiced activity, two short term signal detection techniques are used.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 03 | Mar -2020 ISSN: 2582-3930

DISCRETE FOURIER TRANSFORM

The discrete-time Fourier transform of a discrete set of real or complex numbers $x[n]$, for all integers n, is a Fourier series, which produces a periodic function of a frequency variable. When the frequency variable, ω, has normalized units of *radians/sample*, the periodicity is 2π, and the Fourier series is:

$$X(\omega) \triangleq \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n}.$$

We say that X is the *spectrum* of x

LINEARITY OF THE DTFT

$$\mathrm{DTFT}(\alpha x_1 + \beta x_2) = \alpha \cdot \mathrm{DTFT}(x_1) + \beta \cdot \mathrm{DTFT}(x_2)$$

Where α, β are scalars (real or complex umbers), x1 and x2 are any two *discrete-time* signals (real- or complex-valued functions of the integers), and X1,X2 are their corresponding continuous-frequency spectra defined over the unit circle in the complex plane.

IV. METHODOLOGY

A. Audio Features

Through the sense of hearing, people can distinguish similar sounds of various types. This is often done through the human perception of qualitative audio features. There are four primary auditory qualities related to sound: loudness, pitch, timbre, and also the source of the sound [11]. Loudness could be a quantitative measure of the amplitude of the sound compared to a reference level and can be qualitatively described from being quiet to loud. Pitch is a quantitative measure of the particular fundamental frequency of a signal and can be qualitatively described from low to high. Timbre could be a qualitative measure of a sound that can be used to help differentiate between two sounds of equal loudness and pitch through the tonal quality of the sound.

B. libROSA

LibROSA is a python package for music and audio analysis. It provides the building blocks necessary to create music information retrieval systems. *The* libROSA *package is structured as collection of submodules:*

- Librosa.beat
- Librosa.core
- Librosa.display
- Librosa.decompose
- Librosa.effects
- Librosa.feature
- Librosa.output
- Librosa.segment, Librosa.utile

C. AUDIO SIGNALS IN PYTHON

Sound is just pressure waves, and these waves can be represented by numbers over a time period. Music stored as .WAV, are the audio waves stored as numbers, and MP3 files are a compressed version of the .WAV

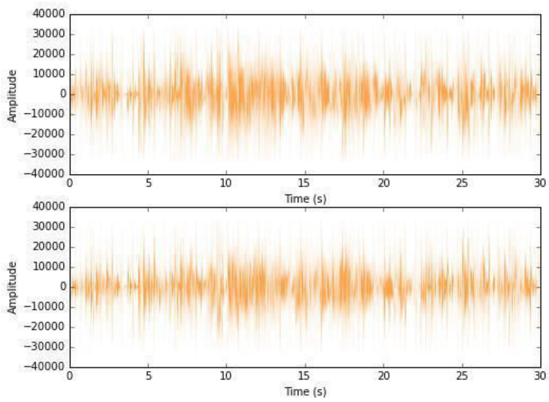

Fig. 2. Signal Processing

The Fig. 2 shows that the audio signal processing. The single audio wave into audio waves at different frequencies. This can be done using a Fourier transform.

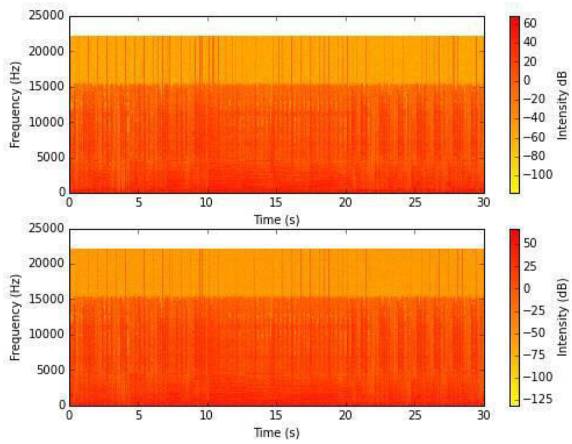

Fig. 3. Frequency Signal

Each Fourier transform over a block, results in the frequencies represented in that block, and to what magnitude. So the resultant array is NFFT times smaller than the original data. The range of frequencies explored relates to half the sample rate. The number of samples in the block (NFFT) determines how many frequencies in that range are considered. So a bigger block results in a greater frequency range, but reduces the information with respect to time.

C. MATPLOTLIB

matplotlib.pyplot is a collection of command style functions that make matplotlib work like MATLAB. Each pyplot function makes some change to a figure: e.g., creates a figure, creates a plotting area in a figure, plots

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 03 | Mar -2020 ISSN: 2582-3930

some lines in a plotting area, decorates the plot with labels, etc.In matplotlib.pyplot various states are preserved across function calls, so that it keeps track of things like the current figure and plotting area, and the plotting functions are directed to the current axes. The pyplot API is generally less-flexible than the object-oriented API. Most of the function calls you see here can also be called as methods from an Axes object. We recommend browsing the tutorials and examples to see how this works.

V. EXPERIMENTS AND RESULTS

When the baby was crying, we placed the microphone around 6¡10 inches away from the infant’s mouth to pick up the cry audio signal. A Sound Digital 722 digital audio recorder was used to record infant cry signals. The sampling frequency was 44.1 kHz with a resolution of 16 bits, and then down sampled to 7350 Hz. The probable reason for each cry signal file was given by experienced neonatal nurses, experienced nurses and caregivers who were able to identify the reason for a baby’s cries after a bit of listening. For example, there are some observed types of newborn cries associated with different audio cues: The “neh” sound is generally related to being “hungry”. Typically, when a baby has the sucking reflex, and his/her tongue is pushed to the roof of the mouth, a “neh” sound is generated. The “owh” sound is made in the reflex of a yawn which means “sleepy”. The “heh” sound means “I need something”, such as: being too cold, being itchy, needing a new diaper, or needing a new body position, etc.. The “eair” is a deeper sound which comes from the abdomen, so it means lower gas pain. It is usually accompanied by a newborn pulling his/her knees up or pushing down his/her legs. The “eh” sound means that a baby needs to burp. Generally speaking, it happens after feeding. Besides listening to cry signals, experienced personnel can confirm the reasons for different cries by considering other cues, such as gesture, facial expressions, and motion. For example, some hunger signs for newborns include fussing, lip smacking, rooting (a newborn reflex that makes babies turn their head toward your hand when you stroke their cheek), and putting their fingers to their mouth [21]. Crying caused by wet diapers can be distinguished by just checking infant’s diaper. The signs for being “sleepy” are yawning, rubbing eyes and nodding. Attention crying can be easily soothed by holding infants or interacting with them. Discomfort crying, such as an injection or blood test, could be associated with a certain medical procedure. All the babies have their own nursing logs containing information including: age, sex,temperature, blood pressure, feeding time, diaper change time, sleep time and so on. Nurses can use the data provided as well as deductive logic to then interpret an infant’s cry. For instance, if a baby was fed a few minutes ago, then their crying is most likely not due to hunger and an infant who just woke up usually does not cry because they are sleepy.

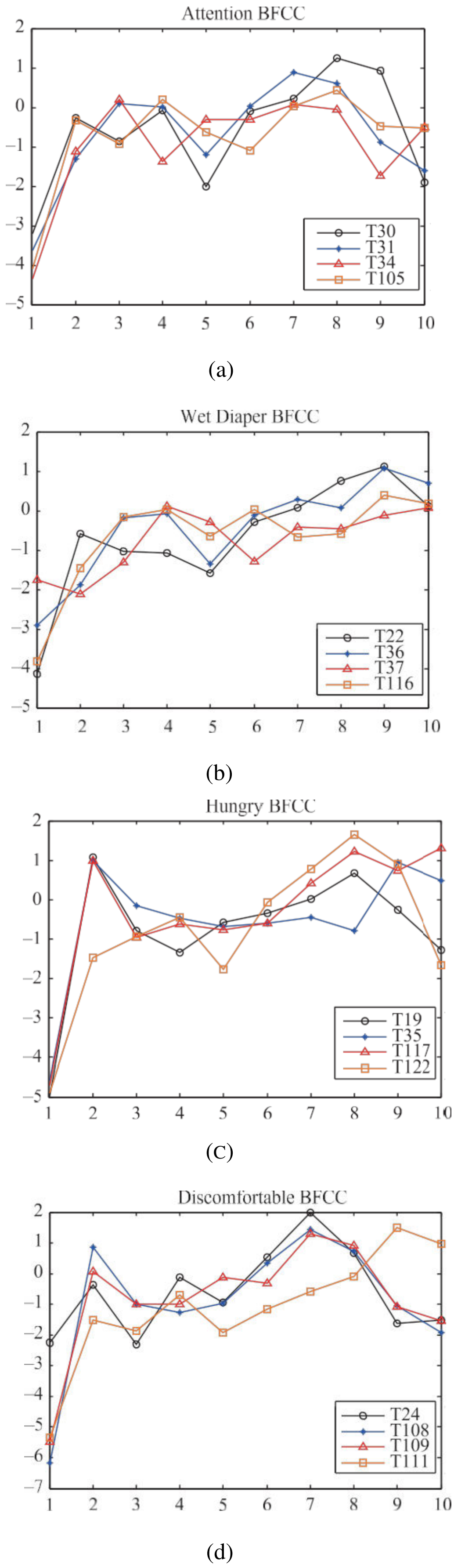

Fig. 4. BFCC features for attention, diaper, hungry and discomfort cry signal

Fig. 4 shows the BFCC features for different catalogs from different infants. BFCC features for attention from 4 different babies are shown in (a). Features from one infant are similar to other infants when they had a similar reason to cry. Subplot (b) shows “Diaper change needed cry units” BFCC features of 4 different cry files. Again, the results show similar features for needing a diaper change across different infants. Since attention-related crying and diaper-related crying both are characterized as normal crying, their intensity levels are similar but less than hunger-related crying. Fig. 4 (c) shows “Hungry cry units” BFCC features of 4 different babies. Hunger-related crying had the highest intensity level in the normal cry catalog. BFCC features obtained from “hungry” is quite different from those of “attention” crying and “diaper” crying. It is shown that the BFCC patterns changed from the low stress level cries to high stress level cries. There is an abrupt jump from coefficient 1 to coefficient 2 which is close to the trend of abnormal cry signals. Fig. 4 (d) shows discomfort-related crying from 4 files associated with 4 different babies. The BFCC features show a similar trend among those infants. They are quite different from normal cry signals, especially the low intensity level cries, such as diaperrelated and attention-related crying. And even compared with hunger-related cry signals, the values of the coefficients were higher which means discomfort-related crying produced higher energy cry signals which matches the experts’ experience.

VI. CONCLUSION

This paper presents a unique cry detection and recognition method for individual independent infant cries in a noisy environment. Audio features of infant cry signals were obtained in time and frequency domains, and were used to perform infant cry language recognition. Practical data from hospitals were used to design and verify the proposed approaches. Experiments proved that the proposed infant cry unit recognition models offer accurate and promising results with far reaching applications medically and societally.

REFERENCES

1. B. Goldand N. Morgan, *Speech and Audio Signal Processing*. New York, NY, USA, John Wiley & Sons, 2011.

2. *C. F. Z. Boukydis and B. M. Lester, Infant Crying: Theoretical and Research Perspectives, Berlin, Germany, Springer Science and Bussiness Media, 2012.*

3. C. Lau, “Development of suck and swallow mechanisms in infants,” Ann. Nutr. Metab., vol. 7, no. 5, pp. 7¡14, July 2015.

4. Dong Yu and Jinyu Li. “Recent progresses in deep learning based acoustic models,” *IEEE/CAA J. Autom. Sinica*, vol. 4, no. 3, pp. 396¡409, April 2017.

5. D. Needell and R. Ward, “Two-subspace projection method for coherent overdetermined systems,” Journal of Fourier Analysis and Applications, vol. 19, no. 2, pp. 256¡269, April, 2013.

6. F. Katzberg, R. Mazur, M. Maass, P. Koch, and A. Mertins, “A compressed sensing framework for dynamic sound-field measurements,” IEEE/ACM Trans. Audio, Speech, and Language Processing, vol. 26, no. 11, pp. 1962¡1975, Jun. 2018.

7. *G. Buonocore and C.V. Bellieni, Neonatal Pain, Suffering, Pain and Risk of Brain Damage in the Fetus and Newborn, Berlin, Germany, Springer, 2008.*

8. H. Karp, *The Happiest Baby on the Block*; Fully Revised and Updated Second Edition: *The New Way to Calm Crying*, New York City, NY, USA, 2015.

9. J. A. Green, P. G. Whitney, and M. Potegalb, “Screaming, yelling, whining and crying: categorical and intensity differences in vocal expressions of anger and sadness in children’s tantrums,” *Emotion*, vol.5, no. 11, pp. 1124¡1133, Oct. 2011.

10. L. LaGasse, A. Neal, and M. Lester, “Assessment of infant cry: acoustic cry analysis and parental perception,” *Ment Retard Dev Disabil Res Rev*., vol. 11, no. 1, pp. 83¡93, 2005.

11. L. Liu and K. Kuo, “Active noise control systems integrated with infant cry detection and classification for infant incubators,” in *Proc. Acoustic*, pp. 1¡6. 2012.

12. L. Liu, K. Kuo, and Sen M. Kuo, “Infant cry classification integrated ANC system for infant incubators,” in *Proc. IEEE International Conf. on Networking, Sensing and Control*, Paris, France, 2013, pp. 383¡387.

13. L. L. LaGasse, R. Neal, and B. M. Lester. “Assessment of infant cry: acoustic cry analysis and parental perception,” *Mental Retardation and Developmental Disabilities Research Reviews*, vol. 11, no. 1. pp. 83¡93, 2005.

14. M.Sahidullah, and G. K. Saha, “Design analysis and experimental evaluation of block based transformation in MFCC computation for speaker recognition,” *Speech Communication*, vol. 54, no. 4, pp. 543¡565, May 2012.

15. P. Dunstan, Calm the Crying: The Secret Baby Language That Reveals the Hidden Meaning Behind an Infant’s Cry, New York City, NY, USA, Avery, 2012.

16. P.Runefors and E. Arnbj¨onsson, “A sound spectrogram analysis of children’s crying after painful stimuli during the first year of life,” Folia

honiatr. Logop, vol. 2, no. 57, pp. 90¡95, Mar–Apr. 2005.

17. S. Ludington-Hoe, X. Cong, and F. Hashemi, “Infant crying: nature, physiologic consequences, and select interventions,” *Neonatal Netw*. vol. 21, no. 2, pp. 29¡36. Mar. 2002.

18. V´arallyay Jr. Gy¨orgy, “Future prospects of the application of the infant cry in the medicine,” *Periodica Polytechnica Ser. El. Eng*, vol. 50, no. 1¡2, pp. 47¡62, 2006.

19. V. R. Fisichelli, S. Karelitz, C. F. Z. Boukydis, and B. M. Lester, “The cry attencies of normal infants and those with brain damage,” *Infant Crying*, Plenum Press, 1985.

20. Y. Kheddache and C. Tadj, “Acoustic measures of the cry characteristics of healthy newborns and newborns with pathologies,” *Journal of Biomedical Science and Engineering*, vol. 6, no. 8, 9 pages, 2013.

21. Z. Ren, K. Qian, Z. X. Zhang, V. Pandit, A. Baird, and B. Schuller “Deep scalogram representations for acoustic scene classification,” *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 3, pp. 662¡669, May 2018.