Vol: 06 Issue: 06

International Journal of Data Science and Advanced Analytics (ISSN: 2563-4429)

Detecting Autism and Epilepsy through Infant Crying Signals: Key Features for Identification

Eng. Mahasin El Dimassi[1,*], Dr.Khodor M. Ozoor [2]

[1] PHD Student of Electrical and Communication Engineering, Beirut Arab University Beirut, Lebanon.
[2] Pediatric Intensive care unit, pediatric emergency department at AUBMC ,American University of Beirut Medical Center, Medical Director at HMC Hospital Beirut Lebanon
[1]m.demassi@bau.edu.lb, [2]dr_ozoor@yahoo.com

Abstract — Crying is an important means of communication for babies during infancy, from birth to three months of age . At this stage of their development, infants are almost entirely dependent on the people who care for them. As a result, crying plays an important role for the survival, health, and development of the child . This paper explores the relationship between autism, epilepsy, and crying patterns in newborn infants. The objective is to investigate the potential of analyzing acoustic features of infant cries as a non-invasive method for early detection and diagnosis of these neurodevelopmental disorders. By utilizing machine learning algorithms, a predictive model can be developed to differentiate between typical cries and those indicative of autism or epilepsy. To conduct this research, a specific dataset comprising crying sounds from newborns diagnosed with autism, epilepsy, and typically developing infants is utilized. Acoustic features, including pitch, intensity, and duration, are extracted from the cry recordings. These features are then used to train machine learning models, such as support vector machines (SVMs) or artificial neural networks (ANNs), enabling the classification of cries into distinct categories. The developed model's accuracy and generalizability are assessed using cross-validation techniques and evaluation metrics such as sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC-ROC). Additionally, the model's reliability and applicability in real-world scenarios are validated using an independent dataset. This study aims to provide a non-invasive and accessible tool for early detection and diagnosis of autism and epilepsy in newborn infants. The utilization of crying patterns as an indicator holds promise for timely intervention and improved outcomes in affected individuals and underscores the importance of early identification in neurodevelopmental disorders.

***Keywords*— autism, epilepsy, newborn infants, crying patterns, acoustic features, machine learning, early detection, diagnosis.**

1. Introduction

All that a newborn can do to express himself is to cry, he cries when he is hungry, when he feels the pain, when he needs to change the dirty layer or even to attract attention, so it is a means of communication. This means of communication plays a vital role during early age .
The analysis of the cry for infants allows the early diagnosis of various pathologies when the symptoms are not yet observable. The cries of babies have been studied for years, and it has become clear that these cries contain important information about the condition of the newborn [1].
Most newborns appear healthy at birth, including those with serious diseases that require special care, or even curable if diagnosed early [2], in order to create a diagnostic system, a database was collected by recording infants who were crying. But as it is important to record cries in the very first days after birth, the final recordings contained noises (slamming doors, talking people, machine beeps, etc.) or silences that may affect the classification system.
So, to eliminate unwanted noises, segmentation is necessary. Segmentation means choosing whether a segment is useful or not to allow the classifier to learn them. This leads us to ask if it is possible to exploit the methods of artificial intelligence in order to create an automatic segmentation system capable of distinguish between useful and non-useful segments instead of using manual segmentation .It is for this reason that we hypothesize that to solve this problem we propose an automatic segmentation system. This system, after a feature extraction we use the MFCC (Mel-Frequency Cepstral Coefficient), used a probabilistic neural network (PNN) to learn in order to classify segments as useful or not automatically.
A cry is composed of two phases, one phase of inspiration which is almost silent and another of expiration during which the baby emits a sound (statement),this statement, which will later be called "a phonation", is produced by the larynx [3].The lungs subsequently emit a stream of air that will vibrate the vocal cords located in the larynx at a fundamental frequency f0 [4][5].Each baby has an acoustically unique cry, in most healthy babies the fundamental frequency F0 varies between 250Hz and 450Hz, with a first RF1 formant of 1100Hz and a second RF2 of 3100Hz [2]. According to (Corwin and Golub, Lester et al.) [2] [6], the cries are classified according to their F0 in five categories:

1. Inspiratory Phonation: audible part during inspiration .
2. Basic cries or expiratory phonation: This is the voiced sound generated by the vibration of the vocal cords at an F0 ranging between 350 and 750Hz. It is defined as a basic mode of oscillation of the glottis. On the spectrogram, she is characterized by well-defined and equidistant harmonics.

3. High Pitch Cry or high scream: F0 between 750 and 1000Hz.
4. Expiratory hyper phonation: Its voiced sound characterized by a very high F0 ranging from 1000 to 2000 Hz.
5. A loud, turbulent cry is a cry due to an aperiodic movement of the vocal cords. This mode involves a significant alteration of the basic mode of screaming. On the spectrogram, it is generally characterized by harmonic non-equidistant harmonics[7].

The earliest analysis of crying used musical notation to describe infant cries (Gardiner, 1838), however more recently a few different acoustic measures have been used to examine infant vocalizations [8] . Analysis has focused on examining spectrographic features of cry [9], and durational measures of cry episodes [8-9] . Finally, long-term average spectra (LTAS) measurements [10], and cepstral analysis [11] have been studied . Across nearly all decades of infant cry research, fundamental frequency (F0) has been the most commonly examined measure [11] . Physiologically, F0 relates to the rate of vibration of the vocal folds, reflecting both tension and mass of the vocal folds. Infant cry F0 analyses gave mixed results.

For example, Murry, Amundson, and Hollein (1977) found no differences in F0 between pain cries and other types of cries, whereas Fuller and Horii (1986) found pain cries to exhibit higher F0 compared to cooing or hunger cries, reflecting the increased arousal associated with infant pain. In addition, the analysis techniques used to measure F0 have varied. Some studies have sampled a limited number of individual cries for each infant, from only 2–3 cries [7], to 8–15 cries [11], whereas other studies have examined full crying episodes [12] .

Li Shihong (1995) using the three-dimensional dynamic mapping method analysis of 20 cases of normal infants less than six months[13], and 40 cases of non-normal baby cry resonance and fundamental frequency, and consider normal baby crying fundamental frequency below abnormal baby; Yuan Fengling (2011) studied characteristic analysis and recognition of baby crying sound through computer signal processing method, using MFCC parameters to analyze the impact of recognition rate of the different growth stages of pain and non-pain cries[14-15] .Gustavoon et al. analyzed and resynthesized a set of infant cries to form natural-sounding cries with varying fundamental frequency (F0), degrees of jitter (period to period variations in F0) , and rise time (time for F0 to reach its maximum value)[16].

2. Why Baby is Crying

Newborns usually spend 2 to 3 hours a day crying. Normal as it may be a bawling baby can be distressing for infants and parents alike.

- ✓ **Reason 1 : Hunger**
- ✓ **Reason 2 : Cold**
- ✓ **Reason 3 : Sleepy**
- ✓ **Reason 4 : Gas**
- ✓ **Reason 5 : Bored**
- ✓ **Reason 6 : (NEW detection of crying due to autism or Epilepsy)**

2.1 Crying Autism and Epilepsy for newborn

There is a strong correlation between newborn infants' crying patterns, autism, and epilepsy. Studies have demonstrated that the acoustic properties of newborn cries can be influenced by neurodevelopmental disorders like autism and epilepsy. As an example, crying patterns that are atypical for infants with autism tend to be higher pitched and last longer (Esposito et al., 2017)[17]. Bolton et al. (2011)[18] discovered that infants suffering from epilepsy frequently exhibit characteristic crying patterns, which include high-pitched cries accompanied by abrupt changes in intensity. According to these results, listening to baby cries and examining their auditory characteristics may be able to detect whether a newborn has autism or epilepsy. These acoustic features have been used by machine learning algorithms to create predictive models, allowing for the distinction of normal crying from that linked to neurodevelopmental disorders (Takahashi et al., 2020)[19]. Healthcare practitioners can obtain important insights for early detection and intervention by utilizing these models. As a non-invasive and easily accessible method of identifying infants at risk of autism or epilepsy, the analysis of crying patterns holds promise for improving outcomes through prompt interventions.

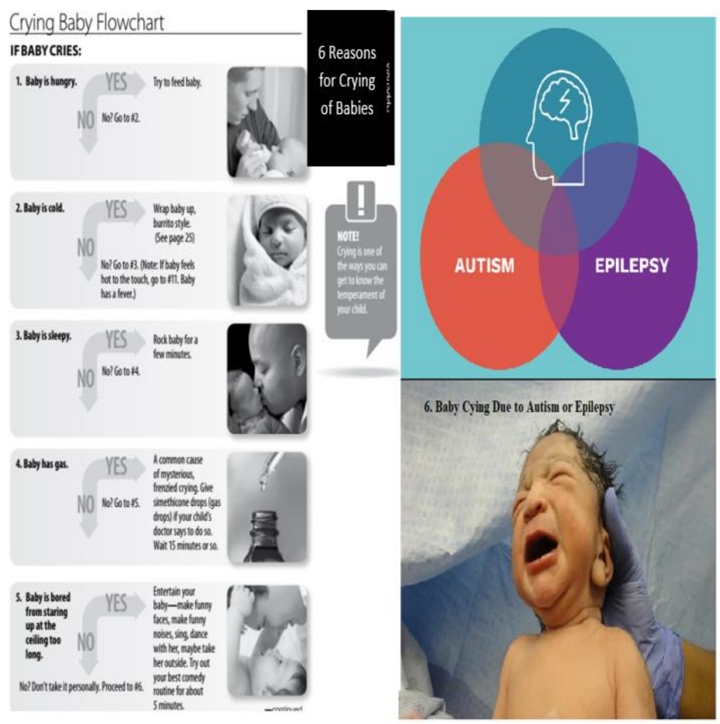

Figure 1: Flow chart to distinguish 6 types of crying in newborn infants ,1(hungry),2(cold),3(sleepy),4(gas),5(bored),6(Last type is a new type for the detection of epilepsy and autism for a crying of a new born).

2. 2 <u>The characteristics of the screams:</u>

The characteristics and patterns of baby cries have been defined as follows [20]:

Characteristics of Cries	**Definition**
Latency	The time between stimulation and the beginning of the cry
Threshold	This is the number of minimum stimulations needed to cause the scream.
Number of statements	Number of expiratory phonation's during the cry.
Inspiratory period	This is the interval in second between two statements
Amplitude	This is the average energy in dB of a statement.
duration	This is the duration of the expiration phonation
Second break	It's time in second between the end of a phonation expiratory and the next inspiration.
Fundamental frequency	The fundamental frequency is the lowest line in the spectrogram. His harmonics appear above like parallel lines [21].
Offset (Shift)	It is the rapid increase or decrease of fundamental frequency.
Melody	The different types of melodies are classified as follows: decreasing, increasing-decreasing, flat or without type with a completely unstable scream representing a sound inaudible or "glottal plosive".
slip	A rapid change from the fundamental frequency of at less than 600Hz in 0.1 second.

Table 1: Characteristics and patterns of baby cries

In a spectrogram, the voiced signals are displayed as black lines. We can observe the different characteristics of the cry on the following diagram:

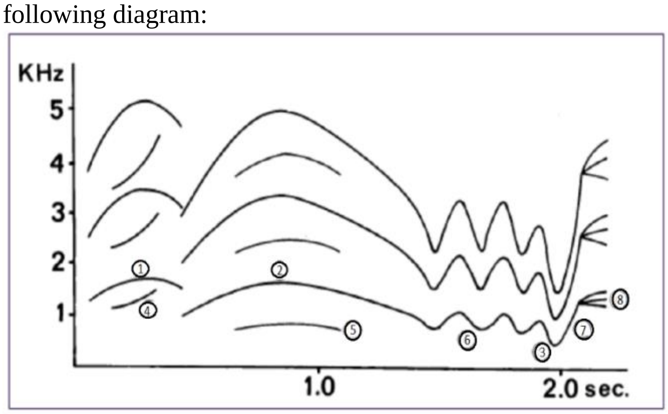

Figure 2. Spectrogram of Cries: 1) shift and max of F0, 2) Max of F0, 3)Min F0, 4) Bi-phonation, 5) double breaks harmonic, 6) Vibrato, 7) Glide, 8)Division.

3. System Layout

The system is based on recording the crying of a babies, that undergo, then, the preprocessing operation, cry detection, segmentation, Filtering and feature points or feature extraction definition. The following Block diagram summarizes the work done:

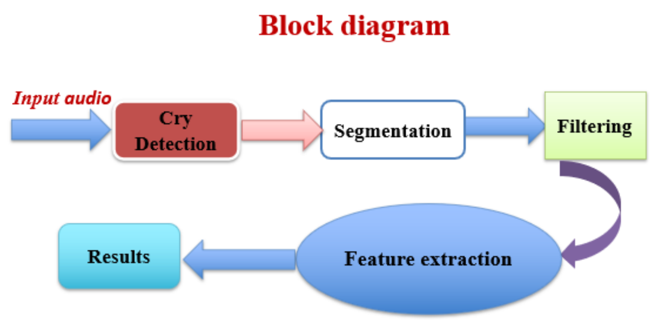

Figure 3: Block diagram.

3.1 Steps:

1- This research proposes a new approach termed here as "Crying recognition" for analyzing the crying of baby.
2- A *MATLAB* Code has been developed to analyze input voice.
3- Load audio file.
4- Plot the input signal.
5- Passing audio to designed high pass filter first.
6- Plot the output signal.
7- Use |fft| to compute the DFT of the signal. Correct the frequency range for the factor of 10 speed-up in the data.
8- Plot the output signal of FFT.
9- Pitch of a signal
10- Plot the output Pitch.

11- Get the fundamental Frequency.

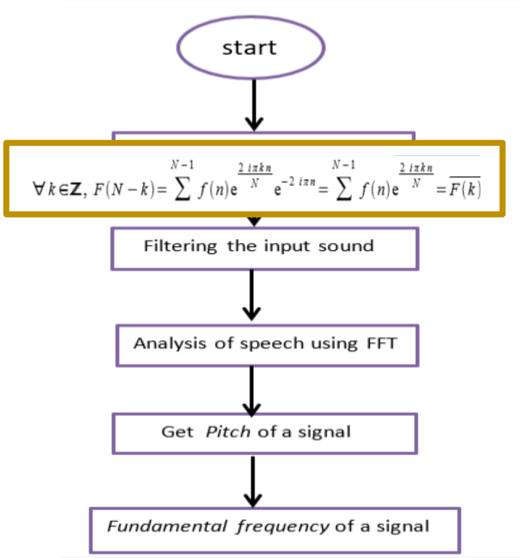

Figure 4: Flow Chart.

4. Th

$$\forall x \in \mathbb{R},\ (f*g)(x)=\int^{+\infty} f(t)g(x-t)$$

4.2- **Fourie**

The Fourier transform makes it possible to carry out a time-frequency analysis with a sufficient resolution for the voice signal, which is quasi-stationary over intervals of the order of 10 to 100 ms .

The Fourier transform F (or TF (*f*)) of a function *f* makes it possible to carry out a projection on the vector space generated by the $\{e_\omega = t \rightarrow e^{i\omega t} / \omega \in \mathbb{R}\}$ sinusoids, by

$$F(\omega)=\langle f \mid e_\omega \rangle=\int_{-\infty}^{+\infty} f(t)e^{-i\omega t}dt$$

first calculating the component of the projection on each sinusoid, given by the scalar product, and not forgetting the complex conjugation compared to the second place . In our study, this transform will never serve us to reconstruct the signal, but in a more qualitative way to examine the contributions of each frequency range in the speech signal by studying the spectrogram, which is in fact a three-dimensional representation indicating the magnitude (module of F (ω)) in the time-frequency plane . (According to Ohm's law, the ear is not sensitive to the phase of the auditory signal.) However, this continuous representation must be adapted to the discrete case, which is that of numerical computation.

4.2- **Discrete Fourier Transform**

The signal is represented by samples taken uniformly over time:

$$\{f(n) / n \in [\![0, N-1]\!]\}$$

The signal of period N is used to avoid edge effects. Then the discrete Fourier transform is then given by:

$$F(k)=\sum^{N-1} f(n)e^{\frac{-2i\pi k}{N}}$$

The frequency of observation is proportional to the factor $\frac{K}{N}$ for $k \in [0, \frac{N}{2}]$ if *fe* is the sampling frequency, so $f=f_e \cdot \frac{k}{N}$. It should be noted that this transformation is redundant because the information obtained is concentrated in half of the interval $[\![0, N-1]\!]$. indeed:

and therefore, the module is the same, the phase being simply of opposite sign. If we want to analyze the signal over a range of 10 kHz , we must take a sampling frequency of 20 kHz . In practice, we implement Fast Fourier Transform (FFT) which has a complexity in O (Nlog2 (N)) instead of O (N^2) by direct computation and which takes advantage of this redundancy to optimize the calculation.

4.2-**Fourier transform and convolution**

In continuous representation, the convolution between two functions is defined by:

By changing the variable in u = x-t, the convolution operator is commutative. It is also associative.

$$\begin{aligned}((f*g)*h)(x) &= \int_{-\infty}^{\infty} f*g(x-y)h(y)\,dy \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x-y-t)g(t)h(y)\,dt\,dy \\ &= \int_{-\infty}^{+\infty} g(t)\Big(\int_{-\infty}^{\infty} f(x-t-y)h(y)\,dy\Big)dt \\ &= \int_{-\infty}^{+\infty} g(t)(f*h)(x-t)\,dt \\ &= (g*(f*h))(x)\end{aligned}$$

and by commutativity we have $(f*g)*h=(f*h)*g$.

We prove that the Fourier transform of the convolute of two functions is usual the product of the Fourier transforms. From the definition of equality: $F(\omega)=(f*e_\omega)(0)$

In addition, we have: $\forall x \in \mathbb{R},\ (f*e_\omega)(x)=F(\omega)\times e_\omega(x)$

because:

$$(f*e_\omega)(x)=\int_{-\infty}^{+\infty} f(t)e^{i\omega(x-t)}=e^{i\omega x}\int_{-\infty}^{+\infty} f(t)e^{-i\omega t}=F(\omega)e_\omega(x)$$

Finally, the Fourier transform of $h = f * g$, using successively the first property, the associativity then the second property:

$$H=(e_\omega * f * g)(0)=((e_\omega * f)*g)(0)=F(\omega)\times(e_\omega * g)(0)=F(\omega)\times G(\omega)$$

and, $\quad TF(f*g)=F\times G$

This result is also valid in discrete representation. This will serve us as a formal analysis because of the modeling adopted for the voice signal.

4.3- **FFT (Fast Fourier Transform):**

The FFT, or fast Fourier transform, is one of the few algorithms whose publication has caused a real revolution in the technical field. Generally associated with the names of JW Cooley and JW Tuckey who published it in 1965, this algorithm for calculating the discrete Fourier transform had been repeatedly "rediscovered" since Gauss, notably by Danielson and Lanczos in 1942. The FFT allows to reduce the computation of the discrete Fourier transform of N^2 to $N\log N$ operations; this reduction in complexity is enough to make many problems impossible to solve. Essentially, the FFT is still the DFT for transforming the discrete-time signal from time domain into its frequency domain. The difference is that the FFT is faster and more efficient on computation . And there are many ways to increase the computation efficiency of the DFT, but the most widely used FFT algorithm is the Radix-2 FFT Algorithm [21]. Since FFT is still the computation of DFT, so it is convenient to investigate FFT by firstly considering the N-point DFT

$$\mathcal{S}(k) = \sum_{n=0}^{N-1} x(n) e^{-2\pi jkn/N} \quad \text{for} \quad k = 0, 1, \ldots, \{N-1\}.$$

2m+1) and he N-point DFT equation also becomes two parts for each N/2 points:

$$X(k)=\sum_{n=0}^{N-1} x(n)W_N^{kn} = \sum_{m=0}^{N/2-1} x(2m)W_N^{2mk} + \sum_{m=0}^{N/2-1} x(2m+1)W_N^{(2m+1)k} = \sum_{m=0}^{N/2-1} x(2m)W_N^{2mk} + W_N^k \sum_{m=0}^{N/2-1} x(2m+1)W_N^{2mk},$$

where m 0, 1, 2, ..., N/2-1

Since:

$$e^{j\omega_k n} = \cos(\omega_k n) + j\sin(\omega_k n).$$

$$e^{j(\omega_k+\pi)n} = \cos\left[(\omega_k+\pi)n\right] + j\cdot\sin\left[(\omega_k+\pi)n\right]$$

$$= -\cos(\omega_k n) - j\cdot\sin(\omega_k n) = -[\cos(\omega_k n) + j\cdot\sin(\omega_k n)] = -e^{j\omega_k n}$$

That is:

$$e^{j(\omega_k+\pi)n} = -e^{j\omega_k n}$$

In speech processing the speech signal should be first transformed and compressed for further processing. There are many signal analysis techniques available which are used to extract the important features and compress the signal without losing any important information. Among the most important techniques are FFT and LPC. FFT is used to convert the word signal into spectrum, but FFT require only complex value[22] .

As simple digital filtering can be applied to a long audio vector in a single pass irrespective of its size (and irrespective of whether it has been segmented) . For example, an FIR (finite impulse response) filter is achieved with:

y=filter(b, 1, x).

Where *b* is a vector holding the filter coefficients. The x and y vectors are equal-length vectors corresponding to the arrays of input and output samples respectively, with each sample of y being calculated from the difference equation :

$$y(n) = b(1)\times x(n) + b(2)\times x(n-1) + b(3)\times x(n-2) + \cdots + b(m+1)\times x(n-m).$$

Frequency-domain operations, by contrast, usually require the audio to be first converted to the frequency domain by using a Fourier transform (or similar, since there are other transforms and spectral estimation methods). An efficient method used by most researchers is the fast Fourier transform (FFT), which is used like this :

a_spec=fft(a_vector)

and implements a transformation that is equivalent to a discrete Fourier transform, *S(k)*,on an original time sequence *x(n)* of length *N* points:

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^{kn}, \quad k = 0,\ 1,\ 2\ldots\ N-1$$

Those who l w that the original f-two sized input windows (vector lengths).

4.4- **Pitch vs fundamental frequency:**

The fundamental frequency F0 and the pitch have a well-known physiological relation . The pitch of the human voice is one of the most easily and quickly learned acoustic attributes. It plays a central role in both the production and the perception of speech . The pitch is the fundamental frequency perceived by the ear. This frequency gives intonation information of the spoken sentence and also a lot of information about the speaker .

A. Fundamental Frequency

The fundamental frequency F0 is the frequency of vibration of the vocal cords. In the time domain, it is the period of a signal voiced at a given moment. This fundamental frequency gives us an index on the pitch of the vocal signal. For the speech signal, its fundamental frequency is nothing more than the frequency of the opening / closing cycle of the vocal cords, determined by the tension of the muscles controlling them. The average range of this frequency varies from speaker to speaker depending on age and gender. It ranges from approximately 80 to 200 Hz for men, 150 to 450 Hz for women, and 200 to 600 Hz for children [23]. The detection of F0 plays an essential role in the field of speech processing and should be, if possible, for modern applications, calculated in real time.

In each case, the associated fundamental frequency *f0*, to a fundamental period T0 is defined as :

$$f0 = \frac{1}{T0}$$

B. Pitch:

It plays an extremely important role in segmenting the information contained in a message. The pitch, or rather the variations of the pitch, also gives us information on the emotivity of a speaker. It is thanks to him that we distinguish a statement from a question or an order . It is also a sign of the elongation of the voice, it is one of the parameters on which we can act in certain cases to give more natural to the pathological voices. In some foreign languages (Chinese for example), its variation encodes the information directly. This variation is imposed by the language and sometimes can distinguish two words which are pronounced in the same way, but they only differentiate in the variation of the pitch. Finally, in music, you can play on this parameter to correct the vocal imperfections or the false notes of certain singers .

The detection of the pitch has some important problems :
1-The variation of the fundamental frequency F0 in time .
2- the appearance of harmonics that distort the detection.
3- the difficulty of making the voiced / unvoiced decision in the pitch contours.

4.5- **Segmentation**

Audio segmentation (often called audio classification) is a preprocessing step in audio analysis that separates different types of sound, for example, speech, music, environmental sounds, silence, and combinations of these sounds[24],[25]

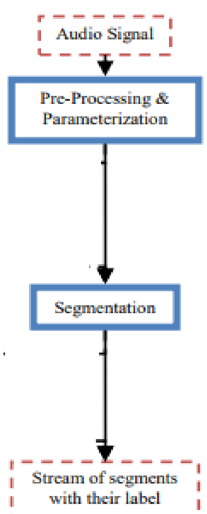

Figure 5 . General architecture of the automatic audio segmentation scheme.

4.6- **Feature extraction of an audio signal**

MFCC alone can be used as the feature for voice recognition. The recorded speech signals are sampled and stored using Audacity.

$$Mel(f) = 2595\log_{10}(1 + \frac{f}{700}).$$

Linear Frequency Cepstral Coefficients **(LFCC)** is a technique similar to **MFCC**, with the exception that it uses a located filter bank on a linear frequency scale.

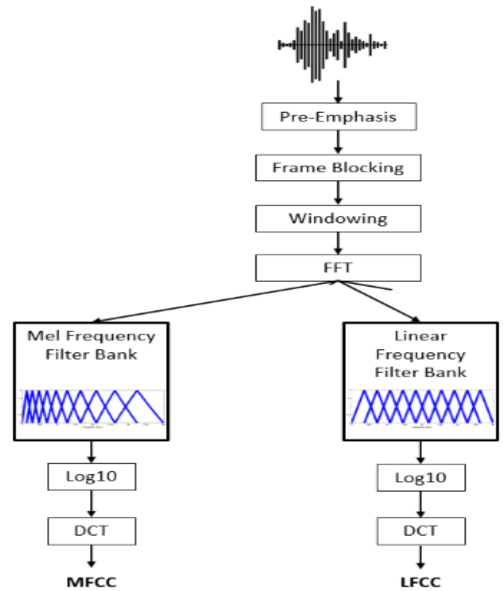

Figure 6. MFCC versus LFCC

4.7- **Feature extraction of an audio signal**

LFCC proved to be more adequate than **MFCC** for that particular task (with performances around **90%** in recognition rate, for each of the five classes corresponding to the Dunstan baby language, compared to **79%** with **MFCC**) [26]. Other authors have chosen MFCC and KNN classifier [27] and reported performances of 80% and 90% recognition rate for the 5 types of crying corresponding to the 5 words of Dunstan baby language. However, the authors have used recordings with children of age 1 day - 6 months, while Dunstan's theory refers to the age under 3 months[27],[28],[29].

4.8- **Infant cry classification models(SVM: support vector machine, ANN: artificial neutral network)**

Traditional machine learning classifiers:
Support Vector Machine (**SVM**) is a widely used probabilistic classifier in infant cry classification. There are different types of SVMs, including multi-class SVMs and binary SVMs with linear and RBF (Radial Basis Function) kernels. Various features such as temporal features, prosodic features, and cepstral features are inputted into the SVM for classification.
In a study conducted by Onu et al. in 2017, SVM was compared to other non-linear classifiers, such as neural networks, for asphyxia classification. The researchers concluded that SVMs are particularly effective in scenarios with limited examples and high-dimensional data .
In another study by Chang et al. in 2015, an incremental SVM learning model was employed for infant cry classification based on FFT (Fast Fourier Transform) features. This model continuously incorporates new data into the dataset during each training step, resulting in an improvement of over 18% in accuracy compared to the original SVM model [30].
Both Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs) have been used for analyzing the crying of

babies in research studies. Here's an overview of how they can be applied in this context:
1. Artificial Neural Networks (ANNs):
ANNs, particularly feed-forward neural networks, can be used for cry analysis by training them on cry-related features. The process involves the following steps:
- Feature Extraction: Cry-related features, such as spectral or cepstral features, are extracted from the cry signals. These features capture important characteristics of the cry, such as pitch, duration, intensity, etc.
- Training Data Preparation: The cry-related features are paired with corresponding labels indicating the cry type or the underlying cause (e.g., hunger, pain, discomfort). This labeled dataset is split into training and testing sets.
- Network Architecture: A suitable architecture for the feed-forward neural network is chosen, which typically consists of an input layer, one or more hidden layers, and an output layer. The number of nodes in each layer and the activation functions are determined based on the specific problem.
- Training: The neural network is trained using the training dataset, where the weights and biases of the network are adjusted iteratively to minimize the difference between the predicted cry type and the actual label. This is typically done using gradient-based optimization algorithms like backpropagation.
- Evaluation: The trained network is evaluated on the testing dataset to assess its performance in classifying the cry types. Metrics such as accuracy, precision, recall, and F1 score are used to measure the network's effectiveness.
2. Support Vector Machines (SVMs):
SVMs are another popular machine learning method used for cry analysis. Here's how SVMs can be applied:
- Feature Extraction: Cry-related features are extracted from the cry signals, similar to the process described for ANNs.
- Training Data Preparation: The cry-related features and their corresponding labels are used to create a labeled dataset for training and testing purposes.
- Training: SVMs aim to find an optimal hyperplane that separates different cry types in the feature space. The training process involves identifying the support vectors (representative samples) from the labeled dataset and determining the hyperplane that maximizes the margin between different cry classes.
- Evaluation: The trained SVM model is tested on the testing dataset to evaluate its classification performance. Metrics such as accuracy, precision, recall, and F1 score are commonly used to assess the SVM's effectiveness.
Both ANNs and SVMs have their own strengths and weaknesses. ANNs are known for their ability to learn complex patterns and capture nonlinear relationships but may require more computational resources and training data. SVMs, on the other hand, are effective in handling high-dimensional data and can handle limited training examples well. The choice between these models depends on factors such as the specific problem, available data, and computational resources.

5. **Results and Discussion**

5.1 Fundamental or Pitch: Frequency

The Baby Speech Recognition app identifies the five types of cries using specific frequency ranges:

a- Autism and epilepsy Cry: If the frequency ranges between 1100 and 1450 Hz,(autism 1110 to 1250 Hz,Epilepsy 1260 to 1450 Hz) the baby is crying due to brain pain and the crying is continuous for a long time. If the sound does not lie in this frequency, the process continues.
b- Pain Cry: If the frequency ranges between 700 and 1100 Hz, the baby is crying due to pain. If the sound does not lie in this frequency, the process continues.
c- Hunger Cry: If the frequency ranges between 400 and 600 Hz, then the baby is crying due to hunger.
d- Fear Cry: If the frequency ranges between 50 and 400, then the baby is crying due to fear.
e- Happy Cry: If the frequency ranges between 600 and 700, then the baby is happy.
f- Undetected Voice: If the frequency of the voice does not lie in any of the above frequency ranges, then it shows that the voice is Undetected.

5.3- **Results**

In the study, a sample of infants ranging in age from 10 days to 3 months was examined to investigate the association between specific crying patterns and various conditions. Out of the total sample 100(65 males and35 female) infants, 20 infants were identified as exhibiting crying patterns suggestive of autism, while 16 infants displayed crying patterns associated with epilepsy. Furthermore, 32 infants were found to cry predominantly due to hunger, while 8 infants exhibited crying patterns indicative of colic, while 11 infants exhibited crying patterns indicative of sleepy, while 13 infants exhibited crying patterns indicative of cold. Additional conditions and factors were also considered and documented.
The study yielded significant findings, providing compelling evidence of the link between specific crying patterns and certain conditions in infants. The identification of distinct crying patterns associated with autism and epilepsy highlights the potential of utilizing cry analysis as a non-invasive screening tool for early detection of these neurodevelopmental disorders. Moreover, the recognition of hunger-related cries and colic-associated crying patterns contributes to our understanding of infant needs and supports the development of targeted interventions.
These results have practical implications for healthcare professionals, caregivers, and parents, as they emphasize the importance of recognizing and responding to specific crying patterns in infants. By identifying and addressing the underlying causes of infant distress, appropriate interventions can be implemented to promote the well-being and overall development of the child.

It is important to note that this study represents an initial exploration into the relationship between crying patterns and various conditions in infants. Further research is needed to validate and expand upon these findings, as well as to explore additional factors that may influence crying patterns. Nonetheless, the study provides valuable insights into the potential of cry analysis as a diagnostic tool and highlights the importance of early detection and intervention in promoting optimal outcomes for infants.

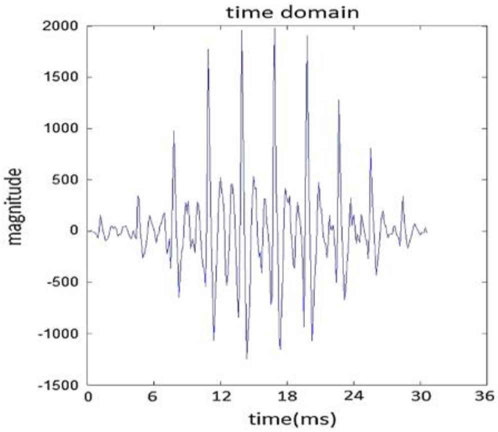

Figure 7. Spectrum of the cry signal in time domain

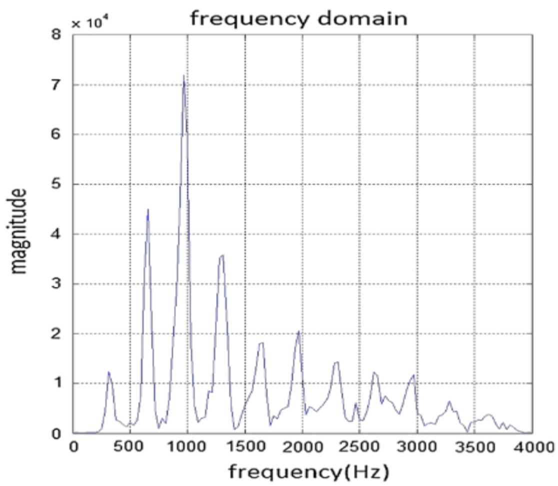

Figure 8.Spectrum of the cry signal after applying FFT

5.3-**Crying output Signals(Due to Pain)**

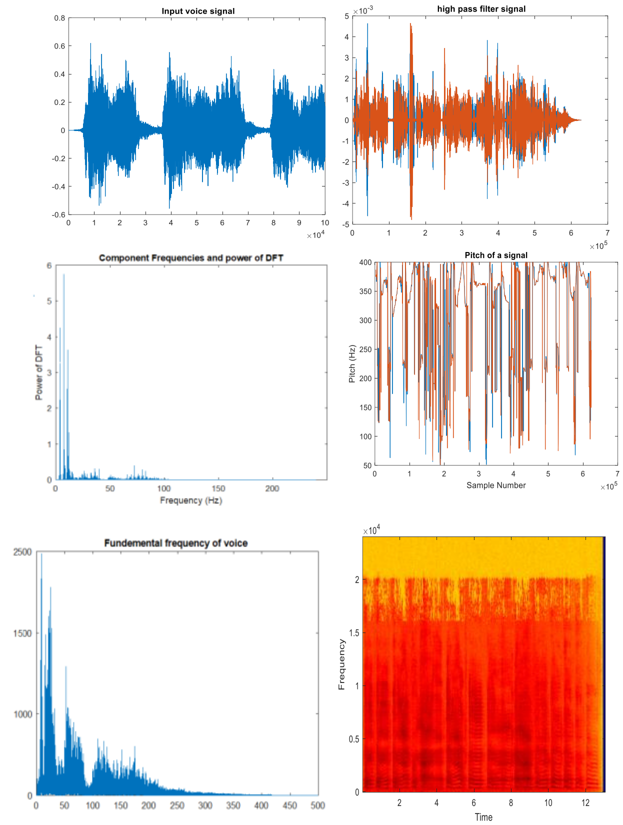

Figure 9.Analyse of crying infant due to disease(Autism)

5.3-**Crying output Signals(Due to Hunger)**

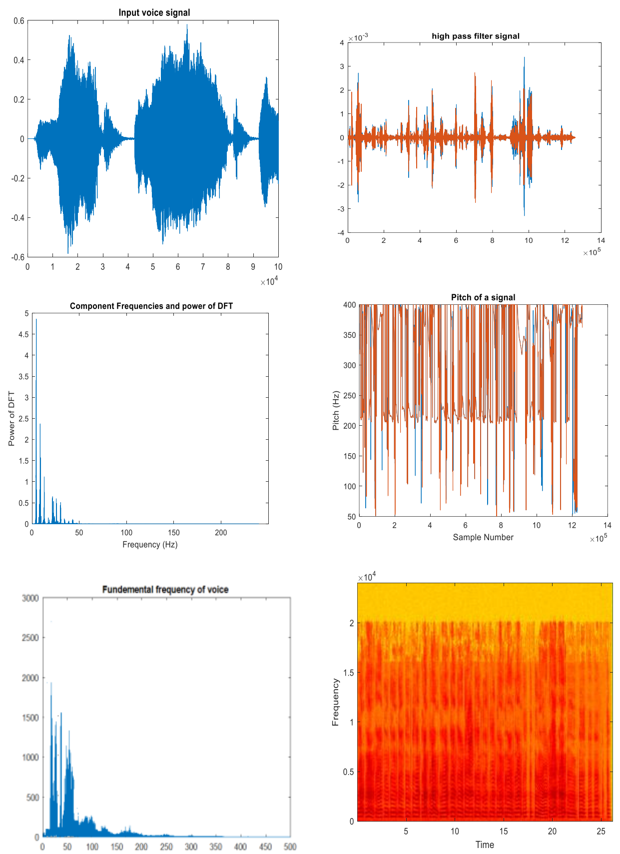

Figure10.Analyse of crying infant due to Hunger.

Table 2 below lists the accuracy of cry detection for upset (due to pain), hunger, fear, and happy emotions.

Feeling	**Training Data**	**Testing Data**	**Total cries**	**Accuracy**
Pain Autism	60cries	60cries	120cries	89.935%
Pain Epilepsy	80cries	80cries	160cries	94.56%
Pain cry	100cries	100cries	200cries	95%
Hunger cry	161cries	161cries	322cries	96%
Fear cry	76cries	76cries	152cries	76%
Happy cry	111cries	111cries	222cries	87%
Cold cry	65cries	65cries	130cries	85%

Table 2: Accuracy of the recognition rates of baby's crying using a continuous cry

In addition, Tab. 3 presents baby cry-speech detection with pitch and deviation for the different frequency ranges.

Cry	**Frequency**	**Pitch**	**Deviation**
Pain Autism	1110 to 1250 hz	1240	1350
Pain Epilepsy	1260 to 1450 hz	1432	1510
Pain cry	700 to 1100 hz	1052	1210
Hunger cry	400 to 600 hz	852	1009
Fear cry	50 to 400 hz	490	498
Happy cry	600 to 700 hz	1491	1249
Cold cry	500 to 650 hz	731	840

Table 3: Detection of baby feelings using pitch and deviation in different frequency ranges

Actual	***System Classification***				
	Autism	Epilepsy	Hungry	Sleepy	accuracy
Autism	***85***	***15***	***5***	***1***	***91.3%***
Epilepsy	***11***	***90***	***11***	***5***	***93.9%***
Hungry	***6***	***10***	***130***	***6***	***95%***
Sleepy	***3***	***4***	***10***	***55***	***86.9%***

Table 4: Classification accuracy using the selected 5 features.

	Correct	Incorrect	Total	Accuracy
Male	160	11	173	92.4%
Female	145	9	154	94.15%

Table 5: Classification accuracy for different genders

6. **Conclusions**

Crying serves as the primary means of communication for newborn infants to express their emotions and physical needs. It can be broadly classified into three types: hunger, sleepiness, and a response to environmental stimuli such as pain from injections or tape removal. Each type of cry corresponds to different requirements of the infant, resulting in distinct crying patterns.

In this study, we proposed a system for classifying infant cries. We employed a feature selection technique called Sequential Floating Forward Selection (SFFS) to identify four discriminative features: peak, contrast, pitch, and spectral centroid. These selected features were then used to train a Support Vector Machine (SVM) classifier on an individual subject basis. Additionally, we incorporated the DAG-SVM algorithm for further verification.

The experimental results demonstrate that our proposed method achieves a high classification accuracy(up to 95%) in distinguishing between different cry types. This suggests the effectiveness of our approach in accurately classifying infant cries, which can have potential applications in identifying specific needs or conditions, such as autism or epilepsy, associated with the crying patterns.

REFERENCES:

[1] L. Abou-Abas, C. Tadj, Christian Gargour, Leila Montazeri. " Expiratory and Inspiratory Cries Detection Using Different Signals Decomposition Techniques". *Journal of Voice*, Vol. 31, No. 2, 2017: pp. 259.e13–259.e28

[2] Y. Kheddache, C. Tadj. " Frequential Characterization of Healthy and Pathologic Newborns *Cries* ", *American Journal of Biomedical Engineering* 2013,3(6): pp. 182-193.

[3] Y. Kheddache, " Caractérisation Des Cris Des Nourrissons En Vue Du Diagnostic Précoce De Différentes Pathologies ". Thèse de doctorat ,Université du Québec, Montréal [2014].

[4] Dror Lederman, " Automatic Classification of the Infant Cry ". Thèse de master , Université Ben Gourion du Néguev[2002].

[5] Lester, B. M., et L. L. LaGasse. 2008. "*Crying* ". In Encyclopedia of Infant and Early Childhood Development, sous la dir. de Haith, Marshall M., et Janette B. Benson. San Diego : Academic Press, pp. 332-343.

[6] L. Abou-Abas, H. F. Alaei and C. Tadj. " *Segmentation of Voiced Newborns Cry Sounds Using Wavelet Packet Based Features* ", *Proceeding of the IEEE 28th*, pp. 796- 800.

[7] Y. Kheddache, C. Tadj. " Acoustic measures of the cry characteristics of healthy newborns and newborns with pathologies ". *J. Biomedical Science and Engineering*, 2013, 6, pp. 796-804.

[8] Gardiner, W. (1838). "*The music of nature"*. Boston, MA; Wilkins & Carter.
[9] Lester, B. M., Corwin, M. J., Sepkoski, C., Seifer, R., Peucker, M., McLaughlin, S., & Golub, H. L. (1991)." Neurobehavioral syndromes in cocaine-exposed newborn infants". *Child Development, 62,* 694–705.
[10] Robb, M. P. & Goberman, A. M. (1997). "Application of an acoustic cry template to evaluate at-risk newborns:Preliminary findings".
[11] Hariharan, M., Chee, L. S., & Yaacob, S. (2012). "Analysis of infant cry through weighted linear prediction cepstral coefficients and probabilistic neural network". *Journal of Medical Systems, 36*(3), 1–7.
[12] Goberman, A. M., & Robb, M. P. (2005). "Acoustic characteristics of crying in infantile laryngomalacia". *Logopedics Phonatrics Vocology, 30*(2), 79–84.
[13]Wu Tianmin, Xu Zhengyuan, "Preliminary analysis of the children's language development record from birth to three year's old", *Journal of psychology*, No.2,pp.153-165,1979
[14] Zhang Yi, Sun Hui, "Research on the third-order nonlinear parameter of bubbly water", JDCTA: *International Journal of Digital Content Technology and its Applications*, Vol. 6, No. 12, pp. 274 -280, 2012
[15]Li Yuming, "1-120 days baby pronunciation " , *Psychological science*,No.5,pp.21-25,1991
[16] Gustafson GE, Green JA, "The importance of fundamental frequency and other acoustic features in cry perception and infant development", *Child Dev*,Vol. 60,pp.772-780, 1989
[17].Esposito, G., Venuti, P., Maestro, S., & Muratori, F. (2017). An exploration of symmetry in early autism spectrum disorders: analysis of lying. Brain Sciences, 7(11), 141.
[18] Bolton, P. F., Carcani-Rathwell, I., Hutton, J., Goode, S., Howlin, P., & Rutter, M. (2011). Epilepsy in autism: features and correlates. The British Journal of Psychiatry, 198(4), 289-294.
[19].Takahashi, H., Nakao, M., Kaga, M., & Hashimoto, R. (2020). Machine learning-based identification of autism using a spectrogram of naturalistic crying. Scientific Reports, 10(1), 17832.
[21]] Y. Kheddache, " Caractérisation Des Cris Des Nourrissons En Vue Du Diagnostic Précoce De Différentes Pathologies ". Thèse de doctorat ,Université du Québec, Montréal [2014].
[20]Y. Kheddache, C. Tadj. " Acoustic measures of the cry characteristics of healthy newborns and newborns with pathologies ". *J. Biomedical Science and Engineering*, 2013, 6, pp. 796-804.
[21] John G.Proakis, Dimitris G.Manolakis, "Digital Signal Processing, Principles, Algorithms, and Applications", 4th edition,Pearson Education inc., Upper Saddle River. *Biology of the Neonate, 71,* 131–136.
[22] M. Frigo and S.G. Johnson. "The design and implementation of FFTW3". *Proc.IEEE,*93 (2):216–231, February 2005. doi: 10.1109/JPROC.2004.840301.
[23] R.Boite, H.Boulard, T.Dutoit, J.Hancq & H.Leich, "Traitement de la parole ". Book, Collection Electricité, *Presses Polytechniques et Universitaires Romandes*, 1999.
[24] Theodoros Theodorou1 , Iosif Mporas1,2 , Nikos Fakotakis1 " An Overview of Automatic Audio Segmentation" I.J. Information Technology and Computer Science, 2014, 11, 1-9 Published Online October 2014 in MECS (http://www.mecs-press.org/) DOI: 10.5815/ijitcs.2014.11.01
[25] T. Butko, C. Nadeu (2011). "Audio Segmentation of Broadcast News in the Albayzin-2010 Evaluation: Overview, Results and Discussion", EURASIP Journal on Audio, Speech and Music Processing 2011, volume 2011 issue 1
[26] R. G. Dandage and P. P. R. Badadapure, "Infant ' s Cry Detection Using Linear Frequency Cepstrum Coefficients," pp. 5377–5383, 2017
[27] Hartarto junaiedi, Infant's cry sound classification using Mel-Frequency Cepstrum Coefficients feature extraction and Backpropagation Neural Network
[28] Chunyan Ji, Thosini Bamunu Mudiyanselage, Yutong Gao and Yi Pan," A review of infant cry analysis and classification".
[29] Hartarto junaiedi, Infant's cry sound classification using Mel-Frequency Cepstrum Coefficients feature extraction and Backpropagation Neural Network . Ji *et al. EURASIP Journal on Audio, Speech, andMusic Processing* (2021) 2021:8 https://doi.org/10.1186/s13636-021-00197-5
[30] C. Y. Chang, Y. C. Hsiao, S. T. Chen, in *Proceedings - 2015 18th International Conference on Network-Based Information Systems, NBiS 2015*. Application of incremental SVM learning for infant cries recognition, (2015), pp. 607–610. https://doi.org/10.1109/NBiS.2015.90.